Neural correlates of dynamic object recognition

Katja Martina Mayer

Institute of Neuroscience, Newcastle University

Thesis submitted for the degree Doctor of Philosophy

September 2010

Table of contents

Acknowledgements

I would like to acknowledge the people without whom I wouldn’t have been able to finish this PhD.

First and foremost I would like to thank my supervisor Dr. Quoc C. Vuong. His support made it possible for me to take on such an ambitious PhD project. Quoc, I can’t thank you enough for your patience when teaching me, your reliability, and for including me in the research on biological motion perception. My gratitude is endless for all you have taught me and for preparing me for my future career in science.

Next, I would like to thank my second supervisor Dr. Yoav Tadmor. Yoav, thank you so much for taking the time to give me advice when I didn’t know how to continue. I wouldn’t have been able to succeed without your help.

As well as Dr. Yoav Tadmor, I would like to thank Dr. Gabi Jordan, Dr. Mike Cox, and Dr. Debbie Riby for the opportunity to demonstrate in their practicals.

Many thanks also to Dr. Michael Firbank, Louise Morris, Carol Smith, and Tim Hodgson from the NMRC for help with the acquisition of MRI data.

I would also like to thank my assessors Dr. Melissa Bateson and Prof. Alex Thiele for their invaluable input to my projects at all stages of my PhD.

Many thanks to Dr. Gesa Feenders for proofreading of earlier versions of this thesis.

Finally, I would like to thank my family and my friends all over the world for being so supportive. Special thanks to Eirini, Milena, Thomas P., Karolina, Bartolo, Mehdi, Willemjin, Ben, Harbaljit, Felipe, Jose, Ilaria, Hannah, Thomas C., Ian, Patrick, Gesa, Alwin, Sascha, Freya, Tascha, and Debborah. You made the 3 years in Newcastle unforgettable. I can’t believe how lucky I was to meet you guys here. I couldn’t have wished for nicer, more supportive, and more fun colleagues than you! I’m going to miss you a lot.

Katja

List of Figures

List of Tables

Abbreviations

2D	two-dimensional
3D	three-dimensional
ANOVA	analysis of variances
AR	autoregressive model
BEDPOSTX	bayesian estimation of diffusion parameters obtained using sampling techniques of crossing fibers
BET	brain extraction tool
BOLD	blood oxygenation level dependent
CingC	cingulate cortex
CoS	collateral sulcus
Cun	cuneus
DTI	diffusion tensor imaging
EPI	echo planar image
FA	fractional anisotropy
FDR	false discovery rate
FDT	FMRIB’s diffusion toolbox
FFA	fusiform face area
FFG	fusiform gyrus
FLIRT	FMRIB’s linear image registration tool
fMRI	functional magnetic resonance imaging
FMRIB	Analysis Group, FMRIB, Oxford, UK
FOV	field of view
FSL	FMRIB software library
FWHM	full width half maximum
GLM	general linear model
hMT/V5	human homologue of the monkey’s MT
HRF	hemodynamic response function
IFG	inferior frontal gyrus
IFOF	inferior fronto-occipital fasciculus
ILF	inferior longitudinal fasciculus
IOG	inferior occipital gyrus
IPL	inferior parietal lobule
IPS	intraparietal sulcus

IT	inferior temporal
ITG	inferior temporal gyrus
ITS	inferior temporal sulcus
LF	lateral frontal
LG	lingual gyrus
LO	lateral occipital
LOa	lateral occipital anterior
LOC	lateral occipital complex
LOS	lateral occipital sulcus
M	mean
MeFG	medial frontal gyrus
MFG	middle frontal gyrus
MNI	Montreal Neurological Institute stereotaxic coordinate system
MOG	middle occipital gyrus
MSE	mean square error
MT	middle temporal
MTG	middle temporal gyrus
OT	occipito-temporal
PCG	post central gyrus
PET	positron emission tomography
pFs	posterior fusiform
PG	parahippocampal gyrus
PPA	parahippocampal place area
PPI	psychophysiological interaction
PreCun	precuneus
PROBTRACKX	probabilistic tracking of crossing fibres
RBC	recognition-by-components
ROI	region of interest
RT	recognition time
SD	standard deviation
SEM	standard error of the mean
SFM	structure from motion
SFS	superior frontal sulcus
SLF	superior longitudinal fasciculus
SLO	superior lateral occipital

SMG	supramargnial gyrus
SPL	superior parietal lobule
SPM	statistic parametric mapping
STS	superior temporal sulcus
TE	time to echo
TR	time to repetition
V1 to V8	early visual areas

1. Abstract

This thesis investigates how colourful dynamic objects are represented in the human brain. The two main theories of object recognition (structural description models and image-based models) make different predictions about which object features (e.g., shape, motion, and colour) are critical for recognition and how individual features can be selectively attended for further processing.

To investigate the relevance of different features for object recognition two sets of novel 3D objects which had different combinations of shape, colour, and motion were created. Selective attention paradigms in which participants attended to single object features were used. The effects of changing unattended features on behavioural performance (Experiment 1 to 3) and large-scale brain responses (Experiments 4 and 5) were measured.

The experiments in this thesis used psychophysical experiments combined with functional magnetic resonance imaging (fMRI) and diffusion tensor imaging (DTI). In addition, the relevance of structural and functional connections between brain areas involved in object recognition was investigated.

The results showed that the representations of shape and motion are closely linked while colour seemed to be processed more independently of other features. At the neural level, shape and motion activated a common network consisting of occipito-temporal, lateral-frontal, and parietal areas whereas colour activated ventral occipital areas. Further support for a common network involved in shape and motion processing was found in structural connections between these areas. For example, the integrity of white matter tracts connecting the occipito-temporal and the lateral frontal areas was correlated with behavioural performance. Although colour activated a distinct set of brain areas compared to shape and motion, activity in colour sensitive brain areas modulated the activity in shape sensitive areas which suggests integration processes of multiple object features.

The results of this thesis suggest a synthesis of elements of both of the two main object recognition theories rather than favouring one of them.

2. Summary of previous research on recognition of multi-featured objects

2.1. Motivation for research in object recognition

Survival would hardly be possible without the ability to recognise objects in our environment. For example, we need to discriminate edible from non-edible foods, recognise our friends from strangers, or identify personal belongings amongst other items. The importance of object recognition in everyday life makes this cognitive process one of the most interesting cognitive processes to study. The amazing flexibility of the recognition system to recognise objects under changing conditions is what makes it an exciting research field as there are still many open questions on how recognition works (Tarr and Bulthoff, 1998).

One of the challenges faced by observers is to be able to recognise the same object despite changes in its retinal projection – a problem often referred to as object constancy (e.g., Marr, 1982). Many factors can contribute to these changes. First, objects can be encountered under different viewing conditions, such as different viewing distance, perspective, or illumination. Second, the relative position between observer and object can change continuously (as when we walk around an object). Lastly, objects can be partially occluded or some features might be more difficult to recover than others under different viewing conditions. Overall, these viewing conditions can dramatically change how the same objects appear to observers, i.e., how visual information projects onto the retinas, and how differences in retinal projections affect the information that is available for object recognition.

How do observers achieve object constancy in a dynamic environment? That is, how might the human brain extract useful visual information about object identity from the dynamic visual input and represent that information so that it can subsequently be used to recognise objects? The guiding hypothesis for this thesis is that there is a large scale network of cortical regions that work together for object recognition (e.g.,Schultz et al., 2008). The different regions within the network process different features (such as colour, shape, or motion) or the conjunction of features, and their activity can be influenced by attention (e.g., Corbetta et al., 1990; 1991; Kanwisher and Wojciulik, 2000; Murray and Wojciulik, 2004; Paradis et al., 2008; Peuskens et al., 2004). This combination of large scale neural networks and attentional modulation of these networks allows observers to recognise objects in a dynamic and complex visual environment.

The experiments in this thesis investigated how colourful dynamic objects are represented by large scale networks in the brain, how attention can modulate these neural representations, and how these representations might ultimately impact on behavioural performance. Shape, colour, and motion have been prominent features studied with respect to both object recognition and attention (e.g., Newell et al., 2004; Paradis et al., 2008; Peuskens et al., 2004). To address these questions and test the guiding hypothesis this thesis combined object recognition and selective attention paradigms with neuroimaging methods.

2.2. The role of shape, colour, and motion in object recognition

The objects we encounter in everyday life are rarely defined by only a single feature. Real-world objects usually consist of a multitude of features such as three-dimensional (3D) shape, surface properties such as colour and texture, and possibly dynamic features (e.g., movements). Any of these features could, in principle, be used to recognise objects. Several researchers argue that observers predominately rely on both 3D and two-dimensional (2D) shape information for recognition (e.g., Biederman, 1987; Biederman and Gerhardstein, 1993; 1995; Marr and Nishihara, 1978; Tarr, 1995; Tarr and Bulthoff, 1995; 1998). Shape is usually the dominant cue because recognition is often carried out at a basic level for many real world objects (e.g., we often recognise an object at a basic level 'dog' as opposed to a more subordinate 'dachshund' or more superordinate 'animal') (Rosch et al., 1976). Moreover, different features such as texture patterns, stereo disparity, shading, and rigid motion are predominantly used to estimate the shape of an object (see Todd et al., 2001, for a review). For instance, humans are able to extract shape from rigid rotation in depth (Ullman, 1979). This ability can be demonstrated with structure-from-motion displays (SFM) in which only dots on the surface of a 3D object are visible. SFM displays appear to be a random 2D spatial distribution of dots when the object is static. When the object moves, however, the 2D motion of the dots induces the percept of the 3D object in the observer. Extraction of shape from SFM displays is possible for simple shapes such as spheres (e.g., Paradis et al., 2000) and even complex shapes such as human faces (Kriegeskorte et al., 2003).

Despite the prominence of shape information for object recognition, other visual features, such as colour and motion, can be used directly for object recognition, rather than only being use to derive shape information. In particular, object features other than shape become relevant for the recognition process when shape does not provide sufficient information about object identity (Vuong and Tarr, 2006). This can be when differentiating between objects that belong to one category (e.g., pens mostly have the same shape and colour can be used to distinguish between them) or when shape information is not sufficiently available (e.g., when objects are occluded).

Several studies have been conducted to define the role of colour for object recognition. The results for whether colour is an important feature to recognise objects were mixed (see Tanaka et al., 2001, for a review). On the one hand, several studies have shown that colour affected object recognition (e.g., Naor-Raz et al., 2003; Price and Humphreys, 1989; Rossion and Pourtois, 2004; Tanaka and Presnell, 1999; Wurm et al., 1993) particularly when colour is a diagnostic feature of that object (e.g., bananas have a diagnostic yellow colour). Price and Humphreys (1989), for example, found that categorisation and naming of real-world objects were facilitated when the objects were presented in the colour they naturally appear in as opposed to objects presented in unexpected colours. Similar results were found by other studies. In their study, Naor-Raz et al. (2003) presented participants with objects with diagnostic colours (such as bananas) in typical (yellow) and atypical (purple) colours. They found that colour naming times were longer for atypical colours which indicated that colour was part of the object representation that observers had formed based on experience with the object. Moreover, Hayward and Williams (2000) showed that colour can increase discriminability between objects and therefore facilitate recognition.

Colour can also contribute to object recognition when shape information is not fully available. For objects with high colour diagnosticity, Tanaka and Presnell (1999) found that recognition performance was less impaired when shape information was degraded and colour information was preserved for objects with high colour diagnosticity compared to objects with low colour diagnosticity. Extending this finding, Rossion and Pourtois (2004) found that even for man-made objects without a particular diagnostic colour, the presence of colour information facilitated object recognition to a larger extent compared to the facilitation due to the presence of texture.

The role of colour may depend on task. For example, Davidoff and Ostergaard (1988), showed that object colour facilitated object naming but not object categorisation performance. Moreover, observers with impaired vision were found to strongly rely on surface information such as colour and texture to identify objects (Cavina-Pratesi et al., 2010; Milner et al., 1991; Wurm et al., 1993).

In contrast to these studies, Biederman and Ju (1988) found no effect of colour for the identification of real-world objects.

Other than shape and colour, observers are able to use dynamic features such as rigid rotation (e.g., the direction in which an object rotates) for object recognition (Liu and Cooper, 2003; Stone, 1998; 1999; Vuong and Tarr, 2004; 2006). Using different object types and experimental paradigms, these studies found that when the rotation direction was reversed for learning and recognising of the objects recognition performance decreased. Rigid rotation reversal preserves the 3D structure of the object as well as the 2D ‘snapshot’ views of the rotating object that fall onto the retina. This means that motion information of the objects was included in the representation that observers formed of the object.

Another type of motion that observers can use for object recognition are semi-rigid articulations of object parts (Setti and Newell, 2010; Vuong et al., 2009; see Aggarwal et al., 1994, for a review). Setti and Newell (2010) used novel objects consisting of multiple parts to investigate whether there were differences between global and local motion in the extent to which they were used to recognise objects. Global motion referred to rigid motion of the object as a whole (e.g., translation or rotation) and local motion referred to rigid motion of the objects’ parts. Overall, they found that both local and global motion are used as cues to object identity. In contrast to global motion, however, the role of local motion depended on its association to global motion and the task. Similarly, Vuong et al. (2009) found that observers were sensitive to the similarity of the local part motion across different viewpoints.

The role of dynamic information is also important for recognising biological objects in the environment. One prominent example is the recognition of biological motion. Biological motion refers to any motion that is produced by animals or humans such as running, climbing, scratching, waving, and many more. Johansson (1973) used stimuli in which the information about the human body was reduced to point lights that were attached to the joints of a walking human. He found that human observers were able to quickly identify a human as soon as the point lights started moving and therefore reflected walking motion. This finding highlighted the relevance of motion for making sense of the natural environment.

Similar results were found for faces (see O'Toole et al., 2002, for a review). Both rigid motion such as head motion and non-rigid motion such as facial expressions have been found to be used for recognition. Facial expressions are motions that change the overall shape of the face. Knight and Johnston (1997) showed that under less optimal conditions for recognition (negatives of face videos were presented) participants showed better recognition performance for faces that moved as if the face was talking as opposed to static images. Consistent with this finding, Lander et al. (1999) demonstrated that facial and head motion helped observers access mental representations of faces. They presented their subjects with video clips of famous people whose faces showed non-rigid (speaking and expressing emotion) and rigid (head turning) motion. These findings were not due to the fact that motion provided further information about the static features of the face (i.e., the shape) but rather due to the fact that observers had encoded the dynamic features of the face (Lander and Bruce, 2000; Lander et al., 1999). The finding that observers combined static information with dynamic information for face recognition was confirmed by another study (Knappmeyer et al., 2003). In their study, Knappmeyer et al. (2003) presented participants with morphed faces. They found that identity judgments were biased by the non-rigid motion that the face showed. In line with these findings, Pilz et al. (2006) showed that non-rigid facial motion facilitated visual search. Participants were asked to search an array of static faces for a specific static target face. When participants were familiarised with the target face while it moved non-rigidly search was quicker than when they were familiarised with the target face when it was static.

The studies on facial motion show that human observers can use non-rigid motion as a cue for identification. There is evidence that the advantage of non-rigid motion for recognition is not limited to biological objects. Pyles et al. (2007; see also Pyles and Grossman, 2009) used ‘creatures’, a set of objects whose motions were adopted from biological motion while the shape was novel, to investigate whether they were processed similarly as biological targets (see also Jastorff et al., 2006; 2009). They found that participants were able to detect these creatures in point light displays with distractor point lights. However, the maximum number of distractor point lights that still allowed for detection of the creature was significantly lower compared to human point light figures. This shows that non-rigid motion can facilitate processing of objects even for novel objects.

The studies reviewed above only addressed how one or two features interact for object recognition. However, real-world objects are usually defined by many different features. Newell et al. (2004) showed that multiple object features (shape, colour, and motion) can be important cues for object classification. Their objects consisted of a shape, a colour, an intrinsic, and an extrinsic motion feature. The intrinsic motion feature was rotation in depth about an axis of the object or swinging motion and the extrinsic motion feature was the path the object moved along in space. Their results indicate that both static (shape and colour) and dynamic features are integrated in the representation that participants formed of the objects.

2.3. Theories of object recognition

Different theories have been developed to date to explain how object recognition is achieved (see Logothetis and Sheinberg, 1996; Pinker, 1984; Tarr and Pinker, 1989). Most theories have focused on the recovery of shape. These theories can be classified in two categories: structural description models and image-based models. Of main interest for the studies in this thesis are the assumptions that the theories make on how object features apart from shape are used for recognition.

2.3.1. Structural description models

Structural description models assumed that object recognition is achieved by recovering shape in a viewpoint-independent manner in which knowledge of the shape is stored in the brain and the shapes of other objects are compared to the stored shape (Biederman, 1987; Marr and Nishihara, 1978).

In Biederman's (1987) recognition-by-components (RBC) model, the mental representation of objects can be decomposed into primitives referred to as 'geons' and the spatial relationship between geons (e.g., 'on top of'; 'left of'). Geons are defined on the basis of non-accidental properties such curvature, collinearity, symmetry, parallelism, and cotermination. Biederman (1987) reported that 36 geons can be identified based on non-accidental properties. A geon is for example a brick or a cylinder. A mug can be represented as a cylinder with an arch attached to its side.

According to RBC, recognition of an object starts with edge detection. The object is then parsed into smaller components (geons) on the basis of non-accidental properties of the object. The arrangement of the components is important because changes in the spatial relation between components can indicate different objects. Once the components of the object are identified the arrangement of components is then compared to stored representations in memory. Importantly, RBC assumes that geons can be recognised in a viewpoint-independent manner. This means that independent of how a brick is presented in space or from which viewing angle an observer might perceive it; it can always be recognised as a brick.

Experimental investigation (Biederman and Bar, 1999; Biederman and Gerhardstein, 1993; 1995) showed that observers can use non-accidental properties to recognise objects that they had learned at a different viewpoint.

At the neural level, Hayworth and Biederman (2006) found that the lateral occipital complex (LOC, an area known to be involved in shape processing, see Chapter 2.4.1.1) was sensitive to parts of objects. In an fMR-adaptation study Hayworth and Biederman (2006) presented participants with line drawings of objects such as airplanes of which line parts were deleted. fMR-adaptation is the neural mechanism of decreased response to repeatedly presented stimuli (see Chapter 5.2). They found that adaptation effects were found when parts of the objects (such as a wing of an airplane) were presented repeatedly. This result suggests that the LOC represents object parts providing neural evidence for the relevance of basic shapes such as geons for object processing.

In summary, RBC postulates that observers rely mainly on shape when recognising objects. RCB and other structural description models discard information about non-shape features such as colour and motion, using them only to help recover shape representations that are invariant to viewing conditions. Other features have a secondary role, for instance, colour and motion can be used to extract non-accidental properties. Thus, when edges are easily extracted from the retinal image, then the addition of other features will not facilitate recognition performance. This hypothesis was supported by Biederman and Ju (1988), who showed that object naming times were identical for full-coloured images of objects and line drawings of objects. This showed that the presence of colour in the stimuli did not facilitate the object recognition process. According to structural description theories, non-shape object features, such as colour and motion, would be used only under difficult recognition condition such as when parts of the objects are occluded or when the objects to be discriminated belong to the same category.

2.3.2. Image-based models

In contrast to structural description models image-based models assume that objects are encoded under specific viewing conditions, such as perspective viewpoint, viewing distance, and illumination (Tarr, 1995; Tarr and Bulthoff, 1995). Image-based models assume that the representation of an object is composed of multiple images (snapshots) of the object in the perspective that the observer had encountered the object before. Recognition is achieved by matching the percept of an object with stored representations. Consequently, recognition performance is best when an object is perceived under the same viewing conditions as when it was first encountered. Generalising from familiar to novel viewing conditions requires additional cognitive processes, such as mental rotation (e.g., Tarr and Pinker, 1989), view interpolation (Edelman and Bulthoff, 1992; Poggio and Edelman, 1990), or processes to achieve object constancy across different lighting conditions (Tarr et al., 1998). Image-based models can explain why recognition performance declines when objects are presented from unfamiliar or non-canonical views (Edelman and Bulthoff, 1992).

In contrast to structural description models, image-based models allow for non-shape features to be incorporated into the object representation. Image-based models postulate that the representation of objects consists of 2D snapshots of the objects as they were encountered. This means that any available visual feature is included in the representation and can therefore affect recognition.

2.3.3. Extensions to image-based models

To extend image-based models Giese and Poggio (2003) built a biologically motivated model and applied it to the recognition of biological motion (e.g., running). Their model is based on anatomical and physiological evidence and extends an earlier image-based model (Riesenhuber and Poggio, 1999). The main assumption is that motion is represented as a series of snapshots that the observer had learned in his/her experience with biological motion. Processing of biological motion happens along a form and a motion pathway. Along both pathways the complexity of the processed features increases. The form pathway analyses shape in the snapshots of the biological target. The motion pathway processes different types of optic flow. For successful recognition of the motion it is essential that the temporal sequence of the snapshots stays in the right order. Altogether, this quantitative model takes neurophysiological evidence into account to describe the recognition of biological motion. Potentially, this model could be used to explain other data such as motion of non-biological objects with similar types of motion such as non-rigid deformation.

Further, Walther and Koch (2007) describe a model that unifies many of the ideas of other models. As with other models (e.g., Giese and Poggio, 2003) Walther and Koch’s (2007) model starts with cells that are sensitive to orientation. Subsequently, the receptive field size increases and the cells become more and more specialised up to cells that encode the critical object such as a certain dog. Importantly, this model differentiates between tasks that require object identification and tasks that require object categorization. This is achieved by taking top-down attentional mechanisms into account. In summary, Walther and Koch (2007) model the processes of object recognition and combine those with processes of visual attention. Therefore, their model is capable of explaining how attention can modulate object recognition.

Besides attentional influences on object recognition temporal aspects can have an impact. In several computational studies Wallis (1996; 1998; 2002) showed that the temporal association between the encoded snapshot images of objects is crucial for recognition. Behaviourally, Wallis and Bulthoff (2001) showed that an image sequence of rotating faces that changed identity throughout the rotation was treated as if the images belonged to the same face.

Wallraven and Bulthoff (2001) developed another model to formalise the processes underlying the recognition of objects and faces. Their model describes recognition as a process of feature tracking. Only key frames of important features of a sequence of available snapshots of a face are encoded providing an efficient way of representing face stimuli for subsequent recognition.

2.4. Object recognition in the brain

In the human and the non-human primate brain, visual information is processed along two specialised pathways (Mishkin et al., 1983). The ventral pathway is the occipito-temporal projection. It follows the white matter tract of the inferior longitudinal fasciculus (ILF) and connects occipital areas to inferior temporal areas. The dorsal pathway is the occipito-parietal projection. It follows the tract of the superior longitudinal fasciculus (SLF) and connects occipital areas to inferior parietal areas.

Each pathway connects specific subsets of visual brain areas. Both pathways originate in the primary visual cortex (V1) and project to the secondary visual cortex (V2). The dorsal pathway then projects to the motion sensitive area middle temporal (MT) and subsequently to the posterior parietal cortex. The ventral pathway projects from V2 to the orientation and colour sensitive V4 and finally to anterior inferior parts of the temporal cortex (van Essen and Gallant, 1994).

Along both pathways, neurons have small receptive fields that are responsive to very basic features such as orientation. The receptive fields of neurons progressively increase in size and they respond to more and more complex configurations of features further down both pathways (see Giese and Poggio, 2003, for a description).

The ventral and dorsal pathways generally carry out different functions, but there are important interactions between them for object recognition. The function of the ventral pathway is the processing of object identity while the dorsal pathway is involved in creating spatial maps and the usage of visual information for motor actions (see Mishkin et al., 1983). Although the ventral pathway is essential for identifying objects (see Mishkin et al., 1983), input from the dorsal pathway can be relevant for identifying objects as well (Konen and Kastner, 2008). Indeed, as reviewed in the following sections, several cortical regions along both pathways contribute to the recognition of colourful dynamic objects (Peuskens et al., 2004; Schultz et al., 2008).

2.4.1. Brain areas involved in object recognition

A large number of experiments have investigated the neural correlates of perception and recognition of objects. Studies have ranged from non-human primate work to brain lesion studies in humans and animals, to more recent neuroimaging techniques. The general idea is that different anatomically segregated regions of the brain are particularly sensitive to different types of information such as shape or motion. However, although a brain region may respond most to a given feature it can respond to other features or combination of features as well. Further, specific features or visual input as complex as objects rarely activates only single brain areas. Most commonly, whole networks of brain areas across the cortex are activated. Within such networks there are both functional (e.g., Friston et al., 1997; Haynes et al., 2005; Macaluso et al., 2000; Schultz et al., 2008) and anatomical connections between areas (e.g., Felleman and van Essen, 1991).

The next sections summarise brain areas that respond to different object features such as shape or colour and brain areas that integrate different features. In addition, networks that are likely to be involved in object recognition are described.

2.4.1.1. Shape sensitive areas

Across different theoretical approaches on how object recognition is achieved it is agreed that shape is the most important feature for object recognition (Biederman, 1987; Tarr and Bulthoff, 1998). This section provides an overview of studies that have helped identify different brain regions involved in processing shape.

Research into non-human primates identified cortical areas that are crucial for shape processing. Dean (1976) reported that lesions in the macaque monkey's inferior temporal (IT) cortex led to impairments in the ability to categorise visual stimuli and in forming associations between them. These results suggested that IT was important for processing of objects. In an electrophysiological study, Logothetis et al. (1995) showed that a set of IT neurons was sensitive to viewpoint changes of unfamiliar objects. The firing rate of these neurons was maximal for one view and decreased when the object was rotated away from the preferred view. They inferred that the viewpoint sensitivity of IT neurons indicated their involvement in recognition of 3D shape.

Initially, positron emission tomography (PET) studies were conducted to identify regions that were activated during object recognition tasks in humans. Motivated by the finding that the visual areas in non-human primates are organised in two pathways (Mishkin et al., 1983) these studies aimed to find homologues in the functional organisation of the human visual cortex. These studies found that mostly temporal and occipital regions of the human cortex were involved in visual object processing (e.g., Corbetta et al., 1990; 1991; Kosslyn et al., 1994; Sergent et al., 1992).

To specify the exact location of the brain areas processing objects, Malach et al. (1995) conducted an fMRI study in which subjects viewed a variety of abstract and real-world objects including faces. They located a brain area in the occipital cortex that was activated more strongly when subjects viewed different types of objects as opposed to textures. As it was unclear whether this region reflected one functional area or a cluster of different functional areas it was termed lateral occpitital complex. Anatomically, they located LOC in the lateral posterior part of the occipital cortex. In the same study, Malach et al. (1995) investigated the localisation of LOC relative to hMT/V5, a brain area involved in processing of motion (see Chapter 2.4.1.2). LOC was located adjacent and posterior to hMT/V5. Further, they found that LOC activity was independent of whether abstract or real-world objects were presented. Importantly, the preference of LOC for objects compared to textures could not be explained by differences in spatial frequencies or contrasts between both types of stimuli (Malach et al., 1995).

Several functional imaging studies that used different types of stimuli found similar results as Malach et al. (1995). Kanwisher et al. (1996), for example, conducted fMRI experiments using different types of line drawings of objects and faces. They showed that an inferior lateral area at the occipito-temporal junction was involved in recognising shapes. Further, their experiments provide evidence that objects and faces are processed by different cortical areas.

In their study, Kraut et al. (1997) presented objects of which only shape information was present. Objects were line drawings of, for example, an airplane while so-called non-objects were line drawings of an airplane on which one wing was replaced by a screwdriver. They found that objects evoked larger activation in occipital cortices and occipito-temporal gyri as opposed to non-objects. Consistently, different studies provided converging evidence that a large portion of the junction of the anterior lateral occipital cortex and the posterior temporal cortex are sensitive to the perception of shape (see Grill-Spector et al., 2001, for a review).

More recent studies attempted to identify subregions within LOC. Grill-Spector et al. (1999) suggest that LOC can be divided into two subregions: a dorsal-caudal region they named lateral occipital (LO) and a ventral-anterior region they named posterior fusiform (pFs) or lateral occipital anterior (LOa). Grill-Spector et al. (1999) found that pFs/LOa showed greater invariance to changes in object position compared to LO. Similarly, Lerner et al. (2001) found a hierarchical organisation within LOC. Stepwise scrambling of pictures of objects and faces led to a decrease in activity in the anterior parts of LOC when compared to intact objects. More posterior parts of LOC were less sensitive to scrambling of pictures of objects (Lerner et al., 2001). Kourtzi et al. (2003) investigated whether the LOC activity to objects differed for the presentation of 2D and 3D shapes. Their results also suggested a subdivision of the LOC. More anterior parts of the LOC seemed to encode 3D shape while more posterior parts responded to 2D features of the shapes.

The studies investigating which brain areas are sensitive to shape have used a wide range of different objects such as cars or tools and presented them in different ways such as line drawings or images (see Grill-Spector et al., 2001, for a review). Another way to present objects is to define the shape by motion. In their study, Murray et al. (2003) aimed at identifying brain areas that responded to line drawings of objects, random dot motion, and SFM. Shape in general activated a set of subregions in the LOC. Only one of these subregions (the superior part of the LOC) responded to both shape defined by line drawings and shape defined by SFM. Additionally, areas in the parietal cortex were involved in processing SFM. Aiming at identifying segregated areas processing shape or motion Paradis et al. (2008) conducted an fMRI study in which they presented participants with 3D SFM objects. They found that shape activated several areas along the ventral pathway.

In addition to the evidence for object selectivity in lateral occipital regions found by PET and fMRI studies, lesion studies support these results. Certain cortical lesions have been found to lead to what is called visual form agnosia (Benson and Greenberg, 1969). These patients lose the ability to recognise objects after cortical lesions while other visual abilities remained preserved to a large extend. The best-known and best-studied case is probably patient D.F. (Cavina-Pratesi et al., 2010; Milner et al., 1991). D.F. suffered from carbon monoxide intoxication. Several months after the accident different neuroimaging methods revealed that amongst other lesions parts of her parieto-occipital cortices and her lateral occipital cortices were damaged (see Milner et al., 1991). After the accident D.F. was not able to visually recognise objects anymore whereas her ability to recognise objects by the means of audition or tactile exploration was unimpaired (Milner et al., 1991).

The case D.F. together with the results of neuroimaging studies with healthy humans provide converging evidence for the importance of lateral occipital cortices for the processing of object shape.

Apart from the well-documented involvement of the LOC in shape processing other brain areas were found to be involved as well. In particular, areas in the parietal cortex can be involved in shape perception. In addition to the occipital and occipito-temporal activation to objects, Kraut et al. (1997) found activity in brain areas that are located along the dorsal pathway, specifically in the superior and the inferior parietal lobules (SPL and IPL). These results provide evidence that both the dorsal and the ventral pathway are involved in visual object recognition. Consistent with this, Murray et al. (2003) showed that an area in the lateral parietal cortex showed stronger response to 3D shapes compared to 2D shapes indicating that the parietal cortex might play a role in processing the 3D shape of objects. Moreover, Peuskens et al. (2004) found that attention to shape activated the ventral and dorsal pathway including areas in parietal cortices. Within the parietal cortex the IPS in particular can be involved in shape recognition (Konen and Kastner, 2008).

Consistent with the results in electrophysiological studies in non-human primates (Murata et al., 2000), studies on human object recognition found that areas along the IPS respond to objects (see Grefkes and Fink, 2005, for a review). In humans, Grefkes et al. (2002) for example, found that the IPS was activated when observers visually and tactilely processed objects.

2.4.1.2. Motion sensitive areas

Several brain areas can be activated when observers view motion (see Culham et al., 2001, for a review). In this section a summary of the brain areas involved in processing visual motion is provided with focus on the activation patterns caused by dynamic objects. Motion is mainly processed along the dorsal pathway (see Mishkin et al., 1983).

A number of brain areas in which cells show sensitivity to motion has been identified using neurophysiological methods. In the hierarchy of visual processing the earliest stage in the primate brain that shows direction sensitivity (and therefore motion sensitivity) are cells in V1 (Hubel and Wiesel, 1977). Direction sensitivity has also been found in V2 and V3 (Felleman and van Essen, 1987). Another motion sensitive brain area was identified in the posterior bank of the superior temporal sulcus (STS) of the monkey (Dubner and Zeki, 1971; Maunsell and van Essen, 1983). The neurons in this region responded to visual motion but were insensitive to form and size changes (Dubner and Zeki, 1971). This area was area MT.

Based on the literature on motion sensitive brain areas in the primate numerous studies were conducted to find homologues for these motion sensitive areas in the human. In a PET study, Zeki et al. (1991) found evidence for functional specialisation of areas of the human visual cortex. They presented participants with random fields of black and white squares in which the squares either moved or were static. They found a bilateral area located at the temporo-occipito-parietal junction that showed increased activity for dynamic compared to static random square fields which was inferred to be the human homologue of the primate MT (Malach et al., 1995; Zeki et al., 1991) and will be referred to as hMT/V5 in this thesis. Watson et al. (1993) reported that apart from the human homologue of MT, the early visual cortex (V1 and V2) and the cuneus showed increased activation when perceiving motion compared to static visual displays. The anatomical location of hMT/V5 was described as the junction of the ascending limb of the inferior temporal sulcus (ITS) and the lateral occipital sulcus (LOS) (Zeki et al., 1991). Later studies replicated hMT/V5 activity in humans when perceiving motion (e.g., Dupont et al., 1994; Watson et al., 1993) and confirmed its anatomical location (Dumoulin et al., 2000). Dupont et al. (1994) found additional activation to visual motion in the cuneus, the cerebellum, the lateral sulcus, the parieto-occipital fissure, parts of the occipital lobe (Brodmann area 19), and parts of the parietal lobes (junction of Brodmann areas 2 and 40).

Tootell et al. (1995) replicated earlier PET studies with fMRI. They used several different MR parameters and consistently found activation in an area corresponding to hMT/V5 that showed stronger activation for random dots moving radially than for static random dot fields. In addition, further brain areas with stronger response to motion compared to static images included areas of the inferior parietal cortex (Tootell et al., 1995). In their fMRI study, Sunaert et al. (1999) found an extensive network of areas that showed increased BOLD activity when observing motion compared to static displays. Apart from hMT/V5 they found that early visual cortices, occipito-temporal, occipito-parietal, parietal, and frontal areas were activated (Sunaert et al., 1999).

Consistently, the hMT/V5 complex was found to respond to motion in healthy subjects. Further support for the importance of hMT/V5 for motion processing comes from a lesion study. Zihl et al. (1983; 1991) reported a case study of a 43 year old female (L.M.) with (among other lesions) bilateral lesions in the posterior part of the middle temporal gyrus after a superior sagittal sinus thrombosis. L.M. reported that she was unable to perceive motion such as a cup filling when pouring tea. She felt insecure in rooms with more than two people because people suddenly appeared in unexpected locations. However, she was able to perceive tactile motion on the skin and she could use auditory cues to hear a train approaching. Other visual functions were preserved. L.M.'s case provides further evidence for the involvement of areas at the occipito-temporal junction in processing of motion.

The results from electrophysiology studies, neuroimaging experiments with healthy humans and lesions studies demonstrate that brain areas in the occipito-temporal junction are involved in processing motion. In addition to that, other parts of the brain can be activated by motion as well. Complex visual stimuli such as 3D dynamic objects were found to activate several parts of the cortex. Paradis et al. (2008) investigated the implications of feature attention on the blood oxygenation level dependent (BOLD) response to SFM objects. They found that, attention to object motion activated temporal areas and the IPL. Wu et al. (2008) asked participants to either attend to the manner or the path of the motion of an animated starfish. When attending to the manner temporal areas were activated while attending to the path along which the starfish moved activated IPL and SPL (Wu et al., 2008). As reported above, Pyles et al. (2007) built a set of creatures that simulated biological motion on non-biological shapes. These creatures activated ventro-temporal areas suggesting that this area is involved in processing motion of novel objects.

2.4.1.3. Colour sensitive areas

As for shape and motion specific areas in the brain have been found to be activated by colour. These brain areas are predominantly located on the ventral surface of the occipital cortex. In non-human primate studies it was shown that the neurons of V4 (Zeki, 1973; 1980) were selective for hue differences. Removal of V4 in macaque monkeys led to impaired hue discrimination along with impairments in the discrimination of patterns and orientations (Heywood and Cowey, 1987). Based on this result it was inferred that V4 is important for colour vision. Later studies, however, questioned the role of V4 as the main cortical colour centre in the monkey (see Gegenfurtner and Kiper, 2003, for a review). Other than in V4 colour sensitive neurons were found in V1 (e.g., Dow and Gouras, 1973) and in V2 (e.g., Baizer et al., 1977). A more recent study that used fMRI (Harada et al., 2009) found colour related brain activation in V1, V2, V3, V4, and IT suggesting that multiple parts of the visual cortex are involved in colour processing. Further, initial assumptions that colour was processed in the ventral pathway only (see Mishkin et al., 1983) were not confirmed by later studies that found residual colour sensitivity in MT neurons (e.g., Thiele et al., 1999). To date, it is assumed that colour processing in the monkey is a process involving neurons in multiple visual areas (see Gegenfurtner and Kiper, 2003).

In humans, the homologue of V4 was located along the collateral sulcus (CoS) a sulcus on the ventral surface of the occipital cortex and the neighbouring gyri; the lingual and fusiform gyrus. The CoS was inferred to be colour sensitive as it responded more strongly to coloured images as opposed to gray-scale versions of the same images (McKeefry and Zeki, 1997). This is consistent with the results by Cant and Goodale (2007) who found that attention to colour and texture of novel objects activated areas along the CoS and additional areas in the inferior occipital gyrus (IOG). The areas responding to colour and texture were distinct from the areas that responded to shape and motion. Further, Cant and Goodale (2007) found larger activation for attention to colour as opposed to texture in V1 and the cuneus.

Another colour sensitive area in the human was located in proximity to the human V4 and referred to as V8 (Hadjikhani et al., 1998). This area was found to be more activated by coloured stimuli than the neighbouring visual areas. Later investigations, however, argued that there was no difference between areas V4 and V8 (Wade et al., 2002). Both V4 and V8 were assumed to belong to the human homologue of V4.

Further support for areas along the CoS as colour processing areas comes from studying lesions that led to achromatopsia (see Zeki, 1990, for a review). Several patients were reported that lost the ability to discriminate hue after bilateral lesions in the inferior part of the occipital cortex (Meadows, 1974; Zeki, 1990).

Similar to hMT/V5, the human homologue of V4 seems to be a complex that can be subdivided. Two subregions were identified for the human homologue of V4: V4 and V4α (Zeki and Bartels, 1999). They reanalysed previous fMRI studies and found that V4 which is located more posterior has topographic organisation in contrast to the more anterior part V4α. Zeki and Bartels (1999) argue that the subdivision of the human V4 can help explaining the diversity of achromatopsia types.

In a recent fMRI study (Cavina-Pratesi et al., 2010), a colour blind patient (M.S.) and D.F. were tested in an object recognition task. They had to identify the odd object in a set of three objects by the texture or by the shape of the object. As expected on the basis of their lesions D.F. was only impaired on the shape task and M.S. was only impaired on the texture task. Importantly, this study showed the clear dissociation of areas processing shape and areas processing surface properties of objects such as texture. Lesions in lateral occipital cortices impaired shape processing while texture processing was preserved. Lesions at the ventral surface of the occipital cortex impaired texture processing but preserved the ability to recognise shape.

2.5. Integrating multiple features in the brain

Although brain areas are often selective to an object feature, certain regions were identified that can integrate multiple features.

The STS is a brain region that can integrate shape and motion of objects. Specifically, this brain area had been found to respond to the motion produced by biological objects such as humans (Grossman et al., 2000). This means that neurons in this brain area respond to a particular type of motion (i.e., biological motion) in combination with a particular type of shape such as the body posture of a human. This finding was replicated with novel objects that rotated in space (Schultz et al., 2008) showing that integration of motion and shape in STS is not limited to integration of biological motion and shape. In their study, task-irrelevant motion direction modulated activity in the left STS when participants performed a same-different task on the shape of novel objects (Schultz et al., 2008). Schultz et al.'s (2008) results support the role of the STS as a brain area that integrates shape and motion of objects.

As reported in Chapter 2.4.1.2, the hMT/V5 complex is a brain area that responds to a wide range of different types of motion. Moreover, this brain area was found to integrate shape and motion of objects. Kourtzi et al. (2002) reported that hMT/V5 shows increased activation to static images given they contain implied motion such as an athlete who is about to throw a disc. This finding indicated that it is not only motion per se that activates hMT/V5 but also the integration of shape and motion information. Consistent with this, Sarkheil et al. (2008) found integration of shape and motion of simple objects in hMT/V5. They conducted a priming study in which they asked participants to identify an object or to judge its rotation direction. On each trial two objects were presented in sequence: first the prime object, then the target object. They measured fMR-adaptation effects (see Chapter 5.2) caused by repeated presentations of features of the prime and the target object. When only the shape of prime and target were identical ventral parts of the occipito-temporal cortex were adapted. When both rotation direction and shape were identical adaptation effects were found in an occipito-temporal area that overlapped with hMT/V5. The adaptation effect found for repeated presentation of shape and rotation direction suggests that anterior parts of hMT/V5 integrate shape and rotation direction of simple objects.

Just as the hMT/V5 complex, LOC was also found to be involved in integrating multiple features of objects. In their fMRI study, Self and Zeki (2005) investigated how objects defined by colour, motion, or a combination of colour and motion were processed. For the condition in which both colour and motion defined the object they found larger activation in the LOC compared to the other two conditions. This indicated that the LOC integrated colour and motion to extract shape information.

2.6. Cortical networks for object recognition

Although distinct brain areas are specialised to process certain aspects of the environment, stimuli as complex as objects are not processed by a single brain area but by large scale networks.

Peuskens et al.'s (2004) study showed how attention can modulate the BOLD response to an object. They used novel objects (deformed spheres) that rotated in depth to investigate the implications of feature attention for the BOLD response elicited by dynamic objects. When participants discriminated the 3D shape of the objects while ignoring their 3D motions and textures, a network consisting of the LOS and parts of the posterior parietal cortex was activated. The activity found in lateral occipital cortices was in proximity of the LOC which is involved in shape processing (Grill-Spector et al., 2001; Malach et al., 1995). In Peuskens et al.'s (2004) study no area was identified that exclusively responded to judgments of 3D motion. This condition activated the same set of brain areas as when participants attended to the 3D shape. When participants performed judgments on the texture of the objects brain areas along the ventral occipito-temporal cortex were activated. Altogether, Peuskens et al.'s (2004) results show that objects can elicit activity in large parts of the cortex. Further, their results suggest that there is a common network processing shape and motion of dynamic objects.

Using a shape discrimination task Schultz et al. (2008) identified a network of brain areas similar to the one that Peuskens et al. (2004) found. In their study, participants performed same-different judgments on pairs of novel objects. They found that a network consisting of lateral occipital, parietal, and frontal areas was involved in processing the 3D shapes of the objects. In particular, they found that parietal and frontal regions were sensitive to the metric variations in the 3D shape of the objects to be discriminated. In contrast, lateral occipital regions were sensitive to the perceived variations in shape (as measured psychophysically). This finding suggests that there was functional segregation within the network, with parietal and frontal regions processing metric changes in shapes while latero-occipital regions processed perceived shape differences. The concept of a network with functionally segregated parts was supported by correlations of the residual BOLD signals between pairs of regions (Haynes et al., 2005; Macaluso et al., 2000). Schultz et al. (2008) found that parietal and frontal regions were highly correlated with each other, as were the bilateral regions in occipital cortex.

Apart from the occipito-temporal brain areas that are well established in processing motion and shape, these studies (Peuskens et al., 2004; Schultz et al., 2008) consistently found that dynamic objects activate brain areas in the parietal cortex. As reported in Chapter 2.4.1.2, Paradis et al. (2008) found that the IPL was involved in processing 3D motion of SFM objects. These findings are consistent with previous research that found that the parietal cortex was activated when recovering SFM (Kriegeskorte et al., 2003; Murray et al., 2003; Paradis et al., 2000; Paradis et al., 2001). Moreover, Murray et al. (2003) found that a parietal area responded more strongly to 3D compared to 2D shapes providing evidence for the involvement of this region in shape processing.

As found by Schultz et al. (2008) frontal brain areas are also involved in object recognition. This is consistent with other studies that found frontal activation on different object processing tasks. Yee et al. (2010) for example used a colour and shape matching task and found that shape trials activated the superior frontal sulcus (SFS) to a larger extend than colour trials. Spencer et al. (2009) found that distinct frontal areas are involved in different object recognition tasks such as object priming and object identification. Moreover, frontal brain areas are known to be involved in cognitive control (Miller, 2000), working memory (Cabeza et al., 2003), and visual attention (Kanwisher and Wojciulik, 2000) all of which are cognitive processes that are necessary for object recognition.

To identify a network for object recognition it is not sufficient to summarise brain areas that are activated when objects are processed. It is also important to investigate the connection between these areas. This can be achieved in two ways. First, functional connections can be measured. These are correlations in the time course of the activity between different areas. Second, structural connections (i.e., the white matter) between brain areas of interest can be reconstructed.

In terms of functional connections numerous studies (Haynes et al., 2005; Macaluso et al., 2000; Schultz et al., 2008) showed correlations that were not driven by the experimental paradigm or stimuli in the BOLD signal between brain areas. This indicated that apart from co-activation of these areas by the task there were further communications between these areas. Moreover, brain areas that are organised in a network can modulate each other's activity. Modulation means that the extent to which a brain area is activated depends on the extent to which a functionally connected area is activated (Friston et al., 1997).

Structural connections are the white matter fibres that connect distant parts of the brain. Several studies that investigated the origin of differences in reaction times between humans (e.g., Bohr et al., 2007; Tuch et al., 2005) showed that differences in the integrity of white matter affected reaction times. From such studies it was inferred that just as certain parts of gray matter are associated with performing certain cognitive tasks (e.g., the LOC for processing of shape) specific fibre tracts in the white matter are associated with task performance as well. This was shown by correlating measurements of integrity of white matter fibres with measurements of behavioural performance such as reaction times and accuracy (e.g., Begre et al., 2007; Bohr et al., 2007; Thomas et al., 2009; Thomas et al., 2008; Tuch et al., 2005).

The integrity of the fibre tracts is usually measured as the fractional anisotropy (FA) value. FA is a ratio of the eigenvalues of the three orthogonal axes along which diffusion can occur. FA is a value between zero and one with zero meaning isotropic diffusion (unrestricted or equally restricted diffusion in all three directions) and one meaning that diffusion only occurs in one direction. A high FA value is related to high integrity of a fibre tract.

Previous studies (Begre et al., 2007; Bohr et al., 2007; Thomas et al., 2009; Thomas et al., 2008; Tuch et al., 2005) have used correlations between behavioural performance and FA value as a means to define parts of the white matter that are involved in certain tasks. The underlying rational is that tracts with high integrity have better myelination and are therefore faster at transmitting signals which could result in faster reaction times (see Tuch et al., 2005). However, some studies investigating the relationship between tract integrity and behavioural performance found that there were positive correlations between the two variables suggesting that other factors than myelination might influence the relationship between tract integrity and behavioural performance (e.g., Tuch et al., 2005; Bohr et al., 2007). In spite of this controversy, research into the dependence of behavioural outcome on white matter integrity showed that there is a relationship between white matter integrity and behavioural performance, making this method a promising approach to understand brain function.

2.7. The role of attention in object recognition

The capacity to process information of the human cognitive system is limited which means that the abundance of information that the environments provides at any given moment poses the potential problem of overload. The ability to extract relevant information from the information flow that an observer has to deal with is what is referred to as attention (see Broadbent, 1958, for example).

Attentional processes allow observers to select information that is potentially important for the task at hand for further processing. Some of these processes operate automatically in a ‘bottom-up’ manner and are typically driven by salient sensory features in the environment (e.g., a sudden loud noise or bright colour in a dim room can automatically draw our attention). Observers can also deploy attention to different locations or objects in the environment in a ‘top-down’ manner (e.g., attend to the colour blue when looking for a blue book) (e.g., Oliva et al., 2003).

The ability to attend to certain features of the environment separately makes attention a very important aspect of object recognition (Hummel and Stankiewicz, 1998; Treisman and Gelade, 1980; Walther and Koch, 2007). Treisman and Gelade’s (1980) influential theory on attentional feature binding postulates that features such as colour and orientation are registered automatically across the visual field and attention is the crucial process needed to bind the correct features together to perceive an object.

As reported above, humans can, in principal, use different features to recognise objects under different viewing conditions and task demands. Although shape tends to be the dominant feature (Tarr and Bulthoff, 1998; Vuong and Tarr, 2006), the human brain is highly flexible and can adapt to different situations. Thus, observers can attend to a particular feature to identify objects. In this regard, attention is important because it enables observers to select features that are the most useful or ‘diagnostic’ for the task at hand, as not all features may be equally available under all viewing conditions (e.g., when object parts are occluded or in dimly lit environments). The ability to attend to relevant features of the whole of visual input allows observers to filter out irrelevant aspects from the extensive stream of information that is available. Depending on the task at hand different features of the object can be relevant. Walther and Koch (2007) account for such attentional processes in their computational model of object recognition.

Several different paradigms have been developed to investigate the implications of feature attention when processing objects. These tasks include visual search paradigms (e.g., Treisman and Gelade, 1980), flanker tasks (Eriksen and Schultz, 1979), cueing tasks (Posner, 1980; Posner et al., 1980), and speeded classification tasks (Garner, 1988; Gottwald and Garner, 1975).

With respect to object recognition, there are different types of attention that are of particular interest. First, a spatial location can be attended to. In this case attention acts as a ‘spotlight’. The location that is highlighted by the spotlight is subject to enhanced processing (Posner, 1980; Treisman and Gelade, 1980). This location could be a Y junction of edges of an object or a particular part of the surface of the object, for example. Second, a feature such as colour can be attended to. Alternatively, whole objects can be selected by attention for further processing (Mounts and Melara, 1999; Scholl, 2001).

This thesis investigates how objects that consist of multiple features are represented in the brain. As described in Chapter 2.3, structural description and image-based models make different assumptions about how object features in addition to shape are represented. One way of investigating these differences is by the means of selective attention. This approach allows for testing what exactly observers are able to selectively attend to. One possibility is that the smallest unit that can be attended to is the object itself. Before attentional processes can be employed all the available features are integrated. This type of selective attention is referred to as object-based selection (e.g., O'Craven et al., 1999). Even when observers are asked to attend to only one object feature such as shape other object features such as the colour are processed automatically. If, in contrast, the features defining the object are the smallest unit that can be attended to (feature-based selection) observers are able to process each feature separately (e.g., Wegener et al., 2008). In case of feature-based selection, the integration of information about different features occurs at a later stage in the recognition process (for reviews see Mounts and Melara, 1999; Scholl, 2001).

Structural description theories (Biederman, 1987; Marr and Nishihara, 1978) postulate that the recognition process is entirely shape-driven and therefore shape is independent of other object features. In selective attention tasks, observers should therefore be able to selectively attend to shape while ignoring other features as predicted by feature-based representation. Other studies (Edelman and Bulthoff, 1992; Tarr, 1995; Tarr and Bulthoff, 1995; 1998; Tarr and Pinker, 1989), however, postulate object representations that are like series of snapshots of objects, including any of the features that were available at the time the object was learned such as viewing angle or colour. Even when participants attend to one object feature when learning and recognising objects other object features are expected to affect recognition as well as predicted by object-based representation of objects.

Different studies investigated whether different object features can be processed separately and are therefore represented independently (e.g., Cant et al., 2008; Ling and Hurlbert, 2004; Wegener et al., 2008). The results to date have been mixed, with studies showing both types of attentional selection. Thus, the specific task and the type of stimuli used appear to play an important role on how object features are represented.

Cant et al. (2008) demonstrated that shape and colour were processed independently in 2D squares. Participants were required to perform speeded classification tasks (Garner, 1988; Gottwald and Garner, 1975) on surface properties (colour or texture) or the size of the squares. Their performance was not impaired by changes in the task-irrelevant feature. This finding suggests that they were able to classify surface properties independently of size and vice versa. Their results suggest independent representations of shape and surface properties.

In Wegener et al.'s (2008) study participants watched gray or pale yellow sine-wave gratings that drifted. Participants were asked to respond to speed or colour changes in the gratings. On every trial two gratings were presented simultaneously. Prior to the onset of the gratings observers were cued to the position of the grating that would most likely undergo a change. In the first experiment Wegener et al. (2008) additionally cued the feature (colour or motion) that would most likely undergo a change. In spite of the cue, participants had to react to changes in attended and unattended locations and unattended features. They found that cueing the feature that would undergo a change resulted in shorter RTs than when the location was cued and therefore the whole grating was attended to. Wegener et al.'s (2008) results showed that participants were able to process motion and colour according to the cued feature independently.

Ling and Hurlbert (2004) investigated interactions between size and colour when comparing solid domes on the basis of either feature separately. They found that participants judged domes with more saturated colours to be bigger than domes with less saturated colours. Further, they found that participants' discrimination thresholds increased when the task-irrelevant feature was different. In contrast to the findings of Cant et al. (2008), Ling and Hurlbert (2004) found that size and colour were both processed even when only one feature was task-relevant.

2.7.1. BOLD modulation by attention

The ability to attend to relevant features allows for more efficient interaction with the environment. At the neural level, attention can affect the response pattern in the brain to visual stimuli (e.g., Corbetta et al., 1990; 1991).

2.7.1.1. Attended features

Although distinct areas of the visual cortex are sensitive to different visual features, BOLD responses can be modulated by cognitive processes such as attention. Attending to specific features was found to increase the BOLD response in brain areas that are sensitive to the attended feature (see Kanwisher and Wojciulik, 2000, for a review). In a series of PET studies, Corbetta et al. (1990; 1991) instructed participants to attend to different features (colour, shape, or motion) of visual displays. These displays consisted of random fields of coloured rectangles that moved horizontally. Across displays these rectangles were presented in different sizes, colours, and moving at different velocities. Participants performed a same-different task on the colour, the shape, or the velocity of two successively presented displays. They found that brain activity patterns depended on which feature was attended to. Attending to shape activated ventral areas of occipital and temporal cortices such as the lingual gyrus (LG), the parahippocampal gyrus (PG), and the fusiform gyrus (FFG). Similar as for attending to shape, attending to colour activated the LG. Additionally, attending to colour activated lateral occipital gyri. Performing the same-different task based on velocity activated the IPL. In summary, the results of Corbetta et al. (1990; 1991) showed that attending to a selected feature of a visual stimulus elicited a larger response the brain areas associated with the processing of this feature compared to conditions in which participants attended to all features of the stimulus.

While Corbetta et al. (1990; 1991) used simple 2D visual displays containing several small objects other studies used single 3D object that consisted of different features. In general, in these studies participants attended to one feature of the object while ignoring the other features.

Similar to the results of Corbetta et al. (1990; 1991), Peuskens et al. (2004) found that different brain areas were activated depending on the attended feature. When participants attended to shape regions along the dorsal and the ventral pathway were activated. Motion activated a similar set of brain areas as shape. Overall, motion activated the parts of the dorsal pathway out of all the areas that were activated by shape. Attention to texture exclusively activated regions along the ventral pathway. In contrast to Peuskens et al.'s (2004) results, Paradis et al. (2008) found that attending to 3D motion or 3D shape of novel objects activated separate sets of brain areas. Attending to shape exclusively activated areas along the ventral pathway; attending to motion activated high-level motion areas such as the IPL and the SPL.

As reported in Chapter 2.4.1.2, Wu et al. (2008) showed that selectively attending to either of two different motion features (manner and path) of a starfish led to a distinct pattern of BOLD activity in the brain. Attending to manner (i.e., motion of parts of the starfish) activated temporal areas; attending to the path the starfish moved along activated IPL and SPL.

Shifting attention between different object features was found to activate the posterior parietal cortex. Schultz and Lennert (2009) found that when participants learned which object feature to attend to in order to detect a target the posterior parietal cortex was involved in the process of guiding attention to the relevant object feature.

2.7.1.2. Unattended features

Attention can enhance activation in brain areas that are sensitive to the attended feature (Corbetta et al., 1990; 1991). This, however, does not mean that unattended features have no influence on the brain activity. Unattended features can modulate the BOLD response pattern. This section summarises studies that measured the response to unattended object features in the brain.

O'Craven et al. (1999) presented their subjects with stimuli in which images of houses and faces were overlaid. One of the overlaid images was moving; the other one was static. Participants attended to the house, the face, or the motion of the moving image. Of particular interest were areas that had previously been identified as sensitive to faces (the fusiform face area, FFA), houses (the parahippocampal place area, PPA), or motion (hMT/V5). The BOLD signal in FFA, PPA, and hMT/V5 was measured in each attention condition. When attending to houses or faces activity in the respective areas increased relative to baseline. Critically, when participants attended to the motion of the image the activity in FFA or PPA also increased relative to baseline. O'Craven et al.'s (1999) results suggest that even task-irrelevant information of a stimulus is processed as reflected in activity changes in the BOLD response of areas processing the task-irrelevant feature.

To separate the underlying neural signature for perception of motion and shape in SFM objects, Paradis et al. (2008) conducted an fMRI study in which participants detected transitions in shape, orientation of the rotation axis, or colour. Importantly, both attended and unattended transitions in these features modulated activation in related brain regions. Attended shape transitions activated inferior temporal, occipito-temporal, and ventral occipital cortices while attended motion transitions activated cuneus, IPL, and STS. Unattended transitions in shape activated LOC and an area at the occipito-parietal junction. Unattended transitions in motion activated IPL, the temporo-parietal junction, and frontal areas.

2.8. Thesis overview

The studies in this thesis provide a comprehensive investigation of (1) how objects consisting of shape, colour, and motion are represented by large scale brain networks, and (2) how top-down attention can modulate these large scale networks. Combined behavioural and imaging techniques were used to investigate which brain areas are involved in processing objects and object features as well as how these areas are functionally and structurally connected to each other to form brain networks. Across five experiments novel objects consisting of shape, colour, and motion were created in a way that ensured that each object feature provided the same qualitative amount of information about object identity. Furthermore, I used a selective attention paradigm to address the second goal. First, three behavioural studies will be reported that addressed how attended and unattended features affected object recognition (Experiments 1 to 3). Second, a neuroimaging study is presented that investigated the neural correlates of attention to objects and object features (Experiment 4). Functional MRI was used to measure regional brain activity and functional connectivity between regions, and diffusion tensor imaging (DTI) was used to measure structural connectivity between functional regions. Lastly, an fMRI study is presented that investigated how unattended object features influence the BOLD response to objects (Experiment 5).

3. Interaction of colour, shape, and motion in object recognition

3.1. Abstract

Three experiments were conducted to investigate how dynamic colourful objects are represented. Participants were presented with novel objects. Each object was defined by a base shape (e.g., a brick), a colour, and a motion profile (a combination of bending, stretching, and twisting motions). Participants attended to either shape or motion only. In the Experiment 1 and 2 they learned and recognised objects by the attended feature. In Experiment 3, participants performed same-different comparisons on the attended feature. Performance on all three experiments was affected by unattended shape or motion. The effects of unattended colour on performance were mixed. The results suggest combined representations of shape and motion while colour might be represented separately. However, the results also suggest strong interconnections between the representation of colour and the representations of the other features.

3.2. Introduction

Real-world objects consist of multiple features all of which can, in principal, be used for object recognition but not all features might be equally useful for recognition. The brain is very adaptive and can select task-relevant information such as object features of interest out of all the available features by attentional mechanisms. The experiments reported in this chapter investigate how object shape, object colour, and object motion are represented by combining object recognition with selective attention tasks.

Although shape is the feature that is predominately used for object recognition (Tarr and Bulthoff, 1998; Vuong and Tarr, 2006), different types of motion (e.g., Knappmeyer et al., 2003; Liu and Cooper, 2003; Newell et al., 2004; Pyles et al., 2007; Stone, 1998; 1999; Vuong and Tarr, 2004; 2006), and in some cases surface properties such as colour and texture (e.g., Naor-Raz et al., 2003; Price and Humphreys, 1989; see Tanaka et al., 2001, for a review; Tanaka and Presnell, 1999; but see also: Biederman and Ju, 1988; Davidoff and Ostergaard, 1988) have been found to be used as cues for object identity as well. Motivated by the different assumptions of the structural description (Biederman, 1987; Marr and Nishihara, 1978) and the image-based theories (Edelman and Bulthoff, 1992; Tarr, 1995; Tarr and Bulthoff, 1995; 1998; Tarr and Pinker, 1989) on how non-shape features are represented the experiments presented in this chapter investigate the representation of colourful dynamic objects.

The guiding research question is whether shape, colour, and motion of novel objects can be used independent of each other when recognising objects (see Cant et al., 2008; Ling and Hurlbert, 2004; Wegener et al., 2008). Dependent or independent processing of multiple object features is investigated by using feature attention paradigms (see Scholl, 2001). In these paradigms participants attend to only one (task-relevant) feature such as the colour while other (unattended) features such as the shape are manipulated.

Feature attention paradigms are well established to investigate whether features of an object are processed independently. Cant et al. (2008) for example, showed that colour and shape of objects are processed independently using a classic Garner paradigm (Garner, 1988; Gottwald and Garner, 1975). Participants were asked to perform a speeded classification task on basic 2D shapes (rectangles) either by their size or by surface properties such as colour or texture. When classifying surface properties, their performance did not decline when width or height of the shapes changed and vice versa (Cant et al., 2008). This finding suggested independent processing of object colour and object shape.

In contrast, Ling and Hurlbert (2004) found interference between the processing of object shape and object colour. In their study participants performed a size and a colour comparison task on illuminated domes. They found that when comparing size objects with more saturated colours appeared to be larger. Further, they showed that the discrimination thresholds for both shape and colour increased when the task irrelevant feature was varied.

Wegener et al. (2008) investigated whether participants were able to independently process motion and colour information of objects. They presented participants with moving gratings in gray or pale yellow. Participants were asked to detect changes in speed of two simultaneously presented gratings or changes in colour. They found that when participants were correctly cued about which feature (i.e., speed or colour) would undergo a change next participants responded more quickly compared to when the location (i.e., object on the left of the screen or object on the right of the screen) that would undergo a change next was cued. This showed that participants were able to process colour and motion information separately.

At the neural level, there is evidence from fMRI studies that unattended or task irrelevant object features affect the activation pattern in the brain. In their fMRI study

O'Craven et al. (1999) presented participants with superimposed images of houses and faces. One of the images moved. Participants attended to the house, the face, or the motion. Attending to one of the features increased the activity in the respective brain area that is sensitive to the attended feature. When participants attended to motion, the brain area sensitive to the ignored feature in the image also increased activity relative to baseline. This indicated that even though participants only attended to one feature of the images the unattended feature was not completely filtered out as indicated by the BOLD activity they elicited.

Similarly, Paradis et al. (2008) found that transitions in unattended shape or motion of objects modulated the activity in shape and motion selective areas, respectively. In summary, the results of the neuroimaging studies show that unattended object features can influence the neural response pattern to objects and therefore are plausible to affect recognition performance at the behavioural level.

Two sets of novel objects were created for the purposes of the experiments in this thesis. The objects were defined by a base shape (e.g., a cylinder), a colour, and a non-rigid motion profile (e.g., bending of the object). In contrast to studies that use rigid rotation as a dynamic object feature (e.g., Liu and Cooper, 2003; Stone, 1998; 1999; Vuong and Tarr, 2004; 2006), motion did not provide any information about the 3D shape of the object. For example, observers could not derive 3D shape from motion (Ullman, 1979) in addition to recovering shape from static cues. Moreover, colour did not define edges of the objects. Therefore, each feature had the same qualitative information about object identity.

Two different recognition paradigms were used. In Experiment 1 and 2 participants learned and recognised a set of objects by either the shape or the motion only. In Experiment 3 participants performed same-different comparisons on either the shape or the motion of two simultaneously presented objects. The difference between the two paradigms is in the time participants have to memorise the objects for. When participants learned the objects and recognised them later on they had to form a longer lasting representation of the objects (see Vuong and Tarr, 2004; 2006) as opposed to the experiment in which they compared two simultaneously presented objects (Vuong et al., 2009).

The experiments did not include a condition in which participants attended to colour as it was assumed that it is very easy to identify the colours used in this thesis.

Colour was manipulated as an unattended feature, however. The underlying rational is as follows. Changing unattended shape or motion necessarily causes a change in the input signal that the visual system receives. To specify effects of an unattended object feature, it is necessary to clarify whether any effects were actually caused by the unattended feature or whether the effect was caused because any change happened.

In all three experiments participants were asked to attend to either the shape or the motion when performing the tasks. This attention manipulation allowed for measuring the extent to which observers incidentally encoded the non-attended features. The unattended object features were manipulated systematically. If objects are represented as postulated by structural description theories (Biederman, 1987; Marr and Nishihara, 1978) it is expected that unattended motion or colour does not affect recognition of object shape whereas unattended shape is expected to have a strong influence on recognition of other object features. If objects are represented as 'snapshots' as postulated by image-based theories (Edelman and Bulthoff, 1992; Tarr, 1995; Tarr and Bulthoff, 1995; 1998; Tarr and Pinker, 1989) any unattended feature is expected to affect recognition performance.

3.3. Experiment 1: object learning

In Experiment 1, observers learned to individually identify four out of 64 objects (learned objects, Figure 3.3). Participants were asked to recognise these objects at a later stage in the experiment.

3.3.1. Method

3.3.1.1. Participants

Twenty-four volunteers (22 female, 2 male) participated in the experiment. The average age was 19 years (SD = 1 year), the range was 18 to 22 years. Most of them were undergraduate psychology students who participated for course credit. Other participants were compensated with 5 GBP. All participants were naïve to the purpose of the experiment and gave informed consent prior to participating. They had never seen the novel objects before the experiment.

3.3.1.2. Stimuli

The stimuli consisted of 64 novel objects (four example objects are shown in Figure 3.1) defined by individual combinations of one out of four basic shapes, one out of four colours, and one out of four non-rigid motion profiles. The stimuli were created using 3D Studio Max 9 (Autodesk, Montreal, Canada).

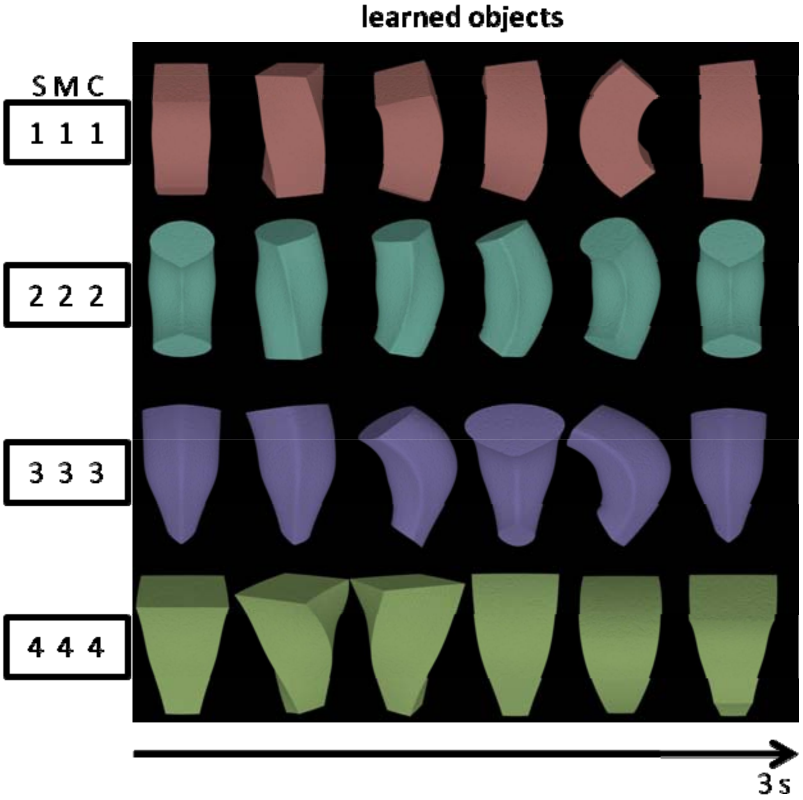


Figure 3.1. Four of the objects out of the set of 64 used for the experiments presented in this thesis. The shapes and colours were the same for all experiments. One row represents one object. Each column represents one frame of the animation. The motion profiles presented in this figure are the motion profiles used for Experiment 2, 3, 4, and 5. The four objects illustrated here are the learned objects in Experiment 2. S: shape; M: motion; C: colour. Numbers are indices for which out of four shapes, motions, or colours was used to define the object.

Shape. The four shapes were a brick, a cylinder, a tapered version of the cylinder, and a tapered version of the brick. Clearly distinguishable shapes were chosen because it was essential to use shapes that were easy to memorise and to identify even when a non-rigid motion profile was applied. Since the non-rigid motion profiles that were applied to the shapes were hard to discriminate on smooth surfaces such as the surface of the cylinder and the tapered cylinder an ‘edge’ was added to their fronts (see Figure 3.1, 2nd and 3rd row). This edge was a fold in the surface of the objects. Despite this modification, the objects were still clearly distinguishable from each other and geon-like (Biederman, 1987).

Colour. Colours were chosen to be easily distinguishable. The four colours were approximately equally spaced hue steps starting with a hue of 0 and then continuing in hue steps of 60. Thus, the four hues were 0 (red), 60 (green), 120 (blue), and 180 (purple). Luminance, saturation, and blackness were identical for the four surfaces. Please note that these values refer to the colour space implemented in 3D Studio Max 9 (Autodesk, Montreal, Canada). The monitor used to present the stimuli was not calibrated. The texture applied to all shapes was the cellular material implemented in 3D Studio Max 9 (Autodesk, Montreal, Canada). This material applies smooth bumps to the surface of the objects. This was chosen to give the objects an elastic, deformable appearance.

Motion profiles. For Experiment 1, the motion profiles were created by bending and twisting the shapes. Two shapes predominately underwent twisting motion; two shapes predominately underwent bending motion. To create a more complex motion profile the other motion type was included in the animations. At eight randomly chosen frames of the animation the amount of bending or twisting was specified. The profiles containing predominately object bending, for example, consisted of six frames on which the amount of bending was specified and two frames on which the amount of twisting was specified. 3D Studio Max 9 uses smooth morphing algorithms to apply the requested amount of bending or twisting to each frame. Importantly, although the non-rigid motion profile changed the shape over time it was ensured that observers were able to identify the underlying basic shape at any frame throughout the animation (see Figure 3.1).

Each motion profile lasted 3 s and consisted of 101 frames in total. In order to allow for smooth looping presentation of the objects, the amount of bending and twisting of frame 1 and frame 101 were identical. The parameters of the motion profiles are shown in Figure 3.2.

The objects were rendered against a black background. Animations were saved in the QuickTime 7 format. Each object had a size of approximately 7.57° visual angle along the vertical axis and 3.79° visual angle along the horizontal axis.

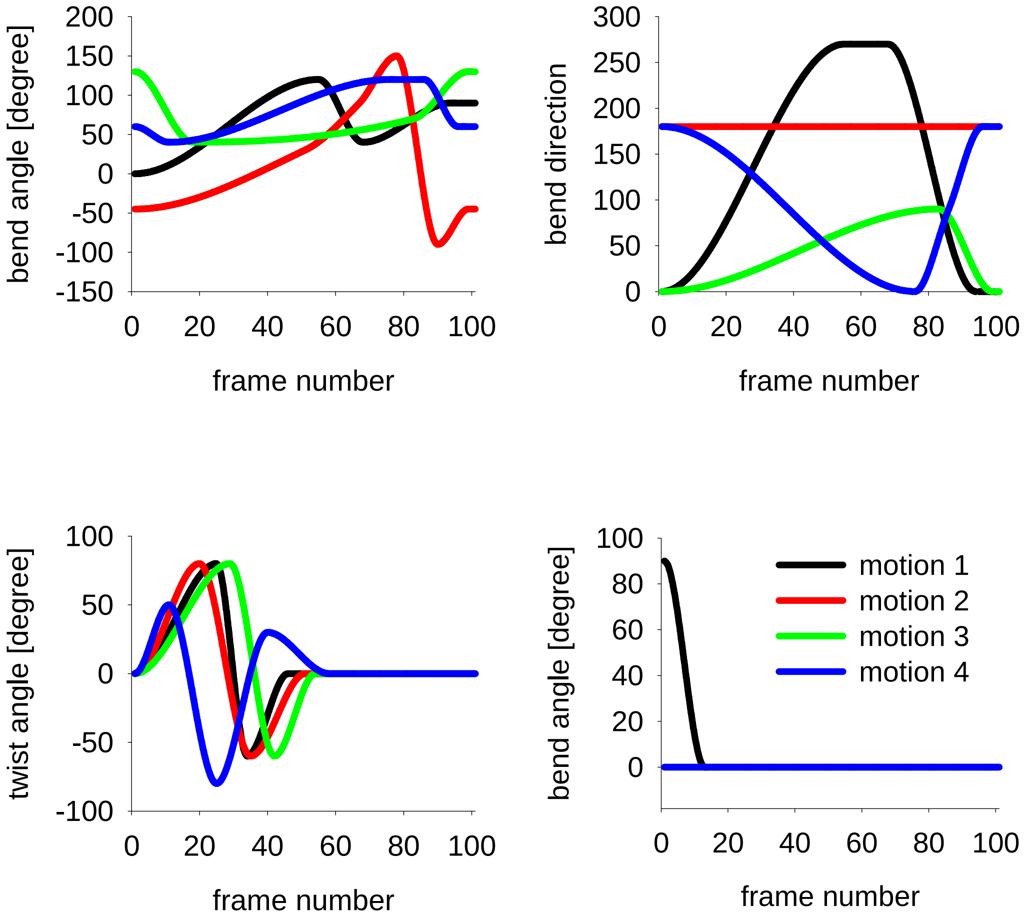


Figure 3.2. The parametric deformers of the motion profiles in Experiment 1. Bend angle: amount of bending. Bend direction: 0: towards the front of the object; 90 and 270 towards the sides of the object; 180 towards the back of the object. Twist angle: amount of rotation of two parallel surfaces in opposite directions. Note that motion 1 consisted of two independent bending profiles as reflected in two separate subplots (top left and bottom right).

3.3.1.3. Design

Participants first learned four objects on the basis of a specific feature (e.g., shape) and then were asked to subsequently identify them. Twelve participants were assigned to identify objects on the basis of their shape (*attend shape* group) and 12 were assigned to identify objects on the basis of their motion (*attend motion* group). For each group, the two unattended features were manipulated with respect to the learned objects in a 2 x 2 factorial design. That is for the *attend shape* group, colour and *motion* could be the same or different. Likewise for the *attend motion* group, colour and *shape* could be the same or different. The non-colour unattended feature is referred to as the complement feature. The dependent variables were the recognition times (RTs) and the accuracy.

3.3.1.4. Procedure

Participants were asked to learn four of the 64 objects (Figure 3.3). The four target objects were chosen so that they differed from each other on the basis of all three features (shape, motion, and colour). The same four objects were used for all participants. The experiment consisted of three phases: a learning phase, a practice phase, and lastly, a test phase.

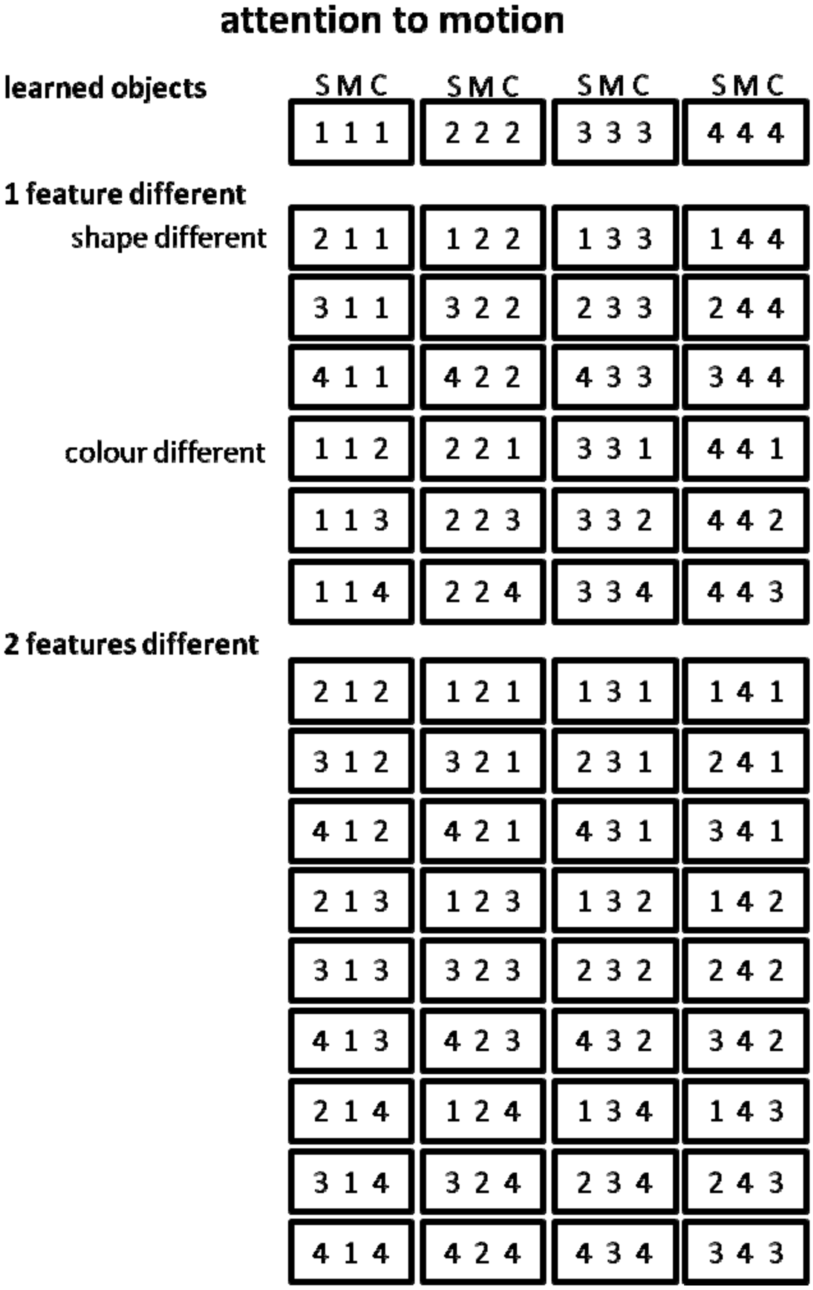
In the learning phase, participants learned to associate a different key ('d', 'f', 'j', or 'k') with either the shape (*attend shape* group) or the motion (*attend motion* group) of each target object. On each trial, a fixation cross was shown at the centre of the screen for 1 s, followed by the key associated with the target feature for 1 s, followed by the object for 3 s (Figure 3.4, left panel), followed by another fixation cross which remained on the screen until participants responded. Participants could only respond by pressing the shown key when the second fixation cross appeared. Feedback was provided at the end of each trial. The allocation of keys to objects was randomised for each participant. Each object was presented eight times. In all three phases, the objects were presented in a randomised order and the starting frame for each object animation was also randomised.

In the practice phase, participants were asked to identify the four learned objects without being shown the associated key first. Otherwise, the trial sequence was identical to the learning phase (Figure 3.4, middle panel). Each object was presented 10 times in this phase.

The test phase was the critical phase for Experiment 1. In this phase, all 64 objects, including the four learned objects, were presented. For the *attend shape* group, the 60 unfamiliar objects could differ from the learned objects in their colour, motion, or colour and motion. For the *attend motion* group, on the other hand, the unfamiliar objects could differ from the learned ones in their colour, shape, or colour and shape.

On each trial, a fixation cross appeared for 1 s. Following that, a test object was shown. The objects were shown for 3 s, or until participants responded. If they did not respond before 3 s, the objects disappeared from the screen. In contrast to the previous phases, participants were asked to respond as quickly and as accurately as possible as soon as they identified the four learned shapes (*attend shape* group) or four learned motions (*attend motion* group). A correct response was entered when participants pressed the key associated with the shape or motion despite differences to the unattended features. RTs were measured from the onset of the object. No feedback was provided during this phase.

The test phase consisted of 240 trials. Each of the four learned objects (i.e., in which the unattended features did not change) was presented 30 times (120 trials). Each of the 60 unfamiliar objects was presented twice (120 trials). On 48 trials, unfamiliar objects were presented in which one unattended feature had changed (colour or motion when participants attended to shape; colour or shape when participants attended to motion). On 72 trials, unfamiliar objects were presented that differed in both unattended features from the target objects. Figure 3.3 shows why trial numbers differ for unfamiliar objects with one or two different features. There was a self-timed break every 60 trials.



attention to motion

	S M C	S M C	S M C	S M C
learned objects	1 1 1	2 2 2	3 3 3	4 4 4
1 feature different				
shape different	2 1 1	1 2 2	1 3 3	1 4 4
	3 1 1	3 2 2	2 3 3	2 4 4
	4 1 1	4 2 2	4 3 3	3 4 4
colour different	1 1 2	2 2 1	3 3 1	4 4 1
	1 1 3	2 2 3	3 3 2	4 4 2
	1 1 4	2 2 4	3 3 4	4 4 3
2 features different				
	2 1 2	1 2 1	1 3 1	1 4 1
	3 1 2	3 2 1	2 3 1	2 4 1
	4 1 2	4 2 1	4 3 1	3 4 1
	2 1 3	1 2 3	1 3 2	1 4 2
	3 1 3	3 2 3	2 3 2	2 4 2
	4 1 3	4 2 3	4 3 2	3 4 2
	2 1 4	1 2 4	1 3 4	1 4 3
	3 1 4	3 2 4	2 3 4	2 4 3
	4 1 4	4 2 4	4 3 4	3 4 3

Figure 3.3. Distribution of learned and unfamiliar objects for the *attend motion group* in Experiment 1 and 2. Each box represents one object. S: shape; M: motion; C: colour of one of the 64 objects. The numbers indicate which out of four shapes, motion-profiles, or colours was used.

The experiment took place in a dimly lit room. It was controlled by a MS Windows PC. The objects were presented on a 16 inch computer monitor. The PsychToolbox (http://www.psychtoolbox.org; Brainard, 1997; Pelli, 1997) extension to Matlab (Mathworks, Natick, MA, USA) was used to control the experiment, present the stimuli, and collect RTs and accuracy. The experiment took approximately one hour. After the experiment participants were debriefed about the purpose of the experiment.

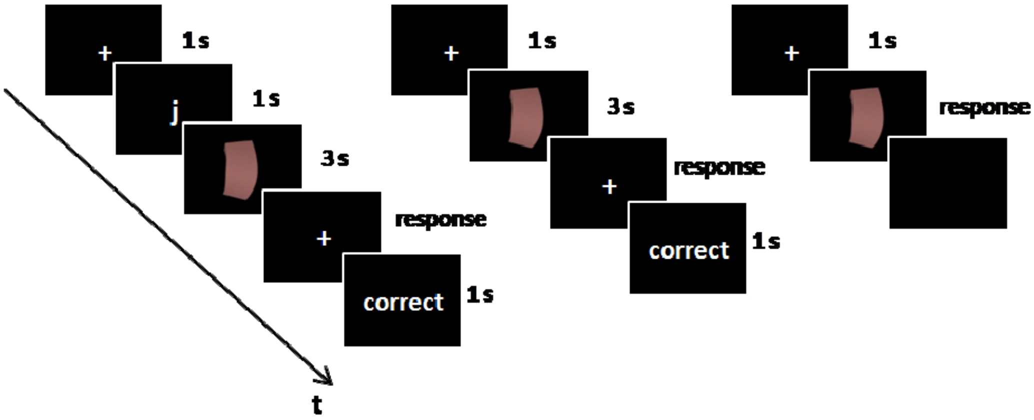


Figure 3.4. Experimental design in Experiment 1 and 2. Left panel: learning phase; ‘j’ refers to the key associated with the attended feature of the object presented next. Middle panel: practicing phase. Right panel: experiment.

3.3.2. Results

In the practising phase participants had reached an averaged accuracy level of 94.5% correct responses in the *attend motion* group and 96.6% correct responses in the *attend shape* group. For two participants (one of the *attend shape* group and one of the *attend motion* group) no accuracy data of the practising phase were available, therefore the aforementioned values are based on N = 11, respectively.

For RTs of the actual experiment, only trials on which participants responded correctly were analysed. To reduce the influence of outliers the following procedure was adopted for this experiment and Experiments 2 and 3. First, RTs greater than 6000 ms and less than 350 ms were discarded and counted as errors. Following that, correct RTs outside +/- 2.5 SDs around the mean for each participant and each condition were further removed and also counted as errors. For the *attend motion* group, the overall error was 17.1% (11.8% of wrong responses, 0.9% were outside +/- 2.5 SDs and 4.3% were smaller than 350 ms or greater than 6000 ms) which means that the overall accuracy was 82.9%. For the *attend shape* group, the overall error was 16.7% (14.4% of wrong responses, 0.3% were outside +/- 2.5 SDs and 2.0% smaller than 350 ms or greater than 6000 ms) which means that the overall accuracy was 83.3%. Please note that performance at chance level is 25%.

Correct RTs and accuracy data were first submitted to an omnibus analysis of variance (ANOVA) with colour feature (same, different) and complement feature (same, different) as within-subjects factors, and group (attend shape, attend motion) as a between-subjects factor. Separate ANOVAs were then computed for each group to further explore simple effects and interactions. The significance level was set to $\alpha = .05$ and η_p^2 was used as a measure of effect size. In this and subsequent experiments, there was no evidence for speed-accuracy tradeoffs.

Table 3.1. Ms and SEMs for Experiment 1. Complement: complement unattended feature (motion for *attend shape* group; shape for *attend motion* group).

	attend shape				attend motion			
colour	same		diff		same		diff	
complement	same	diff	same	diff	same	diff	same	diff
M [s]	1.20	1.24	1.15	1.31	2.34	2.75	2.49	2.78
SEM [s]	0.12	0.13	0.15	0.14	0.12	0.13	0.15	0.14
M [%]	83.4	82.6	84.7	82.8	85.2	81.9	84.7	78.8
SEM [%]	8.2	8.2	7.8	8.0	5.7	6.1	6.4	5.8

Note. M: mean, SEM: standard error of the means, diff: different.

3.3.2.1. Accuracy data

Table 3.1 shows means (M) and standard errors of the means (SEM) for all conditions. Of main interest was the effect of changing the unattended features on accuracy. Therefore, the mean difference in accuracy (and RT) between same and different conditions was computed. The left panel of Figure 3.5 shows the mean difference in accuracy averaged across observers as a function of *group*. In the omnibus ANOVA, there was no main effect of *group* but a main effect of complement feature (F(1, 22) = 6.06; MSE = 35.02; p = .022; η_p^2= .22; complement same: M = 84.5%; SEM = 5.0%; complement different: M = 81.5%; SEM = 5.0%). No other effect reached significance (all ps > .19).

Separate ANOVAs for each group showed that there was a main effect of complement feature (F(1, 11) = 5.45; MSE = 46.27; p = .040; η_p^2 = .33; same complement: M = 85.0%, SEM = 6.1%; different complement: M = 80.4%, SEM = 5.8%) for the *attend motion* group but not for the *attend shape* group (F(1,11) < 1.0). There was no effect of colour or interactions with colour in the omnibus ANOVA or in the separate ANOVAs.

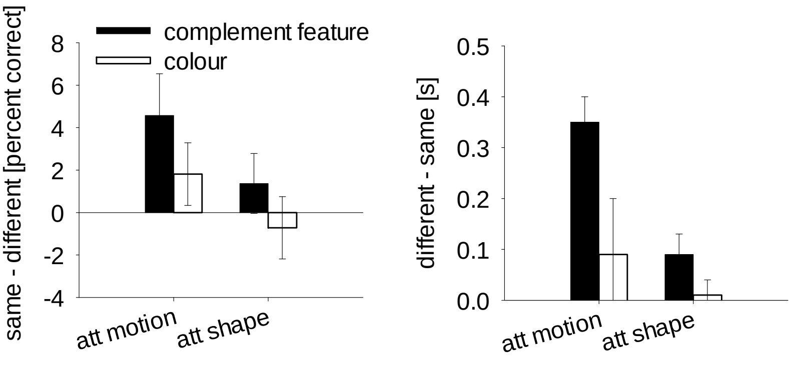


Figure 3.5. Differences between same and different unattended features in Experiment 1 as a function of *group*. Left panel: accuracy; right panel: RTs. Att motion: *attend motion* group; att shape: *attend shape* group.

3.3.2.2. Recognition time data

Differences in mean RTs between the experimental conditions are depicted in Figure 3.5, right panel. Table 3.1 shows the Ms and SEMs for each condition. In contrast to the accuracy data, the omnibus ANOVA showed a main effect of group (F(1, 22) = 62.68; MSE = 0.71; p < .0001; η_p^2= .74): participants responded more quickly when attending to shape (M = 1.23 s, SEM = 0.10 s) than to motion (M = 2.59 s, SEM = 0.14 s). There was also a main effect of complement feature (F(1, 22) = 24.92; MSE = .05; p < .0001 ; η_p^2= .53; complement same: M = 1.80 s; SEM = 0.09 s; complement different: M = 2.02 s; SEM = 0.09 s), and an interaction between group and complement feature (F(1, 22) = 8.21; MSE = .05; p = .009; η_p^2= .27).

The separate ANOVAs showed that there was a main effect of complement feature for both the *attend shape* group (F(1, 11) = 6.40; MSE = 0.02; p = .028; η_p^2= .37; same complement: M = 1.18 s; SEM = 0.09 s; different complement: M = 1.27 s; SEM = 0.11 s) and the *attend motion* group (F(1, 11) = 18.75; MSE = 0.08; p = .001; η_p^2= .63; same complement: M = 2.42 s; SEM = 0.16 s; different complement: M = 2.77 s; SEM = 0.13 s). Similar to the accuracy data, there were no main effects or interactions with colour.

3.3.3. Discussion

We found that for both the *attend shape* and *attend motion* groups, changes to the complement feature but not to colour affected recognition performance. Participants in the *attend shape* group responded more quickly when the complement motion feature was the same in the learning and test phases. Similarly, participants in the *attend motion* group responded more quickly and more accurately when the complement shape feature was the same. These findings suggest that the representation of shape and motion of objects are combined whereas colour might be represented separately.

The independent processing of shape and colour is consistent with previous research using other attentional paradigms (e.g., Cant et al., 2008). Cant et al. (2008) found that participants were able to classify size and surface properties of rectangles independently of changes in the task-irrelevant feature respectively.

Similarly, the results of Experiment 1 showed that participants were able to process object motion and object colour independently. This is consistent with the results by Wegener et al. (2008) who found that speed and colour of sine-wave gratings can be processed independently.

In the present study shape and motion interacted during object recognition. This is consistent with previous studies that found that rotation direction affected shape recognition performance (Liu and Cooper, 2003; Stone, 1998; 1999; Vuong and Tarr, 2004; 2006) and the studies that showed that non-rigid motion is relevant for recognition as well (Knappmeyer et al., 2003; Lander and Bruce, 2000; Lander et al., 1999; Pyles et al., 2007). The results of Experiment 1 show that even when motion does not provide any information about the 3D shape of novel objects (and even deforms shape as for the objects used here) participants still formed object representations in which shape and motion information are combined.

3.4. Experiment 2: object learning with modified objects

In Experiment 1, participants in the *attend motion* group responded more slowly than those in the *attend shape* group, although accuracy for both groups was the same. This difference suggests that observers were able to extract and process shape more quickly than motion. Consistent with this interpretation, the motions of the objects were very similar to each other whereas colours and shapes were easily distinguishable from each other. To further explore if the speed of feature processing would affect how unattended features influence task performance, in Experiment 2, we replicated

Experiment 1 with a new set of stimuli in which the motion profiles were made more distinguishable.

3.4.1. Methods

3.4.1.1. Stimuli

To equate processing time for both shape and motion a new set of objects was created in which the motion profiles were more distinctive. Colours and shapes were the same as in Experiment 1. The motion profiles were adapted to create more distinct motion profiles. In addition to bending and twisting, stretching and skewing of the shapes were included as non-rigid deformers. One object predominately bended, one predominately stretched, and two predominately twisted. As for the first set of objects, one other motion type was included in the animations to create a more complex motion profile. Again, each motion profile lasted 3 s and consisted of 101 frames. Animations were created in the way that at ten frames (frame numbers 1, 20, 22, 37, 40, 60, 63, 77, 80, and 97), the amount of bending, twisting, skewing (motion of two parallel surfaces of the object in opposite directions), or stretching was specified (e.g., 40 % bending at frame 1 and 20 % bending at frame 20). At eight of these frames the amount of bending, twisting, or stretching was specified and at two frames the amount of one other (i.e., bending, twisting, stretching, or skewing) motion profile was specified. Any other object properties remained as described above. The bending, twisting, stretching, and skewing profiles are shown in Figure 3.6.

To measure whether the new stimulus set was indeed easier to discriminate than the previous one Gabor similarity and motion energy of the animations were compared for the four learned objects across the two sets (see Figure 3.1 and Figure 3.3; see Vuong and Friedman, submitted).

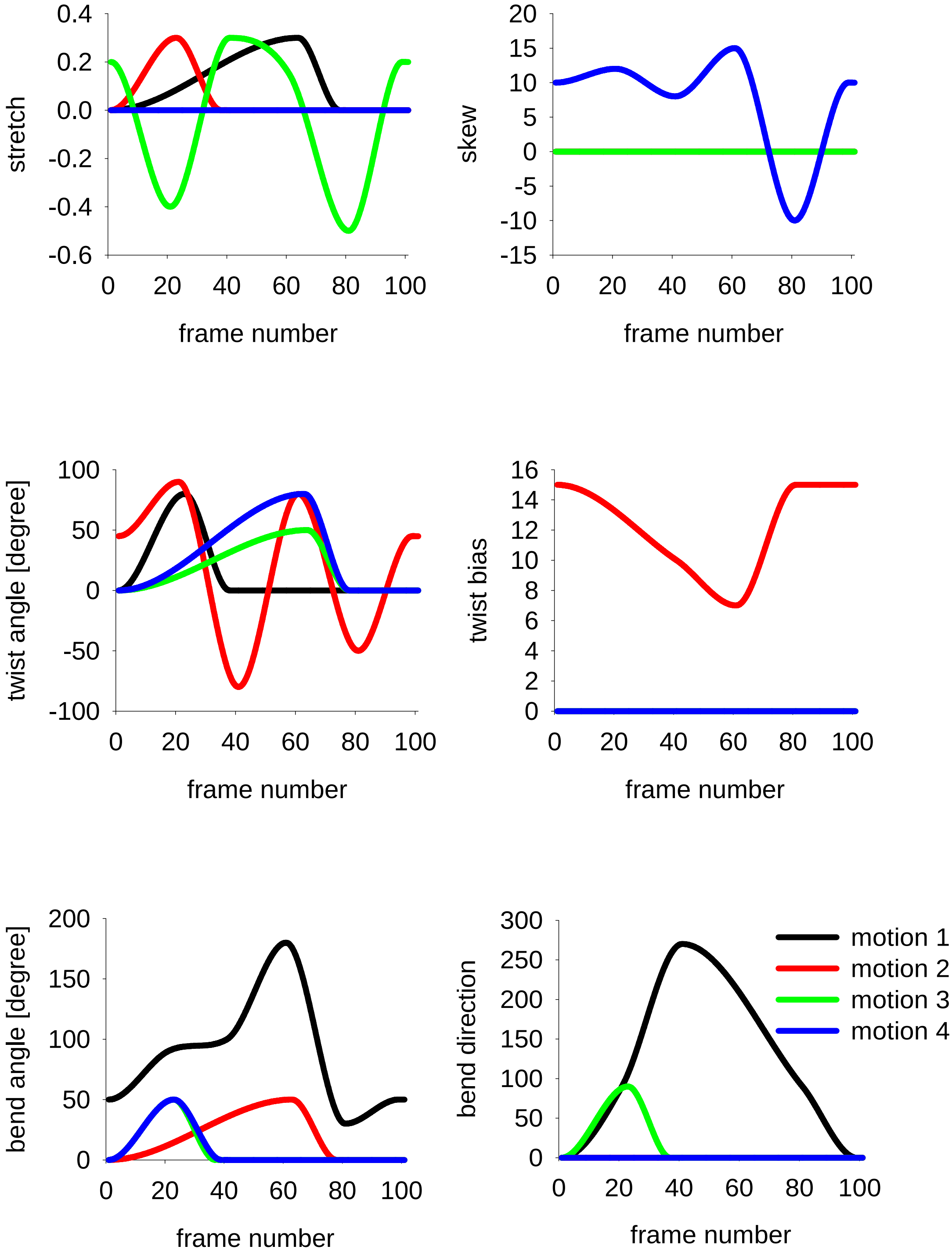


Figure 3.6. Parametric deformers of the four motion profiles in Experiment 2. Stretch: elongating the object along the vertical axis. Skew: shearing of two parallel surfaces in opposite directions. Twist angle: rotation of two parallel surfaces in opposite directions. Twist bias: bunching of both ends of the object during twisting. Bend angle: amount of bending. Bend direction: 0: towards the front of the object; 90 and 270: towards the sides of the object; 180: towards the back of the object.

3.4.1.2. Computing Gabor similarity

For both sets of objects each of the learned objects (Figure 3.1 and Figure 3.3) was compared to the other three objects. Each object pair was compared frame by frame using Gabor similarity computations. Details of this process are described elsewhere (Vuong and Friedman, submitted). In summary, each image is convolved with a Gabor filter and vectors describing the similarity are computed. If the images are identical (Gabor similarity = 0) the vectors are identical. If the images are maximally different (Gabor similarity = 1) the vectors are orthogonal. The value of Gabor similarity is averaged across the image sequence of 101 images. If the objects created for Experiment 2 are more visually dissimilar the mean Gabor similarity should be higher compared to Experiment 2. As shown by Table 3.2 the Gabor similarity was higher for the stimuli in Experiment 2.

Table 3.2. Comparison of Gabor similarity across the two sets of objects used for Experiment 1 and 2.

CSM – CSM	Experiment 1	Experiment 2
1 1 1 – 2 2 2	0.33	0.40
1 1 1 – 3 3 3	0.27	0.38
1 1 1 – 4 4 4	0.34	0.42
2 2 2 – 3 3 3	0.32	0.43
2 2 2 – 4 4 4	0.25	0.21
3 3 3 – 4 4 4	0.30	0.43
mean (SD)	0.30 (0.04)	0.38 (0.08)

Note. C: colour; S: shape; M: motion; SD: standard deviation.

3.4.1.3. Computing motion energy

As for the Gabor similarity the motion energy in animations of the objects the animations were parsed into single frames. Details of the computation of the motion energy are reported elsewhere (Vuong and Friedman, submitted). In summary, sequential pairs of frames were compared to compute the optic flow with the Lucas-Kanade-Algorithm (Lucas and Kanade, 1981). The comparison involved the displacement of each pixel from frame N to frame N+1. Vectors of direction and magnitude of the displacement were computed. Across the sequence, the average magnitude of the displacement was computed to create the optic flow image. Lastly, the mean magnitude of displacement in the optic flow image was computed. The motion energy in both stimuli sets was similar, as shown in Table 3.3.

Table 3.3. Mean motion energy for each animation of the learned objects of Experiment 1 and 2.

CSM	Experiment 1	Experiment 2
1 1 1	0.78	0.77
2 2 2	0.84	0.96
3 3 3	0.66	0.83
4 4 4	0.75	0.62
mean (SD)	0.76 (0.08)	0.80 (0.14)

Note. C: colour; S: shape; M: motion; SD: standard deviation.

3.4.1.4. Participants

Twenty-five (12 female, 13 male) volunteers participated in this experiment. They were naïve to the purpose of the study and had never seen the complete set of objects before. Some participants were members of the Institute of Neuroscience of the Newcastle University and thus were familiar with psychophysical experiments. The average age was 29 years (SD = 10 years; range: 19 to 49 years). Twelve participants were assigned to the *attend motion* group; Thirteen were assigned to the *attend shape* group. One participant was removed from the *attend shape* group for accuracies at the chance level.

3.4.1.5. Design

The same design as in Experiment 1 was used.

3.4.1.6. Procedure

The same procedure as in Experiment 1 was used.

3.4.2. Results

In the practising phase the participants reached an averaged accuracy level of 98.1% in the *attend motion* group and 99.2% in the *attend shape* group. In the actual experiment the overall accuracy for the *attend motion* group was 92.3% and 90.2% for the *attend shape* group. The same error trial procedure as for Experiment 1 was applied. For the *attend motion* group this led to an error rate of 7.7% (averaged across all participants and conditions). These 7.7% are compound of 4.4% of wrong responses, 0.6% of trials on which the RT was outside the interval between 350 and 6000 ms, and 2.7% of trials on which the RTs was outside the range of +/- 2.5 SDs around the mean. For the *attend shape* group the error rate was 9.8%. These 9.8% are composed of 6.7% of wrong responses, 0.4% of trials on which the RTs were outside the interval between

350 and 6000 ms, and 2.7% of trials on which the RTs were outside the range of +/- 2.5 SDs around the mean.

For accuracy and RTs, 2 x 2 x 2 ANOVAs with the within-subject factors colour (same; different) and complement feature (same; different) and the between-subjects factor group (*attend shape*; *attend motion*) were computed. To further explore main effects and interactions for each group separate ANOVAs for the *attend shape* and the *attend motion* group were computed additionally.

Table 3.4. Ms and SEMs of the accuracies and RTs for each experimental condition of Experiment 2. Diff: different; complement: shape for *attend motion* group; motion for *attend shape* group;

	attend motion				attend shape			
colour	same		diff		same		diff	
complement	same	diff	same	diff	same	diff	same	diff
M [s]	1.42	1.82	1.43	1.83	0.99	1.26	1.03	1.45
SEM [s]	0.14	0.18	0.13	0.16	0.05	0.07	0.04	0.05
M [%]	95.5	89.9	95.5	86.0	95.4	86.1	94.1	81.7
SEM [%]	0.6	2.5	1.0	2.3	0.6	3.1	0.6	2.8

Note. M: mean; SEM: standard error of the means; RT: recognition time.

3.4.2.1. Accuracy data

Ms and SEMs are presented in Table 3.4. Differences between the conditions in accuracy are depicted in Figure 3.7. There were no differences between the *attend shape* and the *attend motion* group (p > .20). There was a main effect of colour (F(1, 22) = 17.40; MSE = 8.74; p < .0001; η_p^2 = .44; colour same: M = 91.8%; SEM = 1.0%; colour different: M = 89.3%; SEM = 1.0%). This means that participants were more accurate when colour was the same during learning and at test. There was also a main effect of complement feature (F(1, 22) = 29.22; MSE = 70.85; p < .0001; η_p^2= .57; complement same: M = 95.2%; SEM = 0.4%; complement different: M = 85.9%; SEM = 1.8%) showing that participants were more accurate when the complement feature was the same during learning and at test. The interaction of colour and complement was significant (F(1, 22) = 5.71; MSE = 11.43; p = .026; η_p^2= .21) showing that accuracy was lowest when both unattended features were different at test.

The separate ANOVAs for the two experimental groups showed the following results. The results for the *attend motion* group were consistent with Experiment 1.

There was a main effect of complement feature (F(1, 11) = 13.00; MSE = 55.20; p = .004; η_p^2 = .54; same complement: M = 95.69; SEM = 0.54; different complement: M = 87.96; SE = 2.36) indicating that accuracy was higher when the shape was the same with respect to the learning phase. Unlike Experiment 1, this experiment showed an effect of colour (F(1, 11) = 9.12; MSE = 6.23; p = .012; η_p^2 = .45; same colour: M = 92.9%; SEM = 1.3%; different colour: M = 90.7%; SEM = 1.5%) indicating that accuracy was higher when the colour was the same with respect to the learning phase. Additionally, there was a marginally significant interaction between colour and shape (F(1, 11) = 4.45; MSE = 37.14; p = .059; η_p^2 = .28). This means that accuracy was lowest when both colour and shape were different compared to the learning phase.

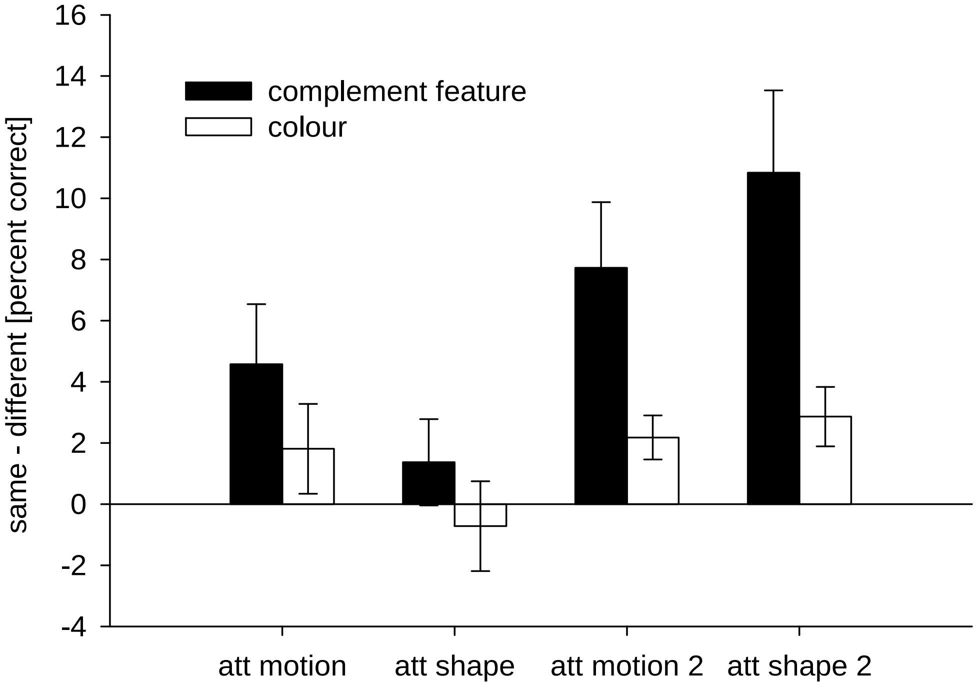


Figure 3.7. Differences in accuracy in Experiment 1 and 2. Error bars in this and subsequent plots are +/- 1 SEM. Att mot: attend motion group in Experiment 1; att shape: attend shape group in Experiment 1; att motion 2: attend motion group in Experiment 2; att shape 2: attend shape group in Experiment 2.

In the *attend shape* group there was a main effect of colour (F(1, 11) = 8.72; MSE = 11.25; p = .013; η_p^2 = .44; colour same: M = 90.8%; SEM = 1.6%; colour different: M = 87.9%; SEM = 1.4%) and a main effect of complement feature (F(1, 11)

= 16.31; MSE = 86.51; p = .002; η_p^2 = .60; complement same: M = 94.8%; SEM = 0.5%; complement different: M = 83.9%; SEM = 2.7%) showing that participants were more accurate when colour and motion, respectively, were the same during learning and at test. The interaction of colour and complement was not significant (p > .18).

3.4.2.2. Recognition time data

Table 3.4 shows the Ms and SEMs of Experiment 2. RTs data are depicted in Figure 3.8. The omnibus ANOVA revealed a main effect of colour (F(1, 22) = 4.80; MSE = 0.02; p = .039; η_p^2 = 0.18; colour same: M = 1.37 s; SEM = 0.08 s; colour different: M = 1.44 s; SEM = 0.07 s) showing that participants responded more quickly when unattended colour was the same during learning and at test. Further, there was a main effect of complement feature (F(1, 22) = 97.36; MSE = 0.03; p < .0001; η_p^2 = .0.82; complement same: M = 1.22 s; SEM = 0.07 s; complement different: M = 1.59 s; SEM = 0.09 s) showing that participants responded more quickly when the complement feature was the same during learning and at test. Overall, participants responded more quickly in the *attend shape* group compared to the *attend motion* group (F(1, 22) = 8.38; MSE = 0.57; p < .0001; η_p^2 = 0.28; *attend shape*: M = 1.18 s; SEM = 0.04 s; *attend motion*: M = 1.63 s; SEM = 0.15 s).

For clarity separate ANOVAs for the *attend motion* and the *attend shape* group were computed. The results of the *attend motion* group were consistent with the results of the *attend motion* group in Experiment 1, there was an effect of complement feature (F(1, 11) = 40.29; MSE = 0.05; p < .0001; η_p^2 = .79; same complement: M = 1.43 s; SEM = 0.13 s; different complement: M = 1.83 s; SEM = 0.17 s). Note that the effect of unattended colour found in the accuracy data was not present in the RTs (all remaining ps > .81). For the *attend shape* group in contrast, there was a main effect of colour (F(1, 11) = 8.79; MSE = 0.02; p = .013; η_p^2 = .44; colour same: M = 1.12 s; SEM = 0.05 s; colour different: M = 1.24 s; SEM = 0.04 s) showing that participants responded more quickly when the colour was the same between learning and testing. Further, there was a main effect of complement showing that RTs were shorter when unattended motion was the same during learning and at test (F(1, 11) = 68.26; MSE = 0.02; p < .0001; η_p^2 = .86; complement same: M = 1.01 s; SEM = 0.04 s; complement different: M = 1.35 s; SEM = 0.05 s). The interaction of colour and complement was significant as well (F(1, 11) =

6.85; MSE = 0.01; p = .024; η_p^2 = .34) showing that RTs were longest when both colour and complement feature were different between learning and testing.

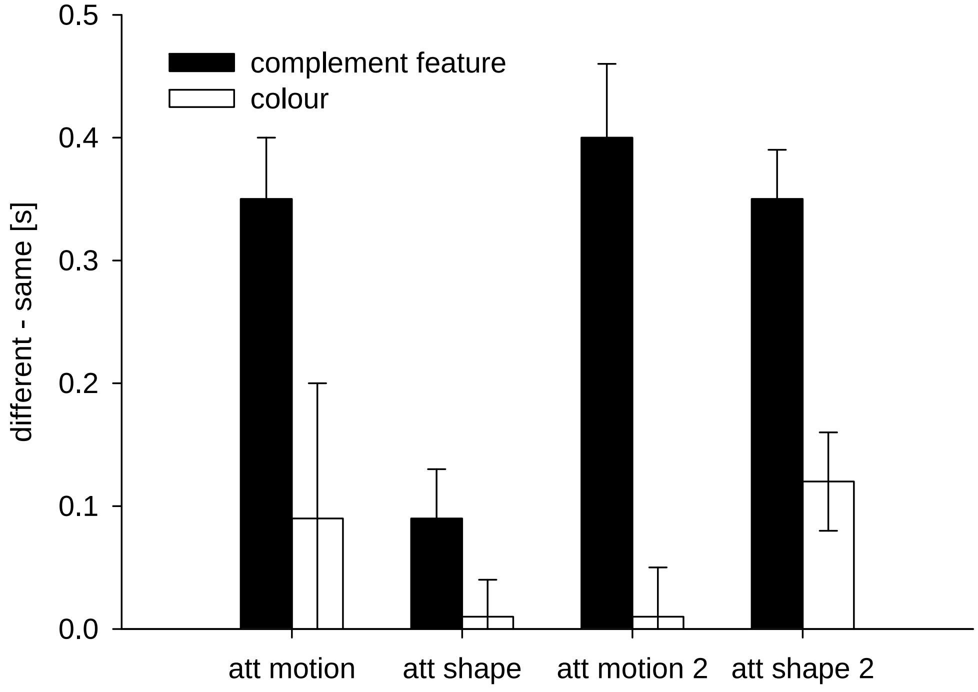


Figure 3.8. Differences in RTs in Experiment 1 and 2. Att mot: attend motion group in Experiment 1; att shape: attend shape group in Experiment 1; att motion 2: attend motion group in Experiment 2; att shape 2: attend shape group in Experiment 2.

3.4.3. Discussion

In terms of RTs, the results of Experiment 2 were in line with Experiment 1. Only in the *attend shape* group of Experiment 2 participants were faster when colour was the same during learning and at test. In this condition RTs were longest when both colour and motion were different at test. Accuracy data of Experiment 2 diverged from Experiment 1 with respect to the influence of unattended colour. Accuracy was decreased when the colour of the objects was different in the learning and the test phase. Further, an interaction (significant in the *attend shape* group and marginally significant in the *attend motion* group) between colour and complement feature indicated that accuracy was worst when both colour and shape had changed.

Overall, performance was better in Experiment 2 compared to Experiment 1. This was expected, as the motion profiles of the objects were more discriminable.

Consequently, the task was easier for the participants which resulted in better performance.

In line with Experiment 1 motion and shape seem to be represented in a combined way which does not allow for independent processing of the two features. The results for unattended colour of this experiment are inconsistent. Although there was no effect of unattended colour in the RTs of the *attend motion* group, Experiment 2 suggests that unattended colour influences processing of object shape and motion which is at odds with the results of Experiment 1. This is remarkable because the colours of the objects were the same as in Experiment 1. Therefore, the saliency of colour cannot explain this effect. Further, task difficulty can be ruled out as an explanation. Vuong and Tarr (2006) showed that participants take more object features in account when recognition becomes more difficult. As the motion profiles of the objects were made more distinct in Experiment 2 compared to Experiment 1 our data surprisingly show the opposite effect. Participants in both attend motion groups were (as expected with stimuli that are easier to discriminate) faster in Experiment 2 compared to Experiment 1.

In summary, Experiment 1 and 2 provide evidence that the representations of shape and motion are combined. The results for processing of colour are diverging. Experiment 1 and the RT data of the *attend motion* group in Experiment 2 suggest independent processing of colour from other object features as found by Cant et al. (2008). The accuracy data of Experiment 2 in general and the RT data of the *attend shape* group in Experiment 2, however, show that unattended colour can affect the processing of motion and suggest links between the representations of all three features.

3.5. Experiment 3: Same-different comparisons

Experiment 1 and 2 suggest that the representations formed for object shape and object motion are combined as changes to the task-irrelevant complement feature affected performance. The results of Experiments 1 and 2 are inconsistent for whether colour can be processed independently. To further clarify the role of colour another experiment was conducted using a different paradigm. The previous experiments required the participants to form a long-term memory representation of the objects presented in the learning phase. During the test phase participants compared the objects on the screen with the representations they had formed. Experiment 3 required participants to compare features of objects that were presented simultaneously.

3.5.1. Method

3.5.1.1. Participants

Twenty-five new volunteers (23 female, 2 male) participated in the experiment. The average age was 20 years (SD = 2 years), the range was 18 to 30 years. One participant of the *attend motion* group was excluded from the analyses due to a technical failure during the experiment. Most of them were undergraduate psychology students, participating for course credit. Other participants were compensated with 5 GBP. All participants were naïve to the purpose of the experiment and gave informed consent prior to participating. They had never seen the novel objects before the experiment.

3.5.1.2. Stimuli

The same set of stimuli as in Experiment 2 was used.

3.5.1.3. Design

The experiment was set up as a 2 x 2 x 2 x 2 factorial design with the between-subjects factor group (*attend shape*, *attend motion*) and the within-subjects factors trial type (same, different), colour (same, different), complement (same, different). Therefore, there were eight experimental conditions. As in Experiment 1 and 2, the dependent variables were the RTs and the accuracy.

3.5.1.4. Procedure

As in Experiment 1, two attention groups were included: twelve participants attended to shape (*attend shape* group) and 12 participants attended to motion (*attend motion* group). Participants were asked to perform a same-different comparison task on the attended feature. Two objects were presented simultaneously on every trial: one object was presented at the centre of the right half of the screen; one object was presented at the centre of the left half of the screen (the centre of the object was approximately 7° visual angle shifted horizontally to the left and the right of centre of the screen). The size of the objects was the same as in Experiment 1.

Each trial started with a fixation cross which was shown for 0.5 s. The objects were presented for 6 s. Participants were asked to respond as fast and as accurately as possible any time after the onset of the objects. The screen turned black if no response had been entered after 6 s. Responses were entered by pressing the ‘c’ or the ‘n’ key on a standard keyboard. Assigning the keys to the ‘same’ and ‘different’ responses was counterbalanced. The duration of each motion profile was 3 s. To present the objects for a total 6 s the animation looped. The starting frame of the animation was randomised on each trial for both objects. Therefore the objects were presented out of phase in their motion. This was necessary to control for differences in task difficulty between trials with same motion and trials with different motion. If the motion profiles had been presented in synchrony it would have been very easy for participants to identify trials on which objects had the same motion.

The setup of the experiment was identical with Experiment 1 and 2. A set of practice trials was conducted prior to the experiment. A total of 480 trials were presented. Due to a technical error the number of trials per condition varied slightly for each participant (each condition was presented between 49 and 71 times). Overall, an experimental session took approximately one hour. After the experiment participants were debriefed about the purpose of the study.

3.5.2. Results

First, RTs and accuracy data were submitted to omnibus repeated measure ANOVAs. For clarity, separate ANOVAs for the two attention groups were computed in addition. The overall accuracy was for the *attend motion* group was 82.6% and 87.8% for the *attend shape* group. The same error procedures as in Experiment 1 and 2 were applied. For the *attend motion* group this led to an error rate of 17.4% (averaged across all participants and conditions). These 17.5% are compound of 12.6% of wrong responses, 3.4% of trials on which the RT was outside the interval between 350 and 6000 ms, and 1.4% of trials on which the RT was outside the range of +/- 2.5 SDs of the mean. For the *attend shape* group, this led to an error rate of 12.1% (averaged across all participants and conditions; please note that due to rounding errors the overall accuracy and the error rate add up to 99.9% instead of 100%). These 12.1% are compound of 9.0% of wrong responses, 0.7% of trials on which the RT was outside the interval between 350 and 6000 ms, and 2.4% of trials on which the RT was outside the range of +/- 2.5 SDs of the mean.

Table 3.5. Ms and SEMs for each experimental condition of Experiment 3.

	attend shape							
trial	same				diff			
colour	same		diff		same		diff	
complement	same	diff	same	diff	same	diff	same	diff
M [s]	1.44	1.59	1.53	1.57	1.52	1.59	1.60	1.57
SEM [s]	0.14	0.15	0.15	0.16	0.14	0.17	0.16	0.14
M [%]	93.5	91.7	93.8	91.4	81.4	84.2	82.2	84.0
SEM [%]	1.0	1.1	1.7	1.5	3.2	2.7	3.0	3.1

	attend motion							
trial	same				diff			
colour	same		diff		same		diff	
complement	same	diff	same	diff	same	diff	same	diff
M [s]	2.38	2.55	2.33	2.57	2.74	2.60	2.81	2.66
SEM [s]	0.10	0.11	0.11	0.13	0.17	0.16	0.18	0.16
M [%]	89.2	79.1	87.9	77.6	79.3	85.3	79.4	82.8
SEM [%]	3.1	5.3	3.3	5.9	4.2	2.6	4.2	3.2

Note. Diff: different; M: mean; SEM: standard error of the means.

3.5.2.1. Accuracy data

Ms and SEMs of the accuracies in Experiment 3 are shown in Table 3.5. The omnibus ANOVA revealed that there were no differences in accuracy between the two groups (F(1, 22) = 1.9; MSE = 356.15; p = .19; η_p^2 = .08; *attend shape* group: M = 87.9; SEM = 2.7%; *attend motion* group: M = 82.5%; SEM = 2.7%). For clarity further results of the omnibus ANOVA are reported in Table 9.1 of the Appendix (Chapter 9.1).

Separate ANOVAs for each group showed the following results. In the *attend shape* group there was an effect of trial type indicating that accuracy was higher on same trials compared to different trials (F(1, 11) = 13.01; MSE = 170.75; p = .004; η_p^2 = .54; same trial: M = 92.6%; SEM = 1.1%; different trial: M = 83.0%; SEM = 2.8%). As shown in Figure 3.9, left panel, trial type interacted with complement feature (F(1, 11) = 6.50; MSE = 18.29; p = .027; η_p^2 = .37). Tukey post hoc tests (see Appendix in Chapter 9.1 for the formula) revealed that this interaction was driven by higher accuracy on same trials compared to different trials when the complement feature was the same. Further, accuracy was higher on same trials compared to different trials when the complement feature was different. No other effect was significant (all ps > .37).

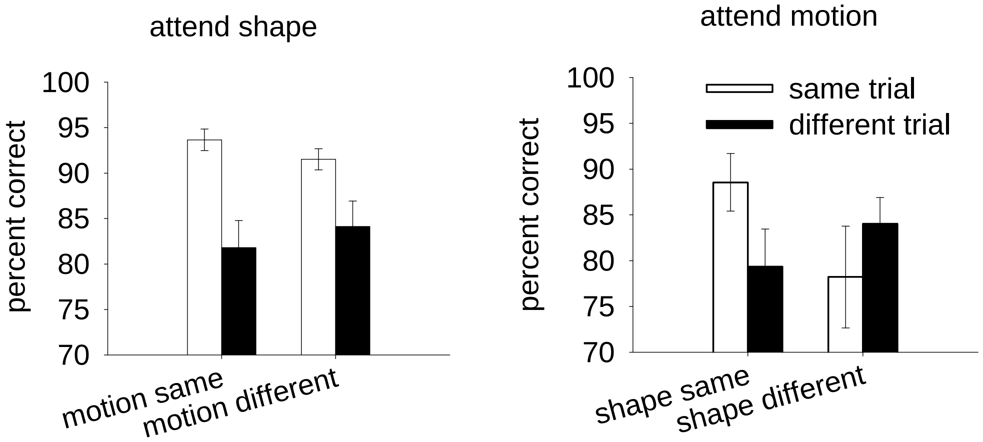


Figure 3.9. Interaction of trial type by complement feature for accuracy data of Experiment 3.

In the *attend motion* group, there was trend that participants were more accurate when the complement feature (i.e., the shape) of the objects was the same (F(1, 11) = 4.33; MSE= 44.32; p = .062; η_p^2 = .28; complement same: M = 84.0%; SEM = 3.2%; complement different: M = 81.1%; SEM = 3.9%). Consistent with the *attend shape* group there was an interaction of trial type and complement feature (F(1, 11) = 12.82; MSE = 105.64; p = .004; η_p^2 = .54) which is shown in Figure 3.9, right panel. Tukey post hoc tests did not reveal significant differences between the four means. However, they revealed a trend suggesting that the interaction was driven by higher accuracy on same trials when the complement feature was the same. Further, accuracy was higher when the complement feature was the same on same trials compared to same trials on which the complement feature was different. No other effect was significant (all ps > .19).

3.5.2.2. Recognition time data

Means and SEMs of the RT data are shown in Table 3.5. The omnibus ANOVA showed that the *attend shape* group responded more quickly than the *attend motion* group (F(1, 22) = 27.53; MSE = .92; p < .0001; η_p^2 = .56; *attend shape* group: M = 1.55 s; SEM = 0.14 s; *attend motion* group: M = 2.58 s; SEM = 0.13 s). Interestingly, there was a main effect of colour (F(1, 22) = 4.62; MSE = 0.01; p = .043; η_p^2 = .17; colour same: M = 2.05 s; SEM = 0.10 s; colour different: M = 2.08 s; SEM = 0.10 s). For clarity other effects found in the omnibus ANOVA are reported in Table 9.1 of the Appendix (Chapter 9.1).

The ANOVAs that were computed for each attention group separately provided the following results. In the *attend shape* group there was a main effect of complement feature indicating that RTs were longer when the complement feature (the motion) of the two objects was different (F(1, 11) = 14.39; MSE = 0.01; p = .003; η_p^2 = .57; complement same: M = 1.52 s; SEM = 0.14 s; complement different: M = 1.58 s; SEM = 0.15 s). There was an interaction of trial type with complement feature (F(1, 11) = 7.61; MSE = .01; p = .019; η_p^2 = .41) as shown in Figure 3.10, left panel. Tukey post hoc tests revealed that this interaction was driven by shorter RTs on same trials compared to different trials when the complement feature was the same. No other effect was significant (all ps > .07).

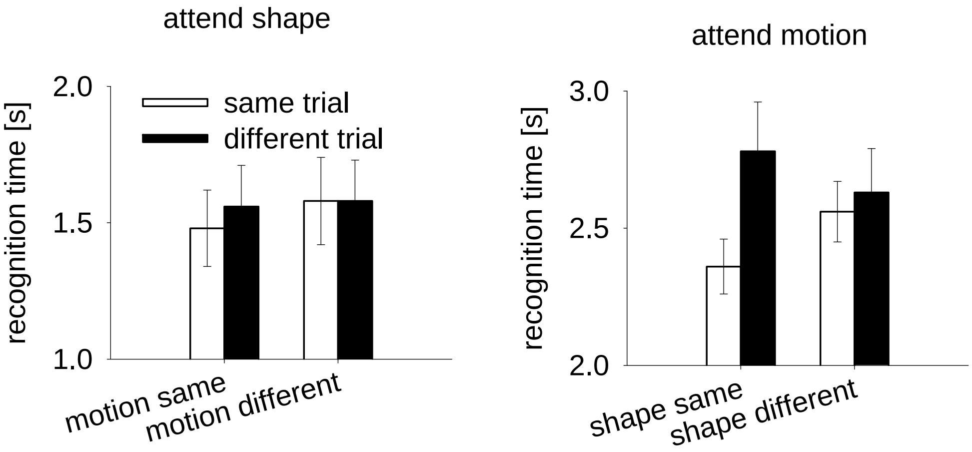


Figure 3.10. Interaction of trial type by complement feature for RTs of Experiment 3.

In the *attend motion* group there was a main effect of trial type indicating that participants were faster on same trials (F(1, 11) = 7.87; MSE = 0.19; p = .017; η_p^2 = .42; same trial: M = 2.46 s; SEM = 0.11 s; different trial: M = 2.70 s; SEM = 0.17 s). Consistent with the *attend shape* group and the accuracy data, trial type interacted with complement feature (F(1, 11) = 27.47; MSE = 0.03; p < .0001; η_p^2 = .71) as shown in Figure 3.10, right panel. Tukey post hoc tests revealed that this interaction was driven by shorter RTs on same trials when the complement feature was the same compared to same trials on which the complement feature was different. Further, shorter RTs were identified on same trials compared to different trials when the complement feature was the same. No other effect was significant (all ps > .10).

3.5.3. Discussion

In Experiment 3 participants performed same-different comparisons on the shape or the motion of two objects that were presented simultaneously. Consistent with Experiment 1 and 2 it was found that unattended shape and unattended motion affected task performance. However, this effect was not found in form of an overall decrease in performance but depended on whether participants responded to a same or a different trial. This interaction of complement feature and trial type showed that performance decreased when the complement features were different on a trial on which participants had to respond ‘same’ and that performance decreased when participants had to respond ‘different’ when the complement features were the same. No such interaction was present when trial type and object colour. These results are in line with the findings of Experiment 1 and suggest combined representations of object shape and object motion. Unattended colour again, was found to affect the processing of the attended feature to a smaller extent compared to unattended shape or motion. Although there as an overall increase in RT for trials on which objects with different colours were presented, this effect was not found when separately analysing RTs for the *attend shape* and the *attend motion* group. No effect of unattended colour was found in the accuracy data. This suggests that unattended colour only had a weak influence on recognition of shape and motion.

3.6. General discussion for Experiments 1 to 3

Three experiments were conducted to investigate how objects consisting of multiple features are represented. In Experiment 1 and 2 participants learned and recognised objects by either shape or motion; in Experiment 3 participants performed same-different comparisons on either the shape or the motion of two simultaneously presented objects. Two sets of 64 novel objects defined by combinations of shape, colour, and non-rigid motion were created to ensure that each object feature contributed the same qualitative information to the recognition process.

Consistently, all three experiments showed that unattended motion and unattended shape affected task performance. Participants responded more slowly when unattended shape or unattended motion were different and accuracy was decreased. The interactions between shape and motion in object recognition found in the present studies are in line with previous research (Liu and Cooper, 2003; Stone, 1998; 1999; Vuong and Tarr, 2004) that showed that shape recognition depended on rotation direction. In contrast to the aforementioned studies, the experiments reported in this thesis used non-rigid motion. Similar effects of non-rigid motion were expected, however, as studies into face perception (Knappmeyer et al., 2003; Lander and Bruce, 2000; Lander et al., 1999; Pilz et al., 2006) reported that non-rigid motion is an object feature that is used to identify complex objects such as faces.

While shape and motion consistently interacted in the present experiments the results for object colour are diverging. Experiment 2 and 3 provide some evidence for representation of object colour in combination with shape and motion whereas Experiment 1 suggests that object colour is represented separately of other object features. There are two possible explanations for the deviation of the results of Experiment 1 and 2. First, the objects in Experiment 1 were harder to discriminate and therefore led to noisier data than Experiment 2. The effect of unattended colour might be so small that it was overshadowed by the noisy data in Experiment 1. Second, the participants in Experiment 1 were mostly first year psychology students who were unfamiliar with participating in psychophysical experiments whereas a large number of the participants in Experiment 2 were employees of the Institute of Neuroscience and therefore well trained in psychophysical experiments. The expertise of the participants in Experiment 2 could have led to less noisy data compared to Experiment 1 which allowed for finding subtle effects such as the dependence of the performance on unattended colour. Overall, Experiments 1 to 3 support the concept of object-based representation of shape and motion. The results for colour were inconsistent.

Previous research provided diverging results on whether colour can be processed independently of other object features (Cant et al., 2008; Ling and Hurlbert, 2004; Wegener et al., 2008). These diverging results might be due to the different paradigms that were used in these studies and the specific types of objects. This cannot explain the inconsistent results found in the three experiments reported here. Especially the different results found for unattended colour between Experiment 1 and Experiment 2 cannot be explained by the stimuli as the same set of objects was used in both experiments.

Price and Humphreys (1989) suggested that shape and colour of objects are encoded in separate but strongly interconnected representations. The results of Experiments 1 to 3 are in line with this idea. While shape and motion might be represented in one integrated representation, information about object colour might be combined by connections between independent representations.

Both the structural description models and the image-based models grant shape as the most important feature for object recognition (Biederman, 1987; Marr and Nishihara, 1978; Tarr and Bulthoff, 1998). Across all three experiments unattended motion affected the recognition of shape suggesting that participants had formed a representation of the shape that included information about the specific motion profile that was applied to the shape. An alternative explanation could be that participants represented only shape information but since the shape was deformed by the motion the effect could be due to the simple fact that participants saw a different shape when the motion profile was different and therefore recognition performance decreased when the motion was different. This is unlikely to account for the whole effect because it was ensured that participants were able to identify the present shape unambiguously at any frame during the animation. For image-based theories the influence of unattended motion on shape recognition can be explained by the decrease in similarity between the stored snapshots of an object and the perceived object. For structural description theories the influence of unattended motion on shape recognition might have emerged from participants encoding the changes in the non-accidental properties caused by a specific non-rigid motion profile along with the shape as an ‘animated geon’. A different non-rigid motion profile led to different changes in the non-accidental properties and was therefore harder to recognise.

Not only did unattended motion impair shape recognition in Experiments 1 to 3. Unattended shape impaired recognition of motion as well. This further supports the concept of a combined representation of shape and motion. Participants were unable to extract the motion profile from its carrier, a particular shape. This is consistent with the both the image-based and the structural description theories. Structural description theories (Biederman, 1987; Marr and Nishihara, 1978) postulate that shape is the feature that is predominately used for object recognition. Even when a less important feature (i.e., motion) is attended to shape is still expected to dominate the recognition process. Therefore, changes in unattended shape impaired the recognition of motion profiles.

The influence of unattended shape on motion processing can also be explained in the framework of the image-based theories (Edelman and Bulthoff, 1992; Tarr, 1995; Tarr and Bulthoff, 1995; 1998; Tarr and Pinker, 1989). As these theories postulate that objects are stored as series of snapshots the same motion profile applied to another shape leads to decreased similarity of the perceived object and the stored snapshots and therefore leads to decreased performance.

Neither the image-based nor the structural description models can explain the effects caused by unattended object colour. Image-based theories would have predicted that unattended colour would affect recognition just as shape and motion did. The structural description theories did not predict an effect of unattended colour on shape recognition. The current results suggest a modification of the object recognition theories that accounts for the small influence of colour on object recognition. Possibly a synthesis of image-based and structural description theories could account for the results found here.

3.7. Conclusion

The three experiments reported here suggest object-based representation of object shape and object motion. Unattended colour affected the recognition of object shape and object motion to a smaller extend compared to unattended shape or motion. The current results support the concept of a separate representation of object colour as suggested by Price and Humphreys (1989). To account for the relations between the representations of shape, colour, and motion found in Experiments 1 to 3 a synthesis of image-based and structural description theories might be necessary.

4. Experiment 4: Structural and functional networks involved in recognising dynamic objects

4.1. Abstract

In this study fMRI and DTI were combined to investigate networks that process colourful dynamic objects. To identify brain areas that respond to certain object features an fMRI study was conducted in which participants were asked to attend to shape, motion, and colour separately. To identify brain areas that integrate multiple object features a condition in which participants attended to whole objects was included. Attending to shape and motion activated a common network consisting of occipito-temporal areas, lateral frontal areas, and areas in the posterior parietal cortex along the IPS whereas attending to colour activated areas along the CoS and medial brain areas. The shape sensitive LOC integrated multiple object features. DTI provided further evidence for a common network that processes object shape and object motion: higher integrity of white matter connections between the occipito-temporal and lateral frontal parts of the network was associated with faster RTs when participants attended to shape, motion, or whole objects but not when they attended to colour. These results suggest a network that processes shape and motion of dynamic objects whereas object colour seems to activate a different network.

4.2. Introduction

Understanding how object recognition systems use different object features to achieve object recognition provides essential knowledge about how cognitive processes are structured. Experiments 1 to 3 found that shape and motion are more tightly linked than shape and colour, and motion and colour. At the neural level, previous research showed that large scale networks, including occipito-temporal, parietal and frontal regions process complex dynamic objects (Peuskens et al., 2004; Schultz et al., 2008). Thus, in this chapter and in Chapter 6, the neural correlates of attention to features of colourful non-rigidly moving objects are explored.

In Experiment 4, fMRI and DTI were used to investigate networks involved in object recognition at a functional and an anatomical level. Further, these neuroimaging techniques were combined with a modified version of the tasks used in Experiments 1 to 3. In particular, across different blocks, participants attended to either the individual features or to the whole object while they were in the scanner. This manipulation allowed for comparing the brain networks involved in dynamic object recognition when participants attended to the individual features relative to the other features (e.g., motion versus shape and colour) and relative to when they attended to the whole objects (i.e., all three features simultaneously). Finally, the paradigm allowed for determining whether the integrity of anatomical connections between regions has an impact on recognition performance for attention to individual features and whole objects.

Functional MRI allows us to measure regional hemodynamic changes as well as how regional responses may share functional connectivity (Haynes et al., 2005; Macaluso et al., 2000). Complementing this, diffusion imaging allows us to measure the structural integrity of large white matter tracts that may connect functional regions. Such structural connectivity allows distant regions to communicate with each other.

Communications between brain areas that are part of a common network can be measured indirectly by investigating temporal correlations in the BOLD responses of these brain areas (Haynes et al., 2005; Macaluso et al., 2000) or directly by investigating the integrity of the white matter that connects these brain areas. Schultz et al. (2008) found correlations in the residual BOLD response between the brain areas that were activated during their object recognition experiment. The residual BOLD response is the change in BOLD signal that cannot be explained by the experimental paradigm (i.e., the error between the expected and the measured BOLD response). This suggests that these brain areas interacted during the experiment in a way that was independent of the experimental paradigm. The correlations indicate that the brain areas activated by the experimental task not only show co-activation but that they are actually part of one network that is engaged in the task.

Previous research (e.g., Begre et al., 2007; Bohr et al., 2007; Thomas et al., 2009; Thomas et al., 2008; Tuch et al., 2005) showed that measurements of behavioural performance such as accuracy and reaction times correlated with the FA value, a measurement of white matter integrity (see Chapter 2.6). Following previous research (Kim and Kim, 2005), we were interested in fibre tracts that connected brain areas that were activated in the fMRI experiment. We investigated the relationship between task performance (i.e., the RT) and the integrity of white matter tracts connecting areas that responded to the objects used in this thesis.

Previous DTI studies (see ffytche and Catani, 2005, for a review) have shown that there are large association tracts of white matter that connect occipital regions to frontal, temporal and parietal regions which were identified as relevant for object processing (e.g., Peuskens et al., 2004; Schultz et al., 2008). In particular, these tracts have been shown to be important for visual memory and face recognition. Thomas et al. (2009) found that patients with congenital prosopagnosia showed less structural integrity of white matter in ventral occipito-temporal areas compared to a control group. Further evidence for the relationship of face recognition with white matter integrity was shown in aging (Thomas et al., 2008). They found that age-related decline in face perception performance was accompanied by a reduction in white matter integrity of the right inferior fronto-occipital fasciculus (IFOF) a fibre tract that connects the occipital lobe with temporal and frontal areas.

Investigating memory and learning, Begre et al. (2007) found that higher white matter connectivity in the superior longitudinal fasciculus (SLF) was related to better performance on a memory test that involved processing of black-and-white figures. The SLF is a fibre bundle that is located in dorsal part of each hemisphere and that connects posterior and anterior parts of the brain. Its fibres project into the occipital, parietal, temporal, and frontal lobe.

Research into memory decline in aging showed that performance on non-verbal memory tasks was related to reduction of white matter integrity in the ILF which connects occipital and temporal brain areas (Sasson et al., 2010). This is consistent with a case study of a patient with partial disconnection of the right ILF who showed extensive impairments in visual memory (Shinoura et al., 2007).

None of the DTI studies conducted so far directly investigated which white matter tracts are involved in object recognition. The evidence from face perception and visual memory, however, suggests that the white matter tracts that connect the occipital lobe to temporal and frontal areas (IFOF, SLF, ILF) could play a crucial role in object recognition as faces are a special type of object and memory processes are crucial for object recognition.

Following Peuskens et al. (2004), we used a paradigm in which participants were asked to attend to a single feature of the objects to identify regions and networks responsive to specific objects features. Moreover, we included a condition in which participants were asked to attend to the whole object. This condition provides knowledge on whether object features are integrated or whether the BOLD response to whole objects is similar to the sum of BOLD response patterns for each single object feature.

Previous research identified brain areas that integrate different features in occipital and temporal areas. Evidence for integration of shape and motion of objects was provided by research into biological motion perception (Grossman et al., 2000). Grossman et al.'s (2000) study showed that the STS integrated body posture with motion of humans. Consistent with this, Schultz et al. (2008) found that integration of shape and motion of novel objects also activated STS.

Other brain areas that have been found to integrate shape and motion of objects were the anterior fusiform area and the motion sensitive hMT/V5 (Sarkheil et al., 2008). This is in line with findings on implied motion (Kourtzi et al., 2002). When participants viewed static images displaying a type of motion (e.g., an athlete about to throw a discus compared to an athlete standing) hMT/V5 showed increased activation. This suggests that hMT/V5 is not only activated by actual motion but is relevant for integration of shape and motion as well.

Self and Zeki (2005) used symmetry judgements of simple objects to investigate integration processes of colour and motion. In an experiment in which both motion and colour defined the object they found that the more anterior portion of LOC integrated motion and colour information.

In addition to brain areas integrating multiple object features there are brain areas that combine information about different object features without integration. Previous research suggested that the posterior parietal cortex plays an important role in memorising multiple object features. Xu et al. (2007) investigated the role of the IPS for visual short term memory. They used simple 2D shapes in different colours and participants had to respond whether a certain colour, a certain shape (experiment 1) or a conjunction of a colour and a shape (experiment 2) was present in a display presented earlier. They found that IPS activity reflected the total number of object features held in visual short term memory suggesting that the IPS combined the information about shape and colour. Consistent with this, Kawasaki et al. (2008) provided further evidence for the importance of the parietal cortex for visual short term memory. Using a set of simple 2D objects defined by shape, colour, and motion of the texture dots they found that the activity in the posterior parietal cortex reflected visual short term memory load caused by all types of features and feature conjunctions. The studies on feature conjunctions required participants to memorise different object features similar to the condition in the present study in which participants attended to whole objects. It is expected that similar memory processes are necessary to perform the task of the present study.

Based on previous work (Peuskens et al., 2004; Schultz et al., 2008), a large scale network consisting of occipital-temporal, frontal, and parietal areas was expected to respond to the objects. Attending to single object features is expected to increase the BOLD response of the part of the network that is sensitive to the attended feature (Kanwisher and Wojciulik, 2000). The condition in which participants attend to the whole object is expected to activate brain areas that are sensitive to shape, colour, and motion simultaneously or to activate integration areas such as STS (Grossman et al., 2000; Schultz et al., 2008). Functional (Haynes et al., 2005; Macaluso et al., 2000) and structural connections between the parts of the network were expected. Structural connections between cortical areas that constitute the network activated in the fMRI experiment were reconstructed. If these connections are crucial for object recognition it is expected that the integrity of the white matter that connects parts of the network is correlated with the behavioural performance on the task. Functional connections were expected in correlations between the activity time course between parts of the network.

4.3. Methods

4.3.1. Participants

Sixteen participants volunteered for this study (8 males; 8 females). Their age ranged from 20 to 36 years (M = 24 years; SD = 4 years). Most of them were either staff or students from Newcastle University. All participants had normal or corrected to normal vision, and were naive to the stimuli used and the purpose of the study. They were informed about the safety precautions for experiments involving a high field MR scanner, and gave informed consent. The study was approved by the Newcastle University ethics committee. In addition, all safety protocols and guidelines set by the Newcastle Magnetic Resonance Centre were followed. After the experiment participants were debriefed about the purpose of the experiment.

4.3.2. Stimuli

The same 64 stimuli as for the behavioural Experiment 2 and 3 were used. We used a canon XEED LCD projector (1280 x 1024 pixel resolution) to project the visual stimuli onto a projection screen at the foot-end of the scanner. Subjects viewed the projection through an angled mirror attached to the head coil approximately 10 cm above the subjects’ eye. The experiment was controlled by a MS Windows PC. Stimuli were presented with MATLAB 7.1 (MathWorks, Naticks, MA, USA) and the Psychophysics toolbox (http://www.psychtoolbox.org, Brainard, 1997; Pelli, 1997). The experiment was conducted with the lights in the scanner room turned off.

4.3.3. Design

The experiment was set up as a 2 (trial type: same, different) x 4 (attended feature: shape, colour, motion, all) within-subjects design. The four attention conditions were run in separate blocks and each block consisted of an equal number of same and different trials randomly ordered. Each functional run consisted of 12 experimental blocks. The order of the blocks in one run was pseudo-randomly determined as follows. Each run was divided into three sets. Each attended feature condition was presented once in each set. It was ensured that each attention condition was preceded by a different attention condition across the three sets in a run. For example, in one run the attention to motion condition was preceded by attention to shape in the first set, by attention to colour in the second set, and was presented as the first condition in the third set. There were also 13 fixation blocks (one at the beginning of the run and one after each experimental block). A fixation block consisted of a white fixation cross rendered against a black background. Participants were instructed to fixate on the cross during these blocks. Each experimental block lasted approximately 31 s and each fixation block lasted 12 s. Participants were tested on three functional runs, each approximately 9 min in length.

4.3.4. Procedure

Figure 4.1 shows the continuous 1-back task used in the present study (see also Schultz and Lennert, 2009). In this task, participants judged whether the attended feature of the current object (on Trial N) matched the attended feature of the preceding object (on Trial N – 1). No response was made to the first object in a block. At the beginning of each experimental block, the word ‘colour’, ‘shape’, ‘motion’ or ‘all’ was shown to indicate the attended feature for that block. The instruction was presented at the centre of screen for 2 s in white letters rendered against a black background. Participants then saw a sequence of nine objects, each shown for 2.5 s. The presentation duration was less than a full cycle of motion for the objects. However, based on our behavioural data, that was a sufficient amount of time to judge all three features. The size of an object was approximately 7.67° visual angle in height and 3.84° visual angle in width.

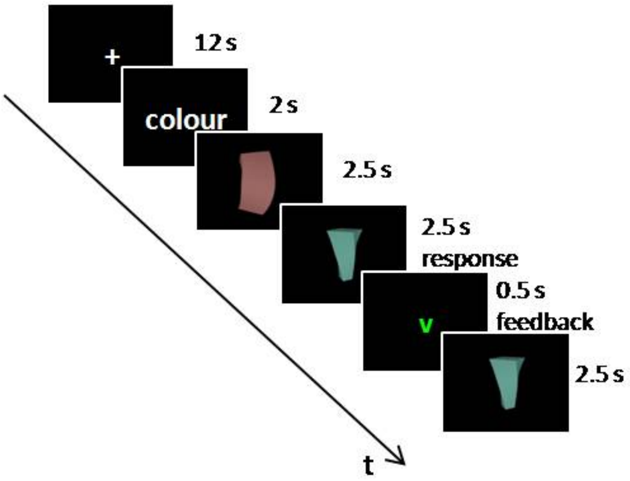


Figure 4.1. Experimental procedure of Experiment 4. The green ‘v’ was used as feedback for a correct response. Note that no feedback was provided after the first object as no response was required. There were 0.5 s blanks between the objects and the feedback that are not shown in the figure.

Participants could respond at any time after the onset of an object. If participants did not respond while the object was present, the trial was counted as an error trial. They made their responses with an MRI compatible response box (LumiTouch™) using the index and middle finger of their dominant hand. The assignment of finger to same or different response was counterbalanced across participants. Feedback was provided after the object disappeared. Correct responses were indicated by a green ‘v’, while errors and misses were indicated by a red ‘x’ which appeared on the screen for 500 ms. In between feedback and the next object there was a 500 ms blank.

There were four same trials and four different trials on each block. On same trials all features matched on two consecutive objects independent of the attended feature. On different trials either one or two unattended features were different in addition to the attended feature. In 25% of the different trials all three features were different between consecutive objects. On the remaining 75% of the different trials the attended feature and one other feature differed between consecutive objects. The non-attended feature that was different was randomly determined on each trial.

Participants practiced the continuous 1-back task outside the scanner to familiarise themselves with the stimuli, block sequence, and response mapping. They practiced with at least one block of each of each condition. They also received a few practice trials while they were in the scanner to ensure that they could see the stimuli and to become familiar with the response box used.

4.3.5. Image acquisition

Scans were acquired at the Newcastle Magnetic Resonance Imaging Centre with a 3 T Philips Achieva Intera scanner. The signal was received with an 8-channel head coil (SENSE by Philips). Scanning sessions started with a high resolution T1-weighted scan, followed by three functional runs and closed with a diffusion-weighted scan. For structural scans 150 T1-weighted images with a resolution of 208 x 208 voxels (the field of view (FOV): 240 mm x 240 mm; thickness: 1.2 mm) were acquired. For the functional scans T2* weighted echo planar images (EPIs) were acquired. These consisted of 29 axial slices (TR = 2 s; flip angle = 90º; 64 x 64 voxels; FOV = 192 mm x 192 mm; thickness: 3 mm; 1 mm gap in between slices; TE = 40 ms). Slices were acquired from the bottom to the top of the head. Before each functional run, preparation steps were performed to allow for equilibration of the T1 signal.

4.3.6. fMRI data preprocessing

The data were processed and analysed using the SPM8 toolbox (Wellcome Department of Imaging Neuroscience, http://www.fil.ion.ucl.ac.uk/spm/). To preprocess the functional data, images were realigned to the first functional image that was acquired of each participant and resliced with a resolution of 3 mm x 3 mm x 3 mm. Images were then smoothed with a 6 mm full width half maximum (FWHM) Gaussian kernel. The smoothed images were normalized to the Montreal Neurological Institute (MNI) EPI T2*-weighted template and resampled with a voxel size of 3 mm x 3 mm x 3 mm (Friston et al., 2005). Lastly, low-level drifts in the preprocessed data were removed with a cutoff of 128 s for each data set. An autoregressive model (AR(1)) was used to estimate serial correlations and adjust the degrees of freedom appropriately.

4.3.7. fMRI data analysis

The GLM approach implemented in SPM8 is a mixed model that uses two steps for the analysis. The first step is a fixed-effects model at the single subject level; the second step is a random-effects model at the group level. A design matrix consisting of a linear combination of regressors was built to estimate the contribution of each regressor to the observed data (Friston et al., 1995). Each experimental condition, the instructions, and the fixation condition were modelled as boxcar functions convolved with the hemodynamic response function (HRF, in SPM8 the difference of two gamma functions). In each session each of the 10 conditions (instruction, eight experimental and one fixation condition) was a regressor in the GLM. The six movement parameters (yaw, pitch, roll, and 3 translation terms) that were computed during realignment were modelled as regressors of no interest. There were also three regressors to model the separate runs and a constant term for each run. Thus, there were a total 51 regressors in the design matrix. The duration of each trial was modelled as the time in which the object was present (2.5 s). A beta-weight was estimated for each regressor in order to minimize the residual error between the modelled and the measured BOLD response. Contrasts of interest were created by subtracting the beta-weights of two or more regressors. These contrasts involved main effects for each attended feature compared to the other features, differences in trial type, and interactions between trial type and attended feature. For example, to test for voxels with a larger BOLD response during the attention to colour blocks, we used the contrast in which the regressor for colour was weighted with 1, both the regressor for attention to shape and motion was weighted with -0.5, and all other regressors were weighted with 0 (so that the contrast weights summed to zero).

For the group-level analysis, statistic parametric mapping (SPM) T maps were then created to allow for statistical inferences. No further smoothing of the data was applied at the group level. Statistical tests were performed on the contrasts images by testing the contrast against zero in a one-sample t-test at each voxel. Unless reported differently we set the significance value to $\alpha = .05$ (corrected for multiple comparisons at the cluster level). Minimum cluster size was set to 10 voxels.

4.4. Results

4.4.1. Behavioural results

RTs and accuracies were calculated for the four experimental conditions. Due to a technical error no behavioural data were available for two participants. Thus, the behavioural results were based on 14 participants. The outlier trimming procedures of Experiments 1 to 3 were not applied because responses were only collected during the presentation time of an object (2.5 s). The overall accuracy was 94.9%. The worst performance was found on different trials when participants attended to shape (86.5%). Although the shapes were chosen to be easily discriminable (Biederman, 1987) this decline in accuracy might reflect that the non-rigid motion deformed the shape in a way that impaired recognition.

The results for the accuracy are plotted in Figure 4.2, left panel. There was a main effect of attended feature ($F(3, 39) = 13.56$; MSE = 11.54; $\eta_p^2 = .51$; $p < .0001$).

Tukey post hoc tests revealed that this was caused by lower accuracy when attending to shape compared to attention to colour or whole objects. Further, participants were more accurate on same trials as opposed to different trials (F(1, 13) = 10.52; MSE = 21.29; η_p^2 = .45; p = .006). The interaction of attended feature and trial type was significant (F(3, 39) = 14.13; MSE = 13.61; η_p^2 = .52; p < .0001). Tukey post hoc tests revealed that this interaction was caused by significantly lower accuracy on different trials when participants attended to shape.

A similar pattern was found for the RTs (Figure 4.2, right panel). There was a main effect of attended feature (F(3, 39) = 39.15; MSE = .02; η_p^2 = .75; p < .0001). Tukey post hoc tests showed that this main effect was driven by shorter RTs on colour trials. There was no main effect of trial type (F(1, 13) = 0.30; η_p^2 = .02; p = .59). The interaction of attended feature and trial type was significant (F(3, 39) = 4.77; MSE = .01; η_p^2 = .27; p = .006).

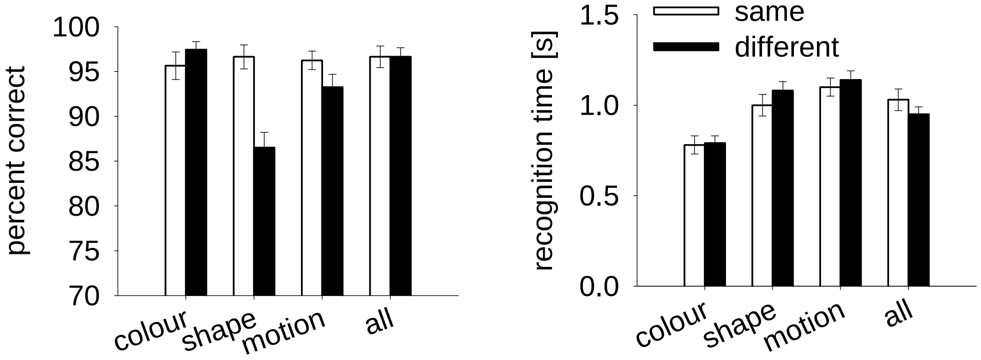


Figure 4.2. Behavioural data of Experiment 4. The attended feature is plotted on the x-axis. Left panel: accuracy; right panel: RTs.

4.4.2. fMRI results: attending to features

As a first step we analysed which brain areas showed larger responses when participants attended to one object feature compared to the other object features. As the activity patterns found for same and different trials were widely consistent we collapsed the data for both trial types. Please note that for all the results reported here MNI space was used with positive x-coordinates referring to the right hemisphere and negative x-coordinates referring to the left hemisphere. In images the right hand side refers to the right hemisphere of the participant. The WFU pickatlas toolbox for SPM5 was used in Experiment 4 and 5 to help identifying which anatomical landmark corresponded to the voxels that were significantly activated (Maldjian et al., 2004; Maldjian et al., 2003, http://www.nitrc.org/projects/wfu_pickatlas). Table 4.1 shows an overview over key studies that have identified brain areas that responded to shape, motion, or colour. Please note that this table is not supposed to provide a complete summary of previous research. Rather, it is supposed to provide landmarks to which the areas activated in Experiment 4 can be compared to.

Table 4.1. Key studies that reported the coordinates of brain areas that responded to shape, motion, or colour. Decimal places occur because of conversion from Talairach & Tournoux (1988) to MNI space.

		MNI coordinates [mm]		
	brain area	x	y	z
shape areas				
Malach et al., 1995	LO	43.2	-75.8	15.8
Grill-Spector and Malach, 2001	LO (1)	40.4	-74.2	-1.9
	LO (1)	-41.4	-79.4	-1.0
	LO (2)	-37.4	-70.5	-16.0
	LO (2)	-36.4	-72.4	-19.7
Kourtzi et al., 2002	LOC	41.8	-63.2	-17.8
	LOC	-38.3	-67.7	-12.5
motion areas				
Tootell et al., 1995	MT	45.5	-78.4	-1
Kriegeskorte et al., 2003	hMT	46.5	-63.8	-3.7
	hMT	-43.4	-69.9	-7.6
Peuskens et al., 2004	hMT (1)	54	-60	0
	hMT (1)	-45	-72	3
	hMT (2)	51	-66	3
	hMT (2)	-54	-72	3
	hMT (3)	48	-63	-3
	hMT (3)	-51	-69	0
colour areas				
Zeki et al., 1991	V4	20.2	-67.8	-8.7
	V4	-26.3	-69.6	-13.6
McKeefry and Zeki, 1997	V4	30.0	-76.2	-27.1
	V4	-29.3	-69.3	-20.7
Cant & Goodale, 2007	CoS (exp. 1)	23.2	-56.0	-18.7
	CoS (exp. 1)	-37.4	-42.5	-20.3
	CoS (exp. 2)	27.3	-57.0	-18.8
	cuneus	7.1	-88.3	12.8

Note. MT: middle temporal, hMT: human MT, CoS: collateral sulcus, LO: lateral occipital, LOC: lateral occipital complex.

For main effects of trial type we found that only when participants attended to motion there were stronger BOLD responses on different trials compared to same trials. For main effects of feature attention we found that attending to motion or shape activated similar clusters (Figure 4.3). These included occipito-temporal areas, clusters along the IPS, and lateral frontal areas. The occipito-temporal clusters were in brain areas that corresponded to previously reported locations of LOC and hMT/V5 (see Table 4.1). Colour, in contrast, predominately activated medial brain areas along the cingulate gyrus (Figure 4.4). The MNI coordinates of the peak voxel of each activated cluster are shown in Table 4.2, Table 4.3, and Table 4.4 for attention to colour, shape, and motion, respectively. Surprisingly, attending to colour did not activate colour sensitive areas in ventral occipito-temporal cortices. Lowering the initial significance threshold to $\alpha = .01$, however, revealed a cluster in the ventral part of the posterior occipital cortex that extended in both hemispheres (Figure 4.5). The cluster extended along both lingual gyri and in the cuneus bilaterally (MNI coordinates: $x = 12$, $y = -64$ $z = -8$; $k = 192$; $Z = 3.96$; $p < .05$ FDR corrected at the cluster level). The peak voxel of this cluster was located slightly more medial and ventral than the colour sensitive V4 and CoS identified by previous research (see Table 4.1). Additional clusters that were activated in Experiment 4 are reported in Table 9.2, Chapter 9.2.

4.4.3. fMRI results: attending to whole objects

When participants attended to all object features the initial threshold was lowered to $\alpha =.01$. Three clusters in lateral and posterior occipital cortices were activated (Figure 4.6 and Table 4.5). The location of these clusters suggests that the area that is activated to a larger extend when participants attend to whole objects as opposed to separate features is the LOC.

Table 4.2. MNI coordinates of the peak voxel of each cluster showing larger activation when attending to colour as opposed to shape and motion.

	MNI				
region	x	y	z	k	Z
attention colour > attention shape/ motion					
CingC	0	50	-11	1315	5.35*
MFG	-27	29	40	75	4.34*
MFG	-30	41	25	78	4.57*
MTG	-60	-16	-14	155	3.97*
MTG	-42	-67	25	36	3.91+
Insula	48	5	-2	263	4.95*
PG	-21	-13	-23	27	3.28+
Precuneus	-12	-49	58	26	4.38+

Note. * p < .0001, FDR correction for multiple comparisons at the cluster level. + p < .05, FDR correction for multiple comparisons at the cluster level. K: number of voxels in the cluster; CingC: cingulate cortex; MFG: middle frontal gyus; MTG: middle temporal gyrus; PG: parahippocampal gyrus.

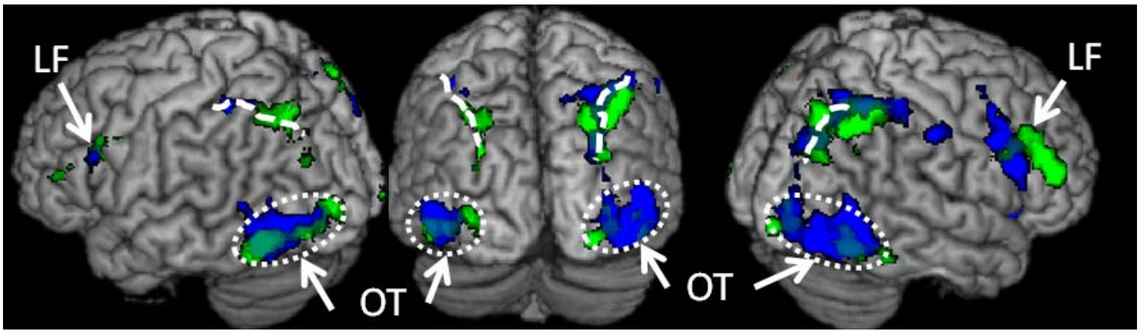


Figure 4.3. Networks that show larger BOLD response for attending to one object feature compared to attending to either of the other two features. Green: attention to shape; blue: attention to motion. Dashed lines indicate the rough outline of the IPS. Dashed circles mark the outline of occipito-temporal areas that include hMT/V5 and LOC. LF: lateral frontal.

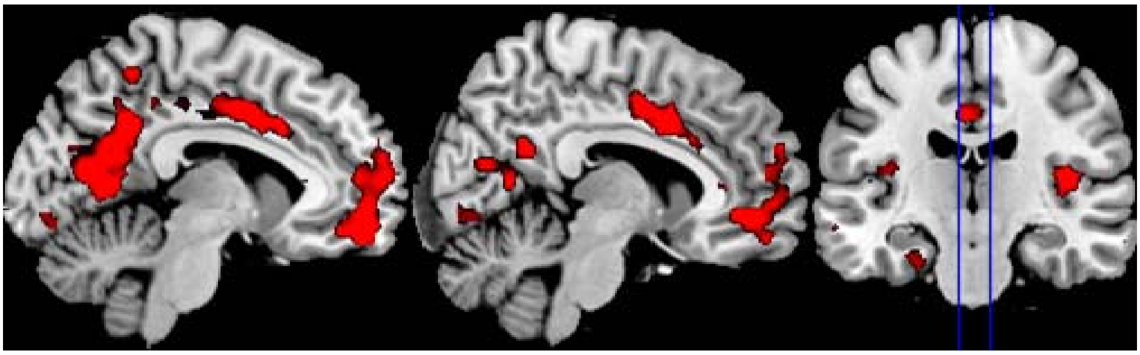


Figure 4.4. Networks that showed stronger activation when participants attended to colour compared to attention to motion or shape. The saggital slices are MNI x coordinates -7 and +7.

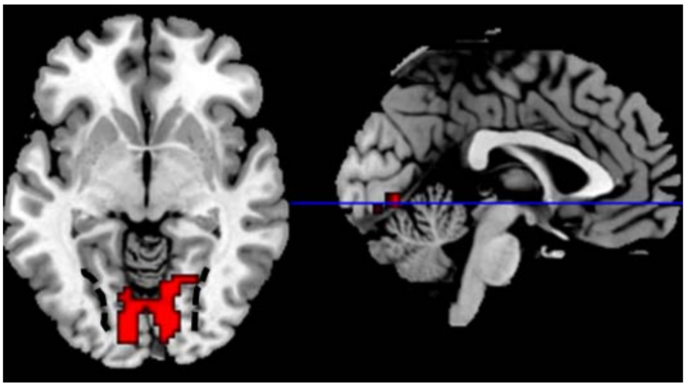


Figure 4.5. Larger activation on attention to colour trials compared to attention to shape or motion trials in bilateral lingual gyri (LG). The initial threshold was lowered to .01 for this result. Dashed lines mark the rough outline of the collateral sulcus (CoS). The LG is located medially relative to the CoS.

Table 4.3. MNI coordinates of the peak voxel of each cluster showing larger activation when attending to shape as opposed to attending to colour or motion.

	MNI				
region	x	y	z	k	Z
attention shape > attention colour/motion					
IFG	-45	11	25	48	3.72+
IFG	48	38	10	224	4.91*
IPS	-27	-73	31	108	4.03*
IPL	39	-61	43	445	4.57*
MOG	-45	-85	-5	151	4.28*
FFG	45	-67	-14	111	3.93*
Cun	27	-100	-5	54	3.97+

Note. * p < .0001, FDR correction for multiple comparisons at the cluster level. + p < .05, FDR correction for multiple comparisons at the cluster level. MNI coordinates in mm. k: number of voxels in the cluster.

Table 4.4. MNI coordinates of the peak voxel of each cluster showing larger activation when attending to motion as opposed to shape or colour

	MNI				
region	x	y	z	k	Z
attention motion > attention colour/shape					
IFG	-57	14	31	28	3.68+
IFG	33	23	-8	42	4.18+
IFG	45	5	31	270	4.85*
IOG	45	-82	-11	439	5.23*
MOG	-48	-76	-2	327	5.08*
IPL	-39	-46	55	38	4.31+
Central s.	27	-61	58	347	4.33*

Note. * p < .0001, FDR correction for multiple comparisons at the cluster level. + p < .05, FDR correction for multiple comparisons at the cluster level. MNI coordinates in mm. k: number of voxels in the cluster.

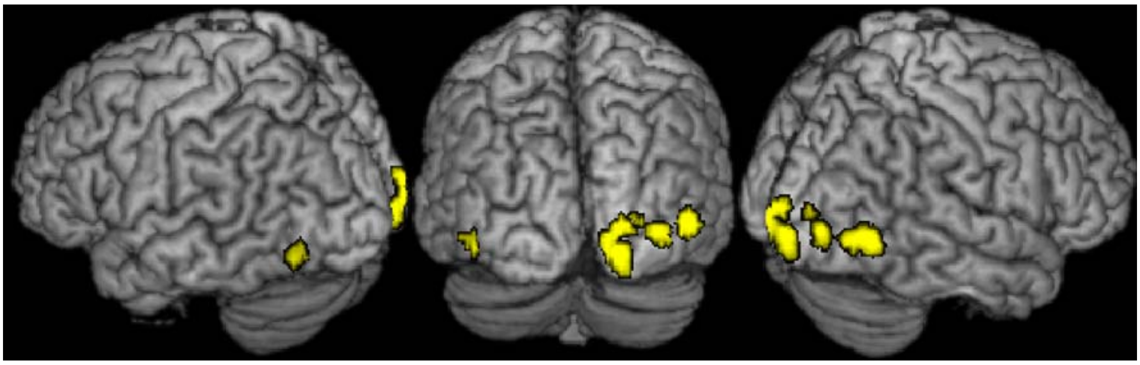


Figure 4.6. Clusters showing larger BOLD response when participants attended to all object features compared to single object features. Initial threshold: α = .01.

Table 4.5. Clusters that were activated when participants attended to whole objects compared to single features.

	MNI				
region	x	y	z	k	Z
attention all > attention colour/shape/motion					
MOG	21	-103	10	87	3.93#
ITG	48	-67	-2	24	3.75@
MOG	-45	-73	-11	12	2.85@

Note. # p < .05, uncorrected. @ p < .05 small volume corrected for multiple comparisons. MOG: middle occipital gyrus. ITG: inferior temporal gyrus. Initial threshold: α = .01.

4.4.4. fMRI results: functional connectivity

The results of the GLM suggest that a common network of brain areas is involved in processing shape and motion of objects. However, a network cannot only be inferred on the basis of brain areas that happen to be activated by the same stimulus. If these brain areas belong to a common network it is expected that they ‘communicate’ with each other. Previous studies (Haynes et al., 2005; Macaluso et al., 2000; Schultz et al., 2008) have suggested that such communications can be measured as functional connections between brain areas. These functional connections are measured as correlations in the activity (measured as the change in BOLD signal) in the brain areas of interest that cannot be explained by the GLM (i.e., the residual BOLD activity) (Haynes et al., 2005; Macaluso et al., 2000; Schultz et al., 2008). In other words, the residual BOLD activity is the error between the activity predicted by the GLM and the measured BOLD response. The residual BOLD activity in one brain area is correlated with the residual BOLD activation of other clusters using the Pearson correlation. The resulting correlations reflect functional connectivity between clusters that is independent of the experimental paradigm.

Correlations in residual activity were computed separately for blocks on which participants attended to shape or motion. A lenient significance criterion of $\alpha = .05$, uncorrected for multiple comparisons, was accepted. For the brain areas activated when participants attended to shape or motion Figure 4.7 and Figure 4.8 show that medium to strong correlations between the occipito-temporal areas, areas in the posterior parietal cortex along the IPS, and the frontal areas were found.

The results suggest that the occipito-temporal areas, areas in the posterior parietal cortex along the IPS, and the frontal areas are functionally connected which supports the evidence that these areas are organised in a network. Similar correlations were found between areas activated when participants attended to colour (see Chapter 9.2, Table 9.3). Importantly, there were no correlations between the residuals of areas activated by colour with any of the areas activated by shape or motion. The lack of correlations in the residuals between areas processing shape or motion with colour areas extends the evidence for processing of colour information in a separate network.

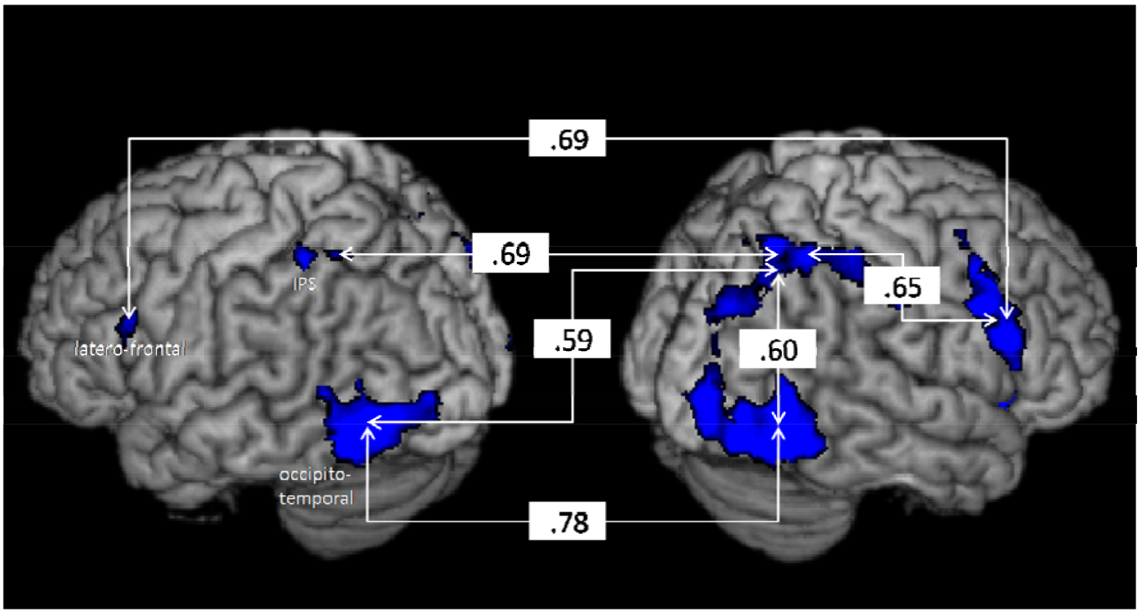


Figure 4.7. Pearson correlation coefficients (N = 16; p < .05, uncorrected for multiple comparisons) between the residuals of brain areas that were activated when participants attended to motion compared to shape and colour.

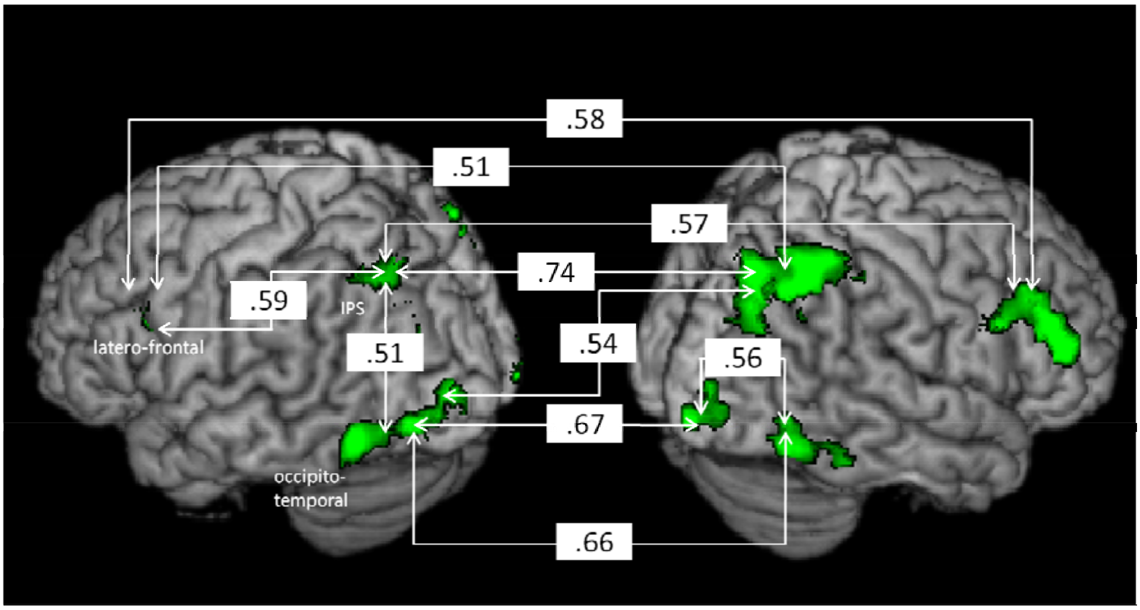


Figure 4.8. Pearson correlation coefficients (N = 16; p < .05, uncorrected for multiple comparisons) between the residuals of brain areas that were activated when participants attended to shape compared to motion and colour.

4.5. Diffusion Tensor Imaging

The fMRI experiment suggested a common network of lateral frontal, areas in the posterior parietal cortex along the IPS, and occipito-temporal areas that is involved in recognising object shape and object motion. Further, correlations in residual BOLD signal between these brain areas suggest functional connectivity between these areas. This DTI study explores the network at the anatomical level. In particular, it is investigated if the anatomical connections between the parts of the network are relevant for object recognition performance.

4.5.1. DTI data acquisition

DTI data were acquired in the same scanning session as the fMRI data. Diffusion-weighted images were acquired in 64 directions isotrophicly sampled on the unit sphere. In addition, there was one non-diffusion weighted scan at the end of the DTI session. The FOV was 270 mm x 270 mm with a resolution of 120 x 124. Fifty-nine axial slices (thickness: 2.11 mm) were acquired during one TR from the bottom to the top of the head. No gap was interleaved between slices. Other scanning parameters were TR = 6108.533 ms; TE = 70.04 ms; flip angle = 90º, diffusion weighting: b = 1000 s/mm^2. The duration of the DTI session was approximately 7.5 min.

4.5.2. DTI data analyses

DTI data were analysed using the Functional Diffusion Toolbox (FDT) implemented in FSL (Smith et al., 2004; Woolrich et al., 2009, http://www.fmrib.ox.ac.uk/fsl/fdt/index.html). The diffusion-weighted images were corrected for eddy current distortions (Leemans and Jones, 2009) with in-house routines. They were also corrected for head motion using the FLIRT (Jenkinson et al., 2002; Jenkinson and Smith, 2001) function in FSL. Then, the robust version of the Brain Extraction Tool (BET, Smith, 2002) was used to extract brain tissue from non-brain tissue (e.g., the skull) for each participant's high resolution T1 scan. BET was also used to extract the brain mask in which no diffusion-weighting was applied. This is a binary image containing ones inside the brain and zeros outside the brain.

This DTI study aims at reconstructing white matter tracts between areas activated in the fMRI experiment. DTI data are acquired and analysed in diffusion space whereas fMRI data were processed in standard space. In order to use the regions identified by the fMRI experiment as seed regions for fibre reconstruction all data had to be transformed to a common space. Using the FLIRT function in the FDT toolbox of

FSL transformation parameters between standard and diffusion space were determined for every participant. Based on inspection we found best registration of images of different spaces when the images were in diffusion space. Therefore, the seed regions that were acquired based on the group level analyses of the fMRI experiment were transformed individually into each participant’s diffusion space.

After the preprocessing FDT was used to reconstruct the white matter tracts of each participant. As a first step, FDT determines the probability distribution of the diffusion direction at each voxel using the BEDPOSTX function. The width of the distribution describes the uncertainty of the diffusion direction at that voxel that is caused by artefacts and noise in the MRI signal. As a second step, the PROBTRACKX function uses these probability distributions to reconstruct white matter fibres. This means that the probability of white matter fibres passing through one specific voxel, a cluster, or white matter fibres connecting clusters can be estimated.

4.5.2.1. Seed regions for probabilistic fibre tracking

Following Kim and Kim (2005) the brain areas that were activated in the fMRI experiment were used as seed regions for the fibre tracking. These seed regions were constructed as follows. Attention to shape and motion activated similar and overlapping regions in the occipito-temporal cortex, the areas along the IPS, and the lateral frontal cortex (Figure 4.3). However, there were small differences in the shape and the size of the activated clusters for attending to shape or motion (Figure 4.3). It was reasoned that if shape and motion indeed activated the same network these differences are unsystematic and do not indicate differences in the activation of functionally segregated brain areas. Therefore, we took the union of the areas activated in the fMRI experiment at the group-level when participants attended to shape or motion (Table 4.3 and Table 4.4). This union resulted in one occipito-temporal, one area along the IPS, and one lateral frontal seed region in each hemisphere. Note that corresponding seed regions in each hemisphere were slightly different across hemispheres in terms of shape and size. Kim and Kim (2005) extended their functional regions a single voxel into the underlying white matter. We did not use this procedure as our seed regions intersected each participant’s white matte (due to the group-level analysis and spatial smoothing at the single subject level) and included white matter that was not at the interface between gray and white matter. Moreover, the probabilistic fibre tracking method we used to reconstruct tracts allowed us to remove spurious tracts reconstructed from gray matter by thresholding (see Chapter 4.5.2.2). Further, the number white matter voxels that overlapped with the seed regions varied between participants.

When participants attended to colour one bilateral cluster in colour sensitive areas was activated (see Figure 4.5). It was assumed that this cluster reflected responses of anatomically distinct colour sensitive regions in each hemisphere. Averaging at the group level and spatial smoothing at the single subject level must have caused the fusion of activation of two distinct clusters. To enter this cluster as a seed region to PROBTRACKX, the cluster was split in a left and a right cluster. Voxels with positive x-coordinates in MNI space were assigned to the cluster that will be referred to as right CoS and voxels with negative x-coordinates were assigned to the cluster that will be referred to as left CoS. Voxels with 0 as x-coordinate were not assigned to any seed region. Please note that the large cluster in the CingC (MNI coordinates: $x = 0$, $y = 50$, $z = -11$, Table 4.2) was not used as a seed region for probabilistic tracking. It was assumed that due to the size of the cluster, its extensions in both the left and the right hemisphere and the nature of the fibre reconstruction procedure used here most parts of the white matter for every participant would have been reconstructed.

4.5.2.2. Probabilistic fibre tracking

FSL uses a probabilistic method in order to reconstruct fibre tracts in the brain. This allows for quantitative measures of fibre tracts by constructing the probability density function of the local mean fibre orientation in each voxel (Behrens et al., 2007; Behrens et al., 2003a; Behrens et al., 2003b). The default parameters implemented in FDT were used for the probabilistic fibre tracking. Five-thousand samples were sent from each voxel in each seed region to initiate fibre tracking. Each voxel across the brain was assigned a counter. Whenever an estimated fibre tract connected two seed regions the counter in each voxel along the fibre tract was increased by 1. This means that the higher the counter of a voxel at the end of the fibre tracking the more estimated tracts had passed this voxel. Therefore, the probability of the existence of a tract connecting two seed regions is higher for voxels with high counter values.

Based on earlier research (Begre et al., 2007; Sasson et al., 2010; Shinoura et al., 2007; Thomas et al., 2009; Thomas et al., 2008) white matter tracts connecting occipital brain areas to frontal areas such as the IFOF, the ILF, and the SLF seemed likely to be involved in object recognition. Therefore, connections between the occipito-temporal and the lateral frontal areas that were activated in the fMRI experiment when participants attended to shape or motion were of particular interest. The results of the residual correlations between the brain areas that were activated in the fMRI experiment suggested that there are functional connections especially between bilaterally activated areas (such as both occipito-temporal areas for example). Therefore, white matter tracts between areas that were activated in both hemispheres were also reconstructed.

Seed regions were entered to probabilistic tracking in pairs. Tracts were reconstructed in between the following pairs of seed regions that were activated when participants attended to shape or motion: occipito-temporal and lateral frontal areas for both hemispheres separately, left cluster along the IPS to right cluster along the IPS, left occipito-temporal area to right occipito-temporal area, and left lateral frontal area to right lateral frontal area. The following pairs of seed regions that were activated when participants attended to colour were used: left CoS to left MFG anterior, left CoS to left MFG posterior, left CoS to left MTG anterior, left CoS to left MTG posterior, and right CoS to right insula (Table 4.2).

For regions that were activated when participants attended to shape or motion the reconstructed fibre tract maps for tracts connecting the occipito-temporal and the frontal seed regions of both hemispheres were thresholded by 1000 to remove paths with low probabilities. This means that only voxels that were passed 1000 times or more during the probabilistic tracking were used for further analyses. In other words, 5000 samples were sent from each voxel in each seed region to estimate tracts that connect to the other seed region. Out of these samples (5000 x number of voxels in the cluster), 1000 or more probabilisitic tracts have to pass a certain voxel for this voxel to be included in the final reconstructed tract. The threshold for the tracts between the areas along the IPS was 800. The threshold for tracts between both occipito-temporal seed regions and both lateral frontal seed regions was 50. Please note that the different thresholds are due to different sizes of seed regions. The smaller the seed region the fewer samples are sent from that seed region and therefore the counters in voxels along estimated tracts are lower as for bigger seed regions.

As the final step in the analysis, the mean FA value of each participant's reconstructed and thresholded fibre tact map was computed.

4.5.3. Results

Based on previous research (Begre et al., 2007; Sasson et al., 2010; Shinoura et al., 2007; Thomas et al., 2009; Thomas et al., 2008) connections between the occipito-temporal and the lateral frontal areas activated in the fMRI experiment were of particular interest as they might be involved in cognitive processes relevant for object recognition. The reconstructed white matter tracts connecting occipito-temporal and lateral frontal brain areas are shown in Figure 4.9 for the left hemisphere and in Figure 4.10 for the right hemisphere. The number of voxels that contained reconstructed fibre tracts for each participant are shown in Table 4.6.

As an initial step, the tracts that were reconstructed for each participant were overlaid with the participant's original FA map to inspect whether the reconstructed tracts corresponded to anatomical tracts. For all participants the reconstructed tracts indeed followed the anatomical tracts, which validated the probabilistic tracking procedure.

Table 4.6. Number of voxels in which white matter fibre were reconstructed after thresholds were applied. Seed regions were clusters that were activated in the fMRI experiment of Experiment 4. Each column reflects a pair of seed regions; each row reflects one participant.

l OT - l fr	r OT - r fr	OT - OT	fr - fr	par - par	l OT - l par	r OT - r par	l par - l fr	r par - r fr
571	1161	451	1508	125	733	2383	1055	2371
384	953	621	1336	776	1265	3377	2039	2742
336	2496	3900	603	1148	698	2524	1386	3955
18	1125	3647	1674	146	1208	2783	1037	3025
262	2843	3041	2460	2279	1975	4631	1467	4164
550	350	2073	31	483	434	4221	1238	2134
317	1726	2061	4210	964	699	3703	656	3265
272	303	1056	117	1322	681	3020	467	1889
1232	1155	1643	3682	2088	1941	5543	1986	2794
404	70	3706	1512	825	316	4399	2927	2532
775	863	3004	1230	2783	2204	5403	1517	3898
1024	907	3566	1240	1543	1421	5144	2578	3599
549	1418	1620	37	1103	868	5177	1971	4952
1151	1253	2371	1463	3036	2615	7693	2469	6976

Note. OT: occipito-temporal; fr: frontal; par: parietal; l: left hemisphere; r: right hemisphere.

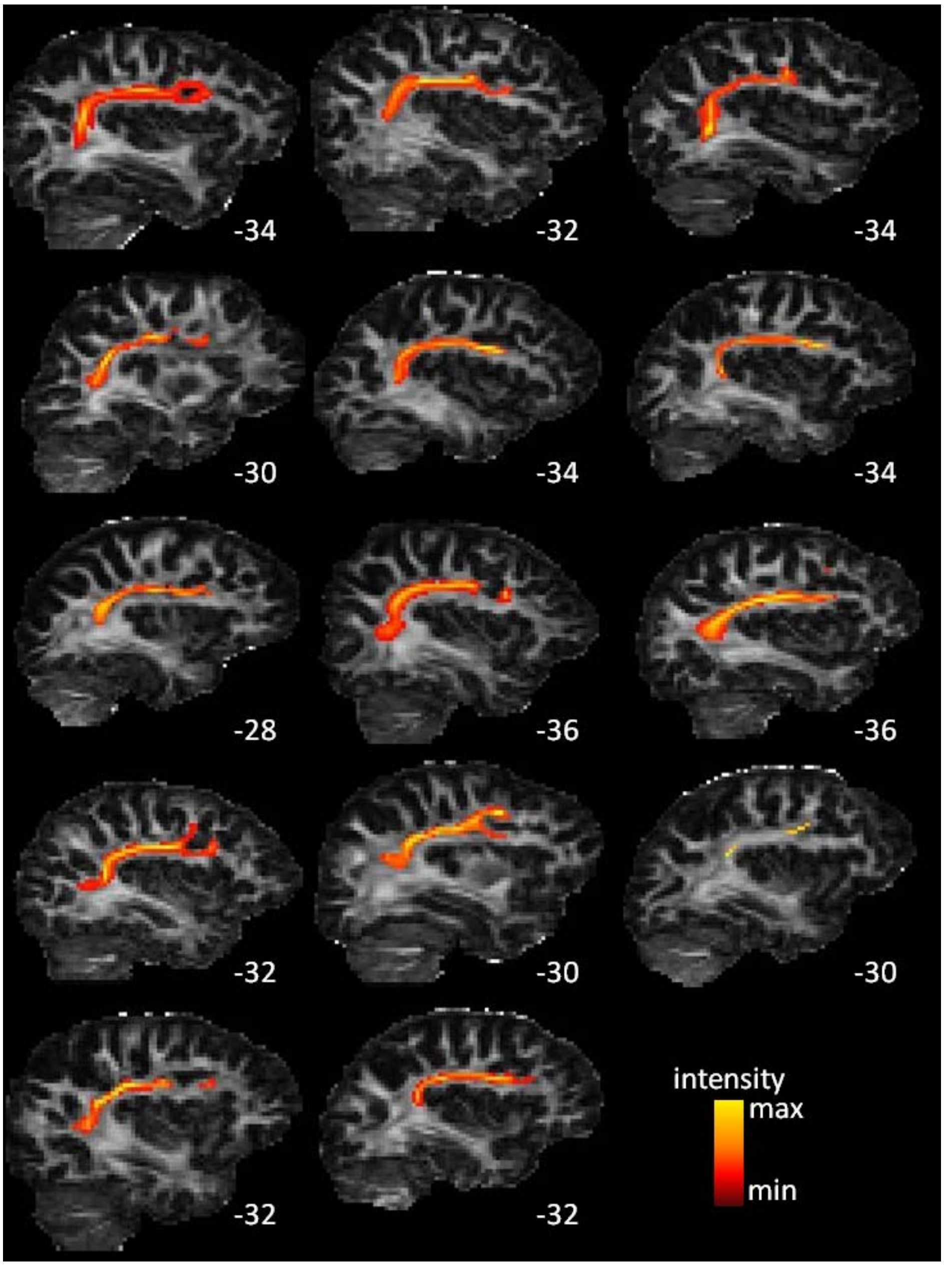


Figure 4.9. Results of the PROBTRACKX analysis for the 14 participants who provided behavioural data. Seed regions were the occipito-temporal and the lateral frontal area in the left hemisphere identified by the fMRI experiment (see text for details). The paths were thresholded with 1000. The white numbers indicate MNI x-coordinates of the presented slice. Intensity refers to the number of estimated tracts that passed through a voxel.

In the left hemisphere the results were consistent across participants. One main tract was reconstructed that expanded from posterior parts of the brain dorsally to the frontal lobe. Using the John Hopkins University atlas tool (Wakana et al., 2004) that is implemented in FSL this tract was identified as the SLF. In one participant (Figure 4.9; 4th row, 3rd column) only small clusters within the SLF were above the threshold of 1000.

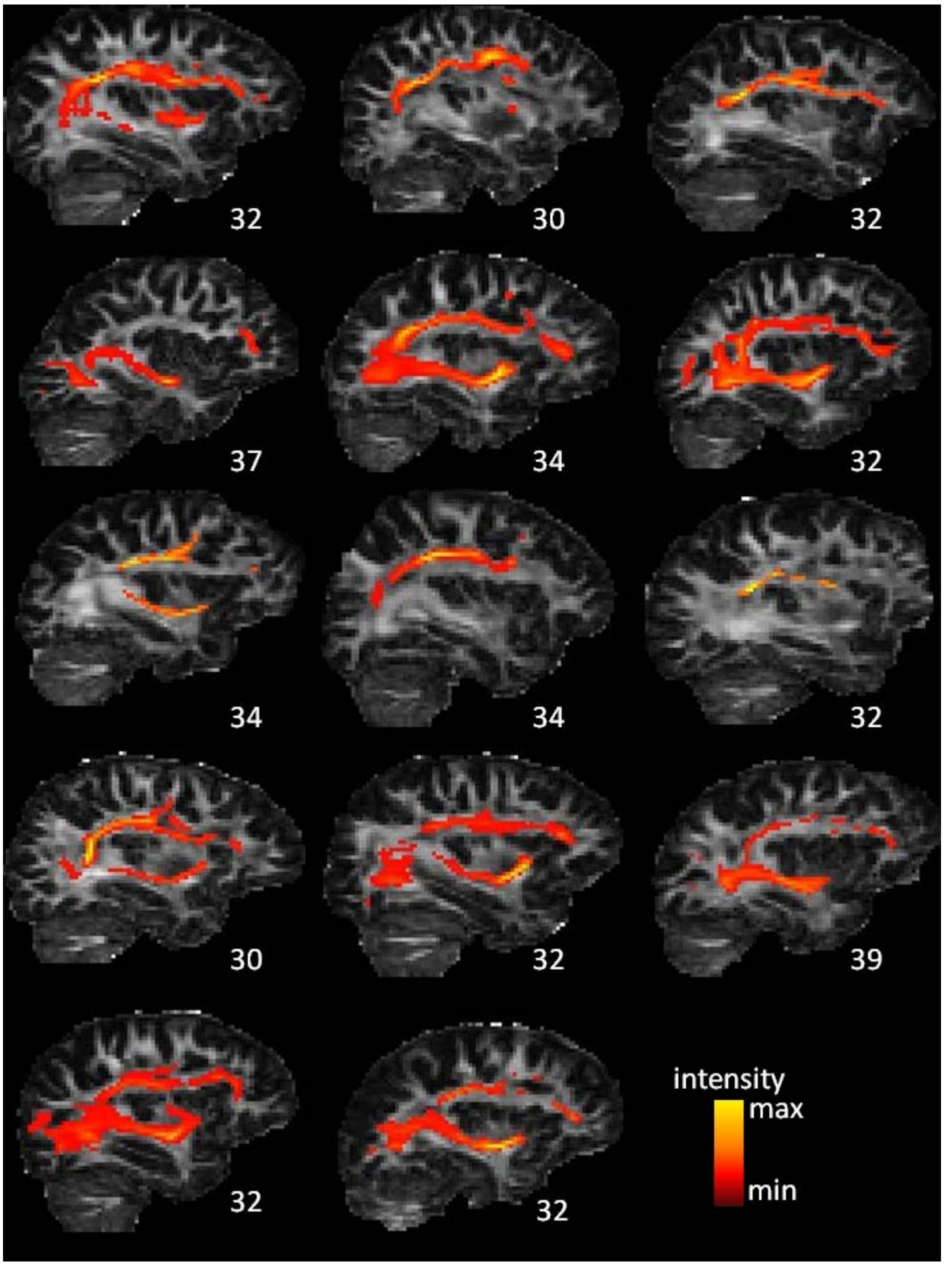


Figure 4.10. Results of the PROBTRACKX for the 14 participants who provided behavioural data. Seed regions were the occipito-temporal and the lateral frontal area in the right hemisphere identified by the fMRI experiment (see text for details). The paths were thresholded with 1000. The white numbers indicate MNI x-coordinates of the presented slice. Intensity refers to the number of estimated tracts that passed through a voxel.

In the right hemisphere the results were less consistent across participants. Consistent with the fibre tracking results of the left hemisphere for 13 out of 14 participants a tract was reconstructed that expanded from posterior brain areas dorsally to the frontal lobe which was identified as the SLF. In nine out of these 13 participants an additional tract was reconstructed that expanded from posterior parts of the brain ventrally to the frontal lobe. This tract was identified as the IFOF. In four participants only the SLF was reconstructed and in one participant only the IFOF was reconstructed.

The number of voxels that was reconstructed between the brain areas that showed larger activation when participants attended to colour compared to shape or motion are presented in Table 4.7.

Table 4.7. Number of voxels of the reconstructed tracts that connect areas that showed larger activation when attending to colour as opposed to attending to shape or motion. Each row represents one participant. The bilateral cluster along the CoS was split in a left and a right part.

l CoS - l MFG anterior	l CoS - l MFG posterior	l CoS - l MTG anterior	l CoS - l MTG posterior	r CoS - r insula
2823	3378	829	5123	2258
8488	3752	1445	3611	27815
3779	4065	241	2230	19921
9575	3483	938	4303	16388
6975	8593	781	8701	19777
368	1151	182	2284	9280
2095	4494	240	7919	9550
2751	2158	192	3832	11079
4868	1670	1337	5740	13097
1968	2642	450	2917	5840
4540	2920	871	3182	10465
1084	724	419	2251	4623
6200	4198	707	2876	12440
4740	3052	870	1888	11503

Note. L: left; r: right; CoS: collateral sulcus; MTG: middle temporal gyrus; MFG: middle frontal gyrus.

The following tracts were reconstructed that connected brain areas that responded when participants attended to colour. Between the left CoS and both of the clusters in the MFG the anterior thalamic radiation was consistently reconstructed in all participants. Additionally, parts of the ILF and the IFOF were reconstructed in some participants. Between the left CoS and both of the clusters in the MTG parts of the SLF, the ILF, and the IFOF were reconstructed. When using the right CoS and the cluster with the peak voxel in the insula the IFOF was reconstructed in all participants.

Due to the choice of seed regions, the nature of the probabilistic fibre tracking, and the thresholding procedure used in this study, the numbers of voxels of the reconstructed tracts varied between participants (see Table 4.6 and Table 4.7). This may limit the explanatory power of the present results, especially with respect to correlations between tract integrity and measurements of performance (see Chapter 4.5.4), because whole fibre tracts are included in further analyses for some participants whereas for other participants only fractions of tracts were included. For the present study, however, it was considered important to apply the same threshold to every dataset in order to avoid arbitrary data modification by choosing individual thresholds for every dataset.

4.5.4. White matter integrity and behavioural response

This study investigated whether fibre tracts connecting brain areas that were activated in the fMRI experiment are relevant for task performance. The mean FA value of reconstructed tacts between areas that were activated in the fMRI experiment of each participant and in each hemisphere was extracted and correlated with the RT and the accuracy for each attention condition separately.

The following correlations were found for seed regions based on the network activated for attention to shape and attention to motion. No consistent patterns were found for correlations between accuracy and the mean FA value of reconstructed tracts based on seed regions activated when participants attended to shape or motion. For RTs, in contrast, a consistent pattern of correlations was found for fibre tracts that were reconstructed between occipito-temporal and lateral frontal areas (see Figure 4.9 and Figure 4.10).

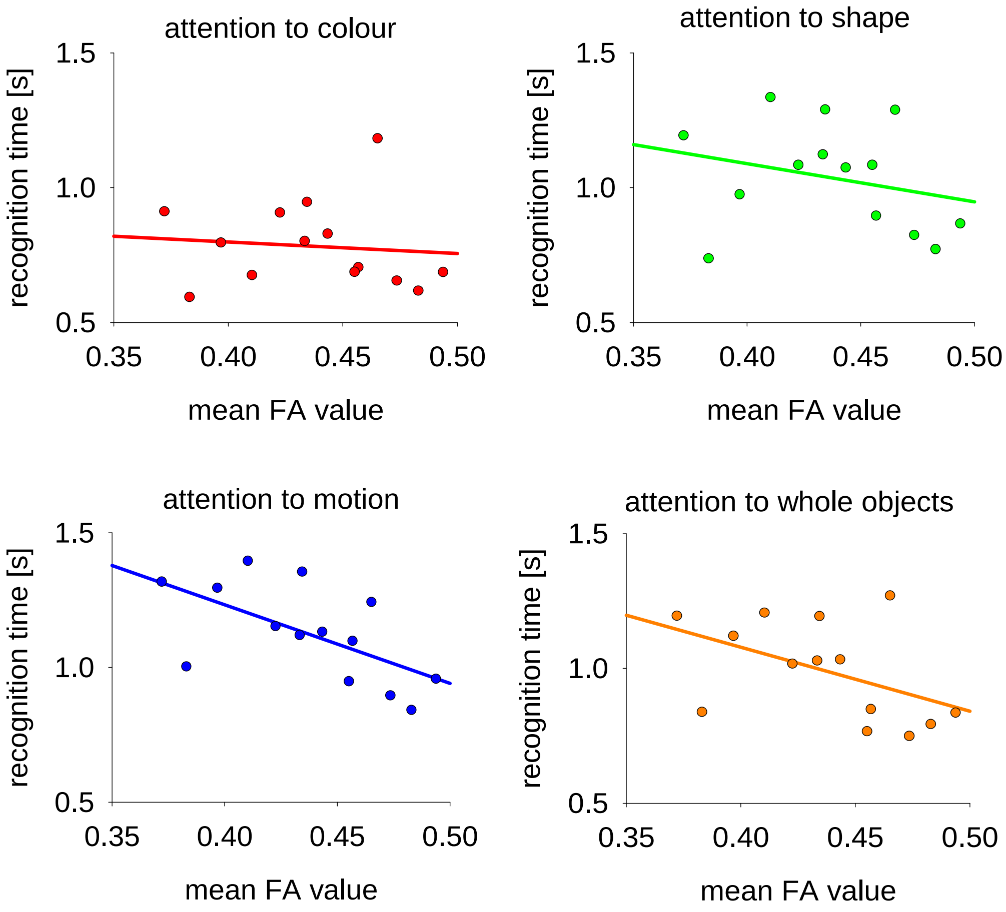


Figure 4.11. Scatter plots with RTs and mean FA value for tracts reconstructed in the left hemisphere between the occipito-temporal and lateral frontal areas that were activated during the fMRI experiment. Lines are linear regression lines. Each dot represents a participant.

Scatter plots with mean FA value of the extracted tracts in the right and left hemisphere and the RT are shown in Figure 4.11 for the left hemisphere and in Figure 4.12 for the right hemisphere.

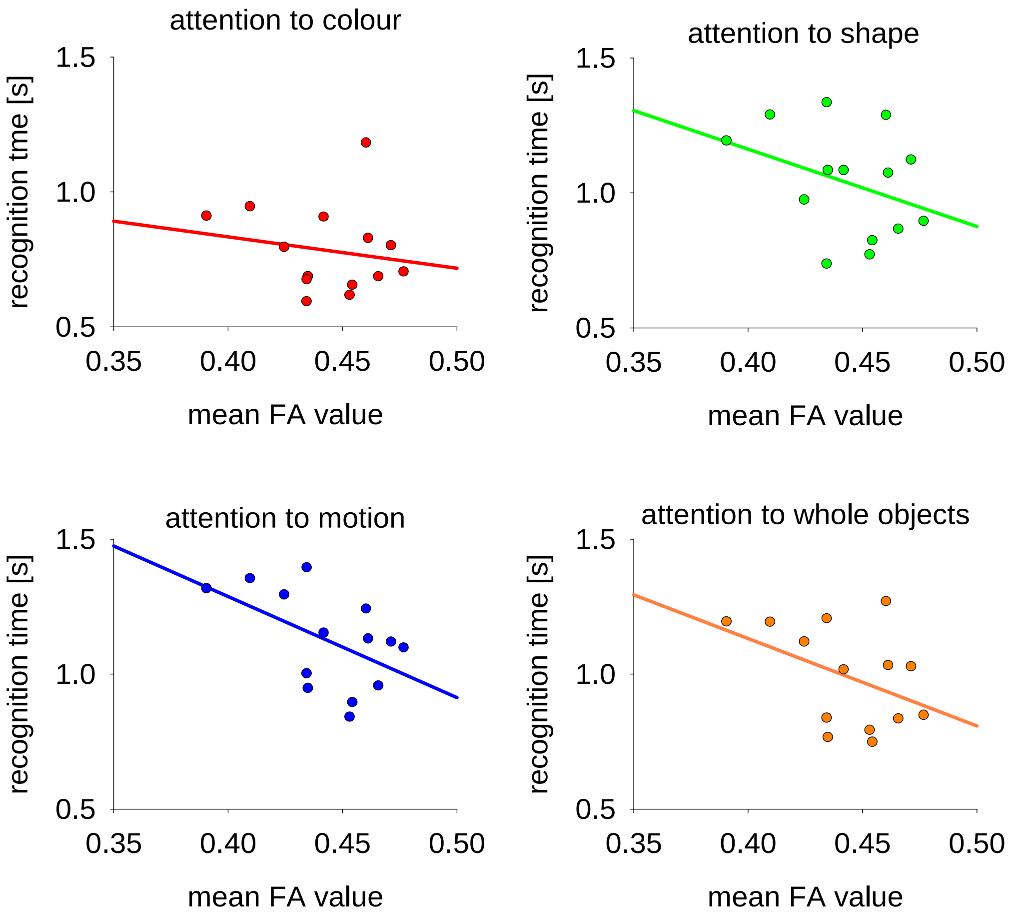


Figure 4.12. Scatter plots with RT and mean FA value for tracts reconstructed in the right hemisphere between the occipito-temporal and lateral frontal areas that were activated during the fMRI experiment. Lines are linear regression lines. Each dot represents a participant.

The correlation coefficients are shown in Table 4.8 for FA with RTs and in Table 4.9 for FA with accuracies. The pattern of correlations showed that for attention to shape, motion, and all features there were small to medium negative correlations between FA value and RT. This indicated that the integrity of the reconstructed fibre tracts was related to the behavioural performance. The correlations of the RTs for colour with the mean FA values were smaller than for the other attention conditions indicating that this network was not crucial for processing of object colour.

The correlations of the FA values of fibre tracts connecting parts of the network that were activated when participants attended to colour with RTs and accuracies are shown in Table 4.9 and Table 4.10, respectively.

Table 4.8. Pearson correlation coefficients of RT and mean FA value for tracts reconstructed in each hemisphere between the areas that were activated during the fMRI experiment when attending to shape and motion.

	colour	shape	motion	all
left OT - left frontal	-0.01	-0.27	-0.61	-0.48
right OT - right frontal	-0.18	-0.35	-0.51	-0.43
right OT - right parietal	0.16	0.58	0.29	0.24
right parietal - right frontal	0.20	-0.24	-0.46	-0.19
left parietal - right parietal	0.17	-0.02	0.32	0.25
left OT - right OT	0.29	0.05	0.13	0.28
left frontal - right frontal	0.01	-0.01	-0.06	-0.02
left OT - left parietal	0.27	0.20	0.51	0.35
left parietal - left frontal	-0.12	-0.14	-0.44	-0.42

Note. All: attention to whole objects. OT: occipito-temporal. Note that p < .05, uncorrected for N = 14 is r > .53.

Table 4.9. Pearson correlation coefficients of accuracy and mean FA value for tracts reconstructed in each hemisphere between areas that were activated during the fMRI experiment when attending to shape and motion.

	colour	shape	motion	all
left OT - left frontal	-0.12	-0.23	-0.02	-0.10
right OT - right frontal	-0.05	-0.30	-0.27	-0.16
right OT - right parietal	-0.16	-0.54	0.15	-0.11
right parietal - right frontal	-0.17	0.03	-0.56	-0.28
left parietal - right parietal	-0.52	-0.25	-0.29	-0.37
left OT - right OT	-0.57	-0.24	-0.81	-0.62
left frontal - right frontal	-0.01	0.01	-0.07	-0.03
left OT - left parietal	-0.35	-0.23	-0.13	-0.15
left parietal - left frontal	0.40	0.18	-0.13	0.10

Note. p < .05, uncorrected for N = 14 is r > .53. All: attention to whole objects. OT: occipito-temporal.

Table 4.10. Pearson correlation coefficients of RT and mean FA value for tracts reconstructed in each hemisphere between areas that were activated during the fMRI experiment when attending to colour.

	colour	shape	motion	all
left CoS - left MFG ant.	0.11	0.09	0.19	0.19
left CoS - left MFG post.	0.37	0.33	0.25	0.34
left CoS - left MTG ant.	0.06	-0.08	-0.01	-0.06
left CoS - left MTG post.	0.14	0.49	0.42	0.42
right CoS - right insula	-0.20	0.06	-0.04	-0.10

Note. p < .05, uncorrected for N = 14 is r > .53. CoS: collateral sulcus; MFG: middle frontal gyrus; MTG: middle temporal gyrus; all: attention to whole objects; ant: anterior; post: posterior.

Table 4.11. Pearson correlation coefficients of accuracy and mean FA value for tracts reconstructed in each hemisphere between areas that were activated during the fMRI experiment when attending to colour.

	colour	shape	motion	all
left CoS - left MFG ant.	0.00	-0.18	-0.18	-0.28
left CoS - left MFG post.	-0.12	-0.40	-0.20	-0.38
left CoS - left MTG ant.	0.06	0.15	-0.06	-0.2
left CoS - left MTG post.	-0.34	-0.45	-0.09	-0.32
right CoS - right insula	0.11	-0.29	-0.04	-0.06

Note. $p < .05$, uncorrected for $N = 14$ is $r > .53$. All: attention to whole objects; ant: anterior; post: posterior.

The correlations between the FA values of the tracts connecting regions that were activated when participants attended to colour and measurements of behavioural performance were unsystematic. There was a moderate correlation between the RT on colour trials and the integrity of tracts that connect the left CoS with posterior parts of the MFG indicating that this tract might be involved in colour processing. However, the integrity of other connections between parts of the colour processing network (left CoS to posterior parts of the MTG) showed stronger correlations with RTs for attention to shape, motion, and whole objects compared to attention to colour. Therefore, the correlations between white matter integrity of the reconstructed tracts with behavioural performance did not reveal any effects specific to colour processing.

4.6. Discussion

This study combined fMRI to identify brain areas involved in object recognition and analyses of functional connectivity with DTI to investigate how objects that are defined by multiple features are processed in the brain. A feature attention task was used to investigate which brain areas process shape, colour, and motion of objects. A network of brain areas was identified that responded to both object shape and motion. Functional connectivity analyses supported the idea of a network consisting of lateral frontal areas, posterior parietal areas along the IPS, and occipito-temporal areas. The existence of a common network processing object shape and motion was further supported by reconstructing fibre tracts that connect these brain areas using DTI. The integrity of the reconstructed tracts correlated with task performance when processing shape, motion, and whole objects. The correlations indicated that the reconstructed tracts are relevant for object recognition. Object colour activated a different set of brain areas. The LOC was found to integrate different object features.

Same and different trials largely activated the same brain areas for each of the attention conditions. For motion, however, stronger BOLD activity was found on different trials. One possible explanation might be neural adaptation mechanisms (Grill-Spector and Malach, 2001). What is unclear is why these adaptation mechanisms were present when participants attended to motion only as no such differences were found when participants attended to shape or colour. There was no evidence at the behavioural level for differences between same and different trials when participants attended to motion as opposed to other features.

Consistent with previous literature (Peuskens et al. 2004; Schultz et al. 2007) Experiment 4 found similar BOLD activation patterns when participants attended to shape and motion. The network activated when attending to shape or motion consisted of occipito-temporal areas, areas in the posterior parietal cortex along the IPS, and lateral frontal areas. The occipito-temporal areas most likely include the shape sensitive LOC (see Grill-Spector et al., 2001) and the motion sensitive hMT/V5 (see Culham et al., 2001).

The activity found in posterior parietal areas along the IPS when processing shape and motion is consistent with the findings of studies using SFM objects. The posterior parietal cortex was found to be active when observers judge the 3D shape of SFM objects (Kriegeskorte et al., 2003; Paradis et al., 2001). Further, Schultz et al.'s (2008) results showed that the posterior parietal cortex encoded parametric shape changes (e.g., the amount of tapering of a shape) of other rotating objects. This indicated that the activity in the posterior parietal cortex was not exclusively due to observers using the rotation of the objects to gain information about the 3D shape of the objects but that areas along the IPS are actively involved in processing of geometric shape. In line with this finding, in Experiment 4 the motion of the objects did not provide any cue that could facilitate shape recognition. The contrary was the case. The objects underwent non-rigid deformation which changed the shape on every frame of the animation making it harder to extract the exact object shape. Thus, the activation of the posterior parietal cortex when participants attended to shape provided further evidence for its involvement in processing dynamic objects beyond recovering 3D shape from motion.

The areas in the posterior parietal cortex along the IPS were also activated when participants attended to motion. This is in line with the results of Peuskens et al. (2004) who found that the same brain areas along the dorsal pathway were activated when participants attended to shape and motion.

The finding that both attention to shape and attention to motion activated areas in the posterior parietal cortex along the IPS might not only be due to processes that are involved in extracting information about object shape. Kawasaki et al. (2008) found that the activity in the posterior parietal cortex reflected memory load when different numbers of features are stored in working memory. The activity in the posterior parietal cortex found in Experiment 4 could have been caused by the 1-back matching task that participants had to perform. Experiments 1 to 3 suggested that shape and motion are represented in an object-based manner. The activity in the posterior parietal cortex found in Experiment 4 might reflect that participants not only memorised attended shape or motion but memorised information about unattended features as well. The results of the behavioural experiments showed that shape and motion cannot be processed independently. This fits with the finding of Experiment 4 that on block on which participants attended to colour no activation was found in the posterior parietal cortex. The lack of activity in posterior parietal areas could indicate that when participants attended to colour they did not memorise information about shape or motion leaving the working memory load lower as when several object features had been memorised.

Previous research (Cabeza et al., 2003; Kanwisher and Wojciulik, 2000; Miller, 2000) showed that lateral frontal brain areas are involved in cognitive processes that are relevant for object recognition such as working memory, visual attention, and cognitive control. In line with Schultz et al. (2008) who found that lateral frontal areas were involved in processing rotating objects Experiment 4 showed that both attending to motion and attending to shape activated lateral frontal brain areas which might reflect processes of memorising the relevant feature.

A quite different pattern of activity was found when participants attended to colour. Colour sensitive areas on the ventral surface of the occipital cortex along the CoS were activated. Additionally, large clusters in medial brain areas were activated. Experiments 1 to 3 suggest that the representations of shape and motion are more closely linked compared to the representation of shape or motion with colour. The results of the present fMRI study provide a neural correlate for this finding. The smaller influence of unattended colour might be due to the fact that colour activates a distinct network of brain areas whereas motion and shape activate a common network.

Similar to Schultz et al.'s (2008) results Experiment 4 found that areas within the network that was activated by attention to shape and attention to motion were functionally connected to each other. This was shown by correlations in the residual activity between areas belonging to the network (Haynes et al., 2005; Macaluso et al., 2000). These correlations indicate that the brain areas were not only co-activated in the experiment but that some form of communication took place between these areas. This is an important finding to support the idea of a network of brain areas that are involved in processing dynamic objects. Please note that the residual correlations occurred systematically between areas that were activated when certain a feature was attended to. No such correlations were found between areas that were activated when attending to shape or motion and areas that were activated when attending to colour. This supports the idea of specialised networks to process certain object features. Although residual correlations do not allow for any inferences about the hierarchical organisation of the areas belonging to the network such as feedforward and feedback connections between the areas, residual correlations do show that the areas work together in order to achieve recognition of dynamic objects.

Experiment 4 not only identified a network of brain areas that processes dynamic objects at the functional level it also provided evidence for the existence of this network at the anatomical level. Probabilistic fibre tracking reconstructed white matter connections between the parts of the network that responded to both object shape and object motion. Fibre integrity of the reconstructed tracts between occipito-temporal and lateral frontal areas in each hemisphere correlated with RTs showing that the better the integrity of the tract the faster the RTs when participants attended to shape, motion, and whole objects. The tracts were identified as the SLF in the left hemisphere and SLF and IFOF in the right hemisphere. This finding fits the results of other DTI studies that showed that the SLF was involved in working memory tasks (Begre et al., 2007) and that the IFOF is involved in face recognition (Thomas et al., 2009; Thomas et al., 2008), both processes that are related to object recognition. Experiment 4 extends the findings of the previous studies and suggests that these tracts are involved in object recognition. Importantly, the correlations of fibre integrity with RTs for attention to colour were smaller than for the other conditions suggesting that the integrity of the reconstructed tracts is not crucial for processing of object colour. This finding is in line with the finding of the fMRI study of Experiment 4 that showed that colour activated a different network.

Experiment 4 indicated that the SLF in the left hemisphere and the SLF and the IFOF in the right hemisphere are involved in object recognition. This asymmetry suggests that object recognition or at least sub-processes of object recognition are lateralised in the brain. This finding is in accordance with other DTI studies (e.g., Bohr et al., 2007; Thomas et al., 2008; Tuch et al., 2005) that used different tasks and found lateralisation in the correlations of white matter integrity and task performance. Further research will be necessary to explore lateralisation of white matter tracts that are involved in object recognition.

Overall, systematic correlations between FA and performance measurements were only found for RTs and less so for accuracy. The lack of systematic correlations between accuracy and FA value might have been due to the little variance in the data as the overall accuracy was very high.

The present study showed that attending to object features increased the activity in brain areas that are sensitive to the attended features. To perceive an object instead of a summary of separate features the information about different object features must be combined. To investigate this, a condition was included in which participants attended to whole objects instead of object features. Areas in lateral and posterior occipital cortices were activated. These areas are most likely part of the shape sensitive LOC instead. This is consistent with the findings of Self and Zeki (2005) who reported that parts of the LOC are involved in the integration of colour and motion. The results of Experiment 4 extend Self and Zeki's (2005) findings. Experiment 4 confirms the LOC as a brain area that integrates features for object recognition. In Self and Zeki's (2005) study motion and colour had to be integrated in order to recover shape information. This was not necessary in Experiment 4 as the objects were constructed in a way that ensured that the features were independent of each other. Therefore, the results of Experiment 4 show that the LOC is not only involved in integrating multiple features for shape recovery but that the LOC integrates features in order to create holistic representations of objects.

4.7. Conclusion

This study provides further evidence for a cortical network consisting of occipito-temporal areas, lateral frontal areas, and areas in the posterior parietal cortex along the IPS for processing of shape and motion of objects as reported by Schultz et al. (2008) and Peuskens et al. (2004). Object colour was found to be processed in a different network consisting of medial brain areas and colour sensitive areas at the ventral surface of the occipital cortex. Experiment 4 provides a neural correlate for the results found in Experiments 1 to 3 showing that the representations of object shape and object motion are combined whereas the links to the representation of colour are less strong. Further, the results suggest that the LOC plays a central role in integrating shape, colour, and motion of objects. Importantly, this study provides evidence for an object recognition network that is not only based on co-activation of brain areas during object recognition tasks. Functional and structural connections between parts of the network were identified. As revealed by correlations between RTs and the integrity of white matter of the SLF and the IFOF these tracts play an important role in the processing of object shape and object motion.

5. Experiment 5: The role of unattended object features for shaping the BOLD response to novel dynamic objects

5.1. Abstract

The results of Experiments 1 to 3 suggest a combined representation of shape and motion. Colour might be represented separately with strong interconnections to the representations of the other object features. Even when participants attended to one object feature unattended object features were automatically processed as well suggesting object-based object representation. The present fMRI study aims at investigating the neural correlates underlying the processing of unattended object features. The Garner paradigm was adapted for neuroimaging and combined with an fMR-adaptation paradigm. Participants attended to shape, colour, or motion separately while unattended features changed from trial to trial (change block) or remained constant (no change block) throughout an experimental block. Adaptation was expected to reduce BOLD activity elicited by unattended object features on no change blocks. Overall, no brain area was identified that processed unattended object features. Psychophysiological interaction analyses, however, revealed that shape, colour, and motion sensitive brain areas modulated anterior parts of the LOC to a larger extend on change blocks compared to no change blocks. The findings indicate that modulating processes between parts of the object recognition network might play an important role in integrating information about unattended object features.

5.2. Introduction

Humans are able to attend to different features of objects separately to interact with the environment more efficiently. This especially helps when searching for a target object. Selectively attending to the features that distinguish the target object from other objects allows for faster search as cognitive resources are only used to process relevant information.

While selectively attending to certain object features can help to interact with the environment more efficiently other object features are still perceived. The behavioural results provided by Experiments 1 to 3 indicated that unattended object features affect recognition of attended features. The present study investigates the influence of unattended object features for the BOLD response to novel objects. The aim is to gain further knowledge on how the BOLD activation pattern in the brain changes depending on which object feature is attended to and how processing of unattended features is reflected in the activity pattern of the brain.

Behavioural studies report different results on whether object features can be processed independently of each other (Cant et al., 2008; Ling and Hurlbert, 2004; Wegener et al., 2008). Overall, Experiments 1 to 3 suggest object-based representation of shape and motion which means that these features are bound at a pre-attentive stage. Previous fMRI studies provide evidence that brain areas processing unattended object features increase their activity when any feature of the object is attended to (O'Craven et al., 1999; Paradis et al., 2008).

The effect of unattended or task-irrelevant stimuli was not only found in overall increases in activity in certain brain areas but also in the way that brain areas modulate each other's activity. Bingel et al. (2007) for example, found that the impairment of visual object recognition caused by task-irrelevant painful stimulation was related to the way in which subcortical (pain processing) brain areas modulated the activity in visual areas. Similar effects were reported between visual areas. Friston et al. (1997) showed that the modulation of V5 by V1 was stronger when participants attended to motion compared to when no instruction to attend to motion was given.

At the neural level, Experiment 4 showed that there was a large-scale network for processing shape and motion, and a different, relatively independent network for colour processing. So far, this thesis only tested the effects of the unattended features indirectly (i.e., by measuring the effects of attention to a particular feature on recognition performance and BOLD response). In this case, we found that there was a separate network for colour although there was a trend for colour to affect performance in Experiments 1 and 2 (see Figure 3.7 and Figure 3.8) and RTs in Experiment 3. These results provided an opportunity to test the possibility that unattended features can modulate activations in a given brain region (Bingel et al., 2007; Friston et al., 1997; O'Craven et al., 1999; Paradis et al., 2008). This modulation could account for the effects of colour on object recognition in Experiment 2 and 3 and in previous work (Naor-Raz et al., 2003; Price and Humphreys, 1989; Tanaka and Presnell, 1999). Thus, Experiment 5 investigated how the BOLD response in networks or brain areas that are involved in object recognition is modulated by unattended object features. To achieve this aim, we adapted an attentional paradigm that allowed us to directly measure any modulatory effects of unattended features (Garner, 1988; Gottwald and Garner, 1975; see also Cant et al., 2008).

In Experiment 5 the activity elicited by unattended object features is identified by using an adaptation of the Garner paradigm (Garner, 1988; Gottwald and Garner, 1975). A Garner paradigm is a way of assessing whether different features (e.g., colour and shape) of a stimulus are processed independently or whether processing one feature automatically activates processing of other features. Independent processing of two features of a stimulus can be inferred when the performance on a task involving one feature is not impaired by changes in the task irrelevant feature. Originally, the Garner paradigm (Garner, 1988; Gottwald and Garner, 1975) was used to test independent processing at the behavioural level. We assumed that a similar reasoning would be possible at the BOLD level as well.

To test whether unattended object features affected the BOLD response pattern we used fMR-adaptation (Grill-Spector and Malach, 2001) combined with a Garner task (Garner, 1988; Gottwald and Garner, 1975). fMR-adaptation refers to a mechanism in which the BOLD response to a visual stimulus decreases when the same stimulus is presented repeatedly. In the present study participants were asked to separately attend to one of the three object features. On half of the experimental blocks the unattended features changed from trial to trial (e.g., participants attended to colour, while shape and motion varied from trial to trial). On the other half of the blocks the unattended features were the same throughout the block. For the blocks on which unattended features are the same throughout the block it is expected that fMR-adaptation decreases the BOLD activity in brain areas that are selective to these unattended features. In other words it is assumed that when participants attend to one feature, the areas that are sensitive to either of the other two features show increased BOLD activity on change blocks in which no adaptation is expected.

Alternatively, processing of unattended object features might be reflected in the way in which brain areas that process certain object features interact with each other. As reported by Bingel et al. (2007), for example, task-irrelevant stimulation (such as pain in their study) can influence the functional connectivity between brain areas by changes in the modulatory influence of one area on the activity in another area. Experiment 4 showed that the LOC integrates object features. In the present study the fMR-adaptation might lead to differences in the information that brain areas that are sensitive to certain object features send to the LOC as a function of whether unattended features changed or not. To test for this possibility, psychophysiological interaction (PPI) analyses (Friston et al., 1997) were computed in addition to analyses comparing activity differences across the whole brain.

BOLD responses to objects differ depending on which object feature is attended to. Experiment 4 identified a network consisting of occipito-temporal areas, areas in the posterior parietal cortex along the IPS, and lateral frontal areas that responded to shape and motion of objects. Colour activated a different network consisting of medial cortices and ventral parts of the occipital cortex. The activity on the ventral surface of the occipital cortex was consistent with the location of colour sensitive regions that were identified by other studies (e.g., McKeefry and Zeki, 1997). In the present study participants were asked to separately attend to shape, motion, and colour of objects. Based on the results of previous research (Paradis et al., 2008; Peuskens et al., 2004; Schultz et al., 2008) it is expected that attending to shape and motion activates a common network consisting of occipito-temporal, parietal, and frontal areas. Attending to colour is expected to activate colour sensitive areas on the ventral surface of the occipital cortex such as the CoS (Cant and Goodale, 2007).

5.3. Methods

5.3.1. Participants

Twelve volunteers (4 male, 8 female; average age in years: M = 31, SD = 6, age range: 24 to 42 years) participated in this study. All participants were naive to the purpose of the study and gave informed consent.

5.3.2. Stimuli

The same set of stimuli as for Experiment 2, 3, and 4 was used.

5.3.3. Design

The experiment was set up as a 2 x 3 within-subjects design with block type (change, no change) and attended feature (colour, shape, motion) as repeated measures. An fMRI block design was used in which the six experimental conditions were run in separate blocks. The six blocks were presented twice in a single functional run. The 12 blocks were grouped into four sets. Within each set, each attended feature block was presented once and one of the two block type conditions was presented twice. Both the order of the attended feature conditions and the repeated block type were randomly determined for each set. In addition to the six experimental blocks, a 12 s fixation block was presented. The fixation block was presented at the beginning of each functional run and after each set of three experimental blocks, for a total of five fixation blocks (note that a functional run therefore began and ended with a fixation block). There were two functional runs each lasting approximately 7 min in the scanning session.

5.3.4. Procedure

Participants were instructed outside the scanner and ran through a set of practice blocks of the experiment ensuring that every participant had practised every experimental condition. In one experimental block, participants watched a sequence of eight objects. They were instructed to attend to one feature (colour, shape, or motion) and to perform a 1-back feature matching task on the attended feature (see also Schultz and Lennert, 2009). They pressed a button of an MRI compatible response box (LumiTouch™) with the left thumb whenever the attended feature was identical on two consecutive trials (target trials). Unlike Experiment 4 no response was required on non-target trials.

Figure 5.1 (attend colour, change block) and Figure 5.2 (attend colour, no change block) show the sequence of events on an experimental block. The block began with a 2 s instruction (the word: ‘colour’, ‘shape’, or ‘motion’) presented at the centre of the screen, which indicated the attended feature for the block. The instruction was followed by a 0.7 s blank screen. Each object was then shown for 2.5 s at the centre of the screen, each followed by a 0.7 s blank screen. The presentation duration was less than a full cycle of motion for the objects. Based on our behavioural data, there was a sufficient amount of time to judge all three features. In total, an experimental block lasted 23 s.

On change blocks (Figure 5.1), the presented objects were selected in a way that ensured that the unattended features were different for consecutive objects. This means, for example, that for an attend motion change block two consecutive objects were never blue (unless a target trial was presented, see below).

On no change blocks (Figure 5.2), unattended features were the same throughout the block. This means, for example, that on an attend shape no change block all eight objects were red with motion profile 1.

For each participant, the 64 objects were used as equally often as possible. Across participants, the 64 objects occurred in all six experimental as equally often as possible.

There were two target trials in each experimental block. On target trials, the repeated object matched the preceding object on all three features. This was necessary to keep target trials on change and no change blocks equivalent. Participants were asked to respond as soon as they detected a target trial. They could respond at any time during the block. A response was counted as correct if it was made during the 2.5 s presentation of the repeated object on a target trial. Misses (i.e., not responding on a target trial) and responses made on any other occasions were counted as errors. No feedback was provided.

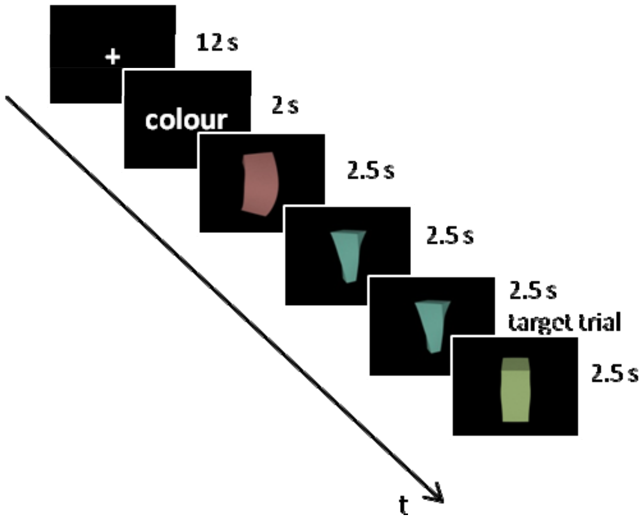


Figure 5.1. Illustration of a block on which participants attend to colour. This block is a change block as motion and shape are different from trial to trial (apart from target trials on which all three features match). Note that the 700 ms blanks between the objects are not shown in this figure for clarity.

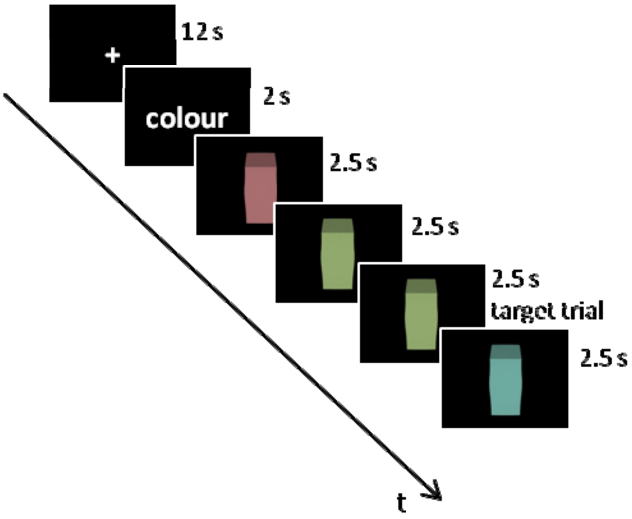


Figure 5.2. Illustration of a block on which participants attended to colour. This block is a no change block as motion and shape are the same throughout the block. Note that the 700 ms blanks between the objects are not shown in this figure for clarity.

5.3.5. Image acquisition

All participants were scanned at the Newcastle Magnetic Resonance Centre. In the scanner, head motion was restrained with foam pads that were placed between the head and the head coil. Scanning sessions started with the acquisition of a T1 high resolution anatomical scan lasting approximately 5 min. After that, two functional runs were presented. Scanning sessions closed with a low-resolution T1-weighted scan.

Anatomical T1-weighted images and functional T2*-weighted EPIs were acquired from a 3 T Phillips Intera Achieva MR scanner using a Philips 8-channel receive-only SENSE head coil. The high resolution T1-weighted scan consisted of 150 images. The FOV was 240 mm x 240 mm x 180 mm with a matrix size of 208 x 208 pixels. One voxel was 1.15 mm x 1.15 mm x 1.2 mm in size. The low resolution T1-weighted scan consisted of 27 slices. Each slice was 3 mm thick. There was a gap of 1 mm between slices. The FOV was 192 mm x 192 mm x 107 mm with a matrix size of 128 x 128 pixels. The parameters of the EPIs were: TR = 1.92 s, TE = 40 ms, flip angle = 90°. The FOV was 192 mm x 192 mm x 107 mm with a matrix size of 64 x 64 pixels. Twenty-seven axial slices were acquired from the bottom to the top of the head during 1 TR. Each voxel was of 3 mm x 3 mm x 3 mm in size. There was a 1 mm gap between slices. For the experiment, a total of 414 functional images were acquired (207 in each run). Before each functional run, preparation steps were performed to allow for equilibration of the T1 signal.

5.3.6. fMRI data preprocessing

Functional images were realigned to the first image of each participant and resliced to correct for head motion. These images were normalised to a standard MNI EPI T2*-weighted template with a resampled voxel size of 3 mm x 3 mm x 3 mm (Friston et al., 2005). They were then spatially smoothed with a 6 mm FWHM Gaussian kernel to improve the signal to noise ratio and to allow for comparisons across participants. For the single-subject analysis, we performed additional preprocessing of the functional data and then estimated how well each experimental condition explained the observed BOLD signal. To remove low-frequency drifts in the signal, we applied a high-pass filter with a cutoff of 128 s. We also applied an autoregressive model (AR(1)) to estimate serial correlations in the data and adjust degrees of freedom accordingly.

5.3.7. fMRI statistical analyses

The preprocessed data were analysed with the SPM5 package (http://www.fil.ion.ucl.ac.uk/spm, Friston et al., 1994). SPM5 implements the GLM framework. As for Experiment 4, within the GLM, we used a two-step mixed-effects approach. First, a fixed-effects model was used to analyse each participant's data set. Second, a random-effects model was used to analyse the individual datasets at the group level. No additional smoothing of the images was used at the group level.

The design matrix for the GLM was constructed as follows. Block onsets and block durations for each of the six experimental conditions (2 block types x 3 attended features) and the fixation block were modelled as boxcar functions. These boxcar functions were then convolved with a canonical HRF. The HRF implemented in SPM5 is the difference of two gamma functions. In addition to these regressors of interest, the six movement parameters (roll, yaw, pitch and three translation terms) acquired during realignment, the time period in which the instructions were presented, and a constant term for each session were included in the design matrix as regressors of no interest. A linear combination of the regressors was fitted to the BOLD signal to estimate the beta-weights (i.e. the contribution of each regressor to the data) for each regressor in order to minimise the residual error.

Contrast images of beta-weights were computed for the main effect of block type, the main effect of attended feature, and interactions of block type and attention. For block type, we took the contrasts of beta-weights for change blocks and no change blocks to identify voxels that showed larger activation on change blocks as opposed to no change blocks. For each attended feature, we computed the contrast for the attended feature versus the weighted average of the two unattended features. For example, to test for voxels with a larger BOLD response during the attention to colour blocks, we used the contrast 1 for the beta-weight of the colour regressor and -0.5 for both the beta-weight of the shape and the motion regressor (so that the contrast weights summed to zero).

For the group-level analysis, SPM T maps were then created to allow for statistical inferences. These maps were based on one-sample t-tests of the contrast images at each voxel. Unless stated differently, $\alpha = .05$ corrected for multiple comparisons at the cluster level was accepted for all statistical tests. The minimal cluster size was set to 10 voxels.

The information about unattended object features might not be reflected in a separate brain area but in changes in the activity in the network of different brain areas. In order to test for differences in the physiological connectivity between different brain areas depending on changes in unattended object features we conducted a PPI analysis (Friston et al., 1997). PPI means that the contribution of a brain area (physiological variable) to the activity of another brain area changes with the experimental paradigm (psychological variable) (Friston et al., 1997). In other words PPIs identify regions that have context dependent connectivity with a region of interest (ROI). In this experiment we were interested in how the modulating influence of ROIs on functionally connected areas changed as a function of whether change or no change blocks were presented. It is important to note that PPIs do not provide any information about the anatomical structure underlying an interaction. They only provide information about interactions in functional connectivity between different areas in the brain.

In a PPI analysis, the time course of the activity in each ROI (the seed regions for the PPI) is deconvolved to get the neural signal without the HRF (Gitelman et al., 2003). The PPI regressor was computed as the product of the neural signal in the seed region and a vector coding 1 for change blocks and -1 for no change blocks. This product was then reconvolved with the HRF implemented in SPM5 (see Figure 5.3 for a schematic overview of steps included in PPI analyses). The movement parameters obtained in the GLM and the original eigenvariate of the BOLD signal were included in the PPI as regressors of no interest. As for the GLM, the PPI analyses consisted of two steps, one at the single subject level and one at the group level. The PPI model implemented in SPM5 (Friston et al., 1997) was run on each participant's dataset. The interactions obtained show the increase (or decrease) in functional connectivity between the seed region and target regions depending on the experimental condition. Contrast images were computed. Of interest were the contrasts that had higher connectivity on change trials compared to no change trials. The reasoning behind this is as follows. If a region receives input from a target region processing unattended features the connectivity between both regions should be stronger on change blocks. On no change blocks it is assumed that the input will be minimal as the activity in the region is decreased due to BOLD adaptation mechanisms. The contrast images were submitted to a GLM analysis at the group level. This created SPM-T maps that can be used to make statistical inferences. Tests for significance were one-sample t-tests testing which beta-values are larger than zero.

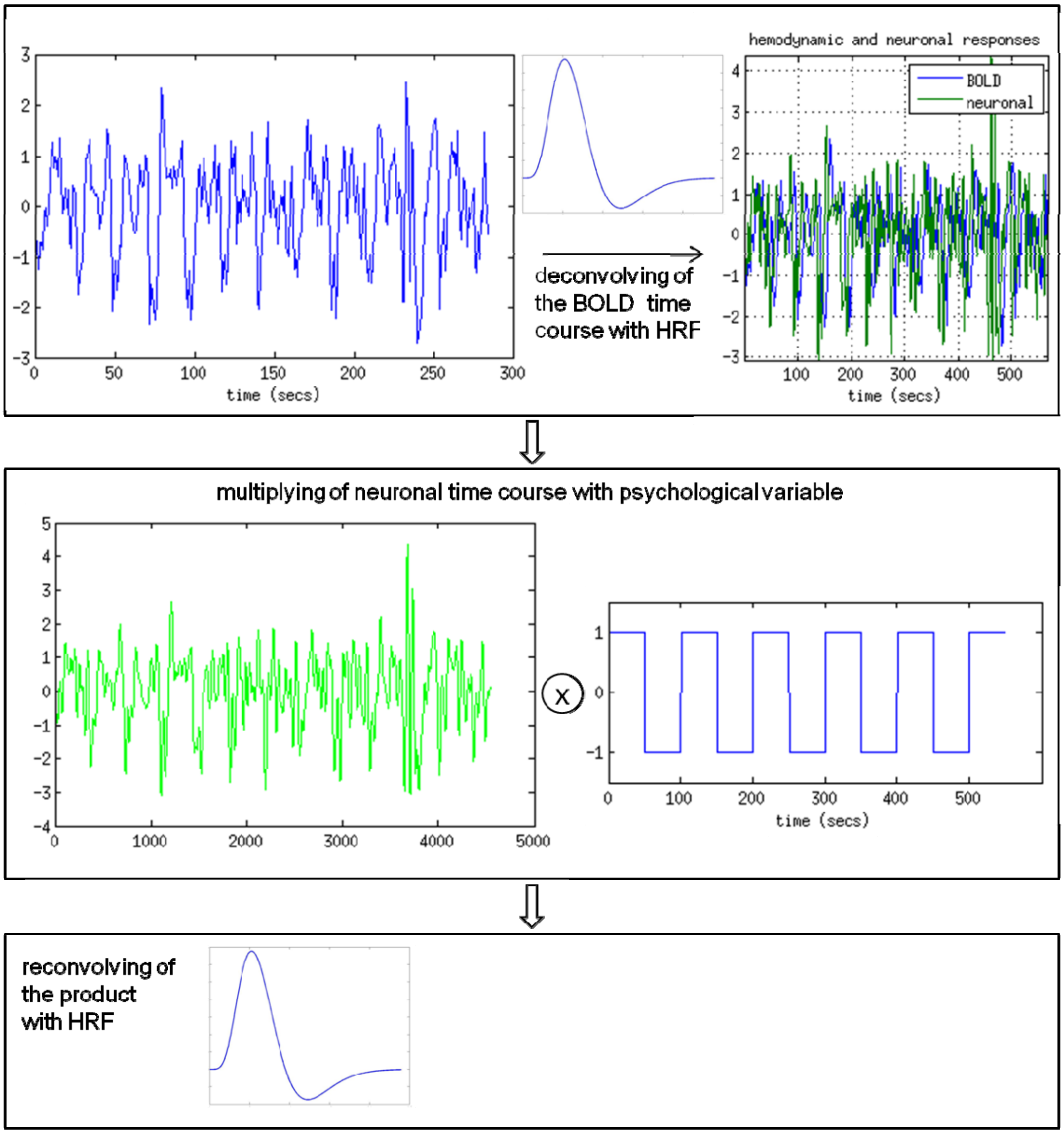


Figure 5.3. Overview of the steps included in PPI analyses. BOLD: blood oxygenation level dependency; HRF: hemodynamic response function.

5.4. Functional localising of brain areas processing shape, colour, or motion

Experiment 4 and 5 investigate BOLD responses to objects that consist of shape, colour and motion. Therefore, it is expected that brain areas that are sensitive to these object features are involved in the tasks presented here. To obtain a measurement of which brain areas are potentially involved in processing objects consisting of different features functional localiser scans were conducted that identified shape, colour, and motion sensitive areas independently of the experimental paradigm. The advantage of this approach is that these independently defined areas can be used as references for the location of a certain function in the brain such as shape sensitive areas. This can help for the interpretation of the activation pattern found in other studies and allows for comparison of the results across different studies. The functional localiser scans were acquired during the same scanning session as Experiment 5.

5.4.1. Motion sensitive brain areas

As reviewed in Chapter 2.4.1.2 motion sensitive areas are expected in middle temporal areas where the hMT/V5 complex is located (e.g., Malach et al., 1995; Zeki et al., 1991; Tootell et al., 1995). In order to functionally localise brain areas that show larger activation when presented with dynamic visual input as opposed to static visual input we presented participants with an adaptation of standard hMT/V5 localiser scans based on previous studies (e.g., Malach et al., 1995; Zeki et al., 1991; Tootell et al., 1995). Most hMT/V5 localisers used simple stimuli that were stationary in one condition and moved in another condition. Brain areas that showed larger response during the condition with motion are considered motion sensitive areas. Tootell et al. (1995) for example used a random dot field rendered against a black background. The dots contracted, expanded, rotated, or remained stationary.

5.4.1.1. Design

The hMT/V5 localiser was set up as an fMRI block design. There were three conditions: static, dynamic, and fixation. Each participant was presented with one run of the hMT/V5 localiser in which 158 volumes were acquired.

5.4.1.2. Stimulus

The stimulus used to localise motion sensitive areas is shown in Figure 5.4. The stimulus consisted of a white dot rendered in the centre of the screen and a random dot field. The dot in the centre was slightly bigger than the dots of the random dot field. The appearance of random dots was restricted to an area that was between circles of approximately 2° and 8° visual angle around the centre dot. On static blocks the participants were presented with a static random dot field. On dynamic blocks the dots of the random dot field expanded and contracted within the restricted area.

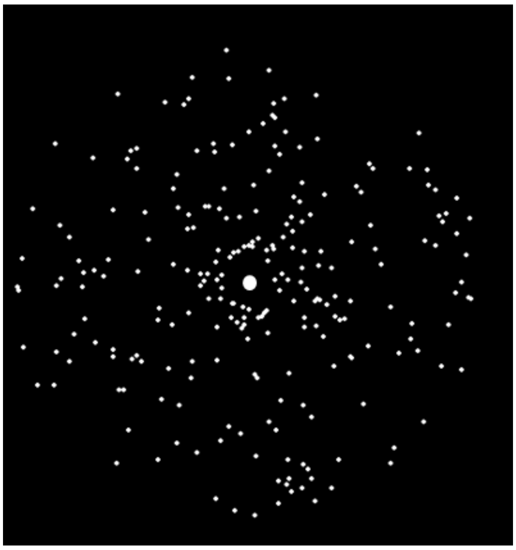


Figure 5.4. Random dot field used to localize hMT/V5. In the dynamic condition the dots of the random dot field expanded and contracted.

5.4.1.3. Procedure

Participants were instructed outside the scanner. Each of them ran through practice trials to ensure they had understood the task. They were asked to fixate on the dot in the centre of the display. They made a button press with their left thumb every time the saw the dot in the centre turning green. Participants were presented with eight dynamic blocks and eight static blocks. These were grouped in four sets of two static and two dynamic blocks. Each block lasted 15 s. Within a set the order of static and dynamic blocks was alternated. Two sets started with a static block and two sets started with a dynamic block. Three to four times within a block the dot in the centre turned green. These times were the same for all participants. At the beginning of the localiser and after each set of four blocks there was a fixation condition where participants remained fixation on a white ‘+’ rendered against a black background for 12 s. The hMT/V5 localizer took approximately 5 min.

5.4.2. Shape and colour sensitive areas

The functional localisers for shape and colour sensitive areas were combined in one run. Following Malach et al. (1995) the localiser for the shape sensitive LOC consisted of blocks of images of different objects and of blocks in which scrambled versions of these images were presented. To localise colour sensitive areas a localiser similar to the one used by Zeki et al. (1991) was presented. In their study colour sensitive areas were localised by presenting participants with coloured Mondrian patches and gray-scale versions of the same images. The brain areas that responded more strongly to the coloured Mondrian images were considered colour sensitive areas.

5.4.2.1. Design

The combined localiser for shape and colour sensitive brain areas was set up as an fMRI block design. Two conditions were included to localise shape sensitive areas; and two conditions were included to localise colour sensitive areas. For shape sensitive areas those were objects and scrambled objects; for colour sensitive areas those were Mondrian images and gray-scale versions of the Mondrian images. Additionally, a fixation condition was included leading to a total five conditions.

5.4.2.2. Stimuli

To localise shape sensitive areas blocks of coloured photographs of everyday objects and blocks of scrambled versions of the same images were presented. These photographs were rendered against a random noise background. The size of the images varied slightly. The scrambled images were created by randomly changing the location of the pixels of the photographs of objects. The images of the objects and the corresponding scrambled versions are shown in Figure 5.5.

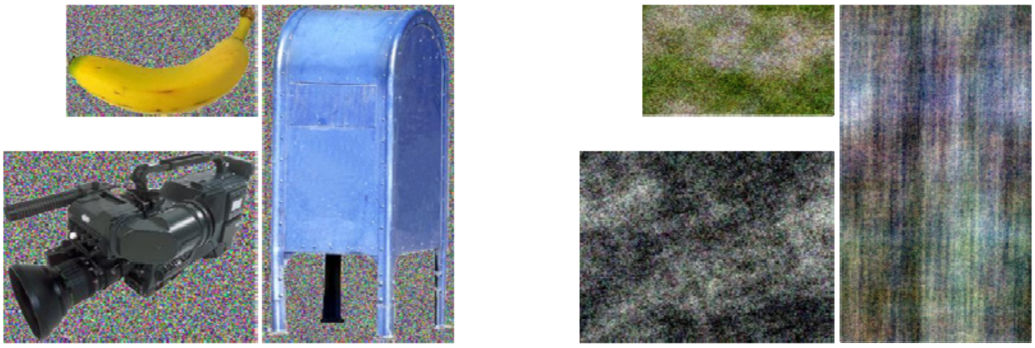


Figure 5.5. Stimuli used to localise shape sensitive areas. Left panel: intact objects; right panel: scrambled versions of the same images.

To localise colour sensitive areas Mondrian images consisting of coloured rectangles that partly occluded each other rendered against a gray background were presented. Additionally, gray-scale versions of the same images were presented. The fixation condition was a white ‘+’ rendered against a black background. Stimuli are shown in Figure 5.6.

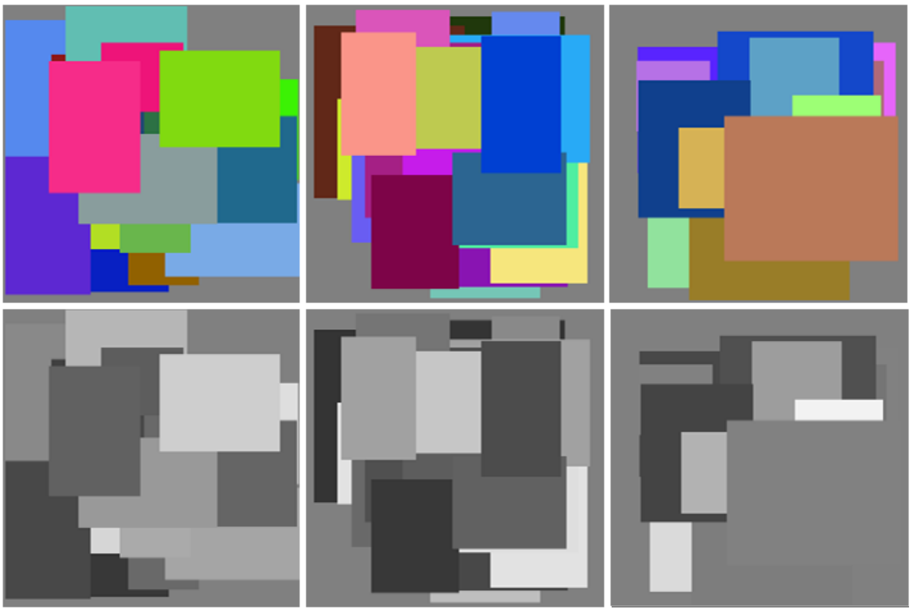


Figure 5.6. Mondrian images (top row) and gray-scale versions of the same images (bottom row).

In the localiser for shape and colour sensitive brain areas the location of the centre of each image was randomly located within a square of 50 pixels around the centre of the screen to prevent retinal after-effects.

5.4.2.3. Procedure

Participants were instructed outside the scanner. Each of them ran through practice trials to ensure they had understood the task. In the scanner, participants used the left thumb to press a button on an MRI compatible response box to respond to target trials. They performed a one-back matching task. A target trial was present when the same image was presented twice in a row. The order of the blocks was the same for all participants and is presented in Figure 5.7. The images of each category were presented in random order. The combined localiser for shape and colour sensitive brain areas took approximately 7 min.

Figure 5.7. Order in which the conditions of the combined localiser for shape and colour sensitive areas were presented. Gray: fixation, red: photographs of objects, yellow: scrambled objects, blue: coloured Mondrian images, green: gray-scale Mondrian images.

5.5. Results

5.5.1. Results: functional localisers

The data of the functional localisers were analysed with the same GLM approach as the data of Experiment 5 (see Chapter 5.3.7).

5.5.1.1. Localiser results: motion sensitive areas

Contrasts were set up in the way that brain areas that showed larger BOLD response during dynamic blocks compared to static blocks were identified. Three clusters were activated for this contrast (Table 5.1). Two of them were located bilaterally in the posterior parts of the occipital lobes and extended to more anterior ventro-lateral areas of the occipital cortex. In the right hemisphere a cluster at the junction of the occipital and the temporal cortex was activated as well. The more posterior regions might reflect motion selectivity in the early visual cortex (see Chapter 2.4.1.2). The occipito-temporal cluster in the right hemisphere is most likely hMT/V5. The results are shown in Figure 5.8.

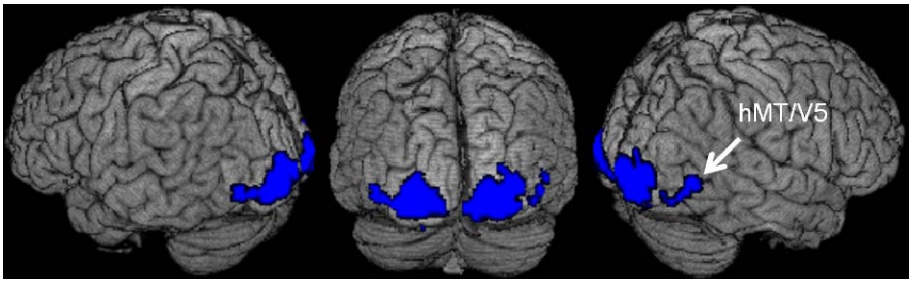


Figure 5.8. Brain areas that showed stronger activation when dynamic dot fields were presented as opposed to static dot fields.

Table 5.1. Coordinates (in millimetres) of the peak voxel of the localiser experiments.

	MNI				
brain area	x	y	z	k	Z
dynamic > static					
IOG	-27	-93	-6	239	4.99*
Cun	27	-99	-6	245	4.80*
IOG	39	-72	-9	56	3.98*
thalamus	24	-30	-3	13	3.83+
objects > scrambled					
PG	30	-45	-12	962	5.07*
cerebellum	-33	-45	-27	1053	6.19*
colour > gray-scale					
LG	18	-90	0	58	4.08*
FFG	-33	-75	-18	27	4.22*
MOG	33	-78	-15	51	4.25*
FFG	-36	-54	-21	13	3.58+

Note. * p < .0001 FDR corrected for multiple comparisons at the cluster level. + p < .05 FDR corrected for multiple comparisons at the cluster level. IOG: inferior occipital gyrus, Cun: cuneus, PG: parahippocampal gyrus, LG: lingual gyrus, FFG: fusiform gyrus, MOG: middle occipital gyrus.

5.5.1.2. Localiser results: shape sensitive areas

Contrasts were set up in the way that brain areas that showed larger BOLD response during object blocks compared to blocks with scrambled objects were identified. Bilateral activity was found in big clusters of the occipito-temporal cortex (Figure 5.9 and Table 5.1). This big cluster might reflect the whole ventral pathway which is associated with processing of object identity. Please note that even though the peak voxel of the activated cluster in the left hemisphere was found in the cerebellum it is inferred that this cluster reflects activity in the shape sensitive LOC as the cluster is expanding dorsal with respect to its peak voxel (Table 5.1).

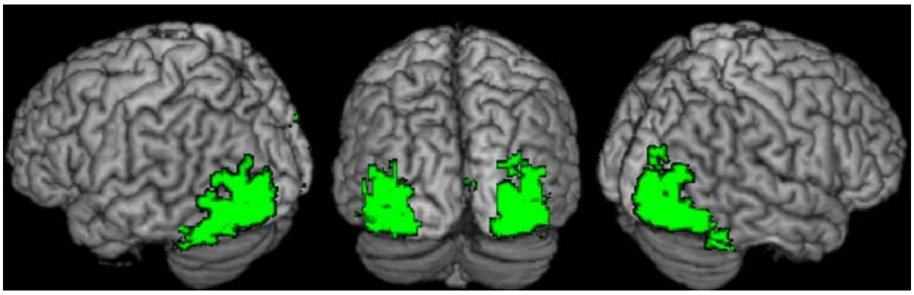


Figure 5.9. Brain areas that showed larger activation on object blocks compared to blocks in which scrambled versions of objects were presented.

5.5.1.3. Localiser results: colour sensitive areas

Contrasts were set up in the way that brain areas that showed larger BOLD activity during blocks with coloured Mondrian images compared to blocks with gray-scale Mondrian images were identified. Bilateral activity was found in areas along the CoS. The results are shown in Figure 5.10 and Table 5.1.

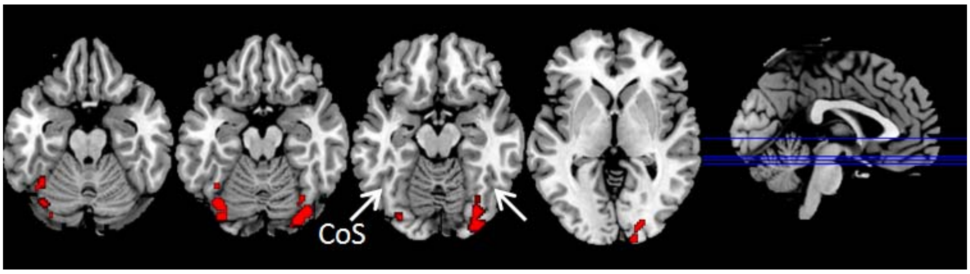


Figure 5.10. Brain areas that responded more strongly to coloured Mondrian images compared to gray-scale Mondrian images. Arrows point to the collateral sulcus (CoS). The slices are MNI z-coordinates -21, -18, -15, and 0.

5.5.2. Behavioural results

RTs and accuracies were analysed to investigate differences between the experimental conditions. Due to a technical failure behavioural results were only available for 11 out of 12 participants. The results are shown in Figure 5.11.

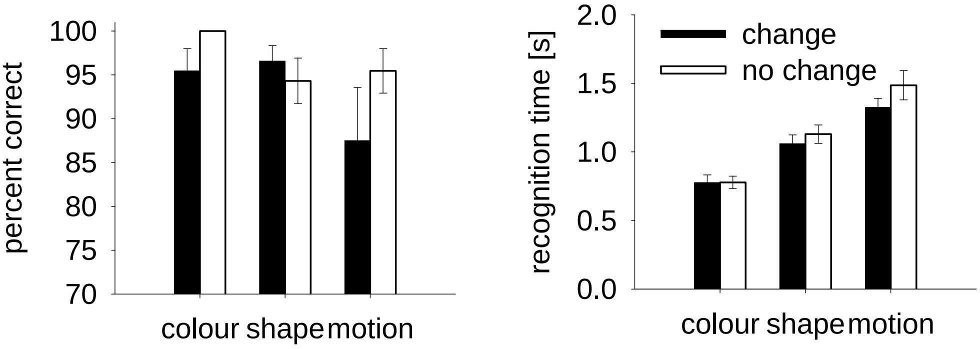


Figure 5.11. Behavioural results of Experiment 5. The accuracy is shown in the left plot; the RTs are shown in the right plot. N = 11.

As in Experiment 4, no outlier trimming procedures were used because responses were only counted as correct when they occurred during the presentation time of the target object (2.5 s). There were no differences in accuracy between the different attended features and whether unattended features changed or not (all ps > .11). For RTs there was a main effect of attended feature (F(2, 20) = 35.86; MSE = 0.06; p < .0001; η_p^2 = .78). Tukey post hoc tests revealed that RTs on attention to colour trials were shorter compared to RTs on attention to shape and attention to motion trials. RTs on attention to shape trials were shorter than RTs on attention to motion trials. There were no differences in RTs depending on the block type. The interaction of block type and attended feature was not significant (all remaining ps > .14).

5.5.3. General linear model (GLM)

We used the GLM to identify main effects and interactions of the two experimental factors (attended feature and block type).

To identify brain areas that showed stronger activation when participants attended to a certain object feature we searched for voxels in which the difference of the beta-weights for attending to a feature and the beta-weights for attended to the other two features was significantly greater than zero. Table 5.2 and Table 5.3 show the MNI coordinates of the peak voxel of clusters that showed significant increase in BOLD activity when one feature was attended to. When attending to shape, one cluster in the right middle occipital gyrus showed significant greater change in BOLD response compared to attention to colour or motion (Table 5.2; Figure 5.12). This cluster is most likely part of the shape sensitive LOC. The peak voxel is located more posterior than found with the functional localiser Table 5.1.

Numerous clusters showed larger activation when participants attended to the motion of the objects compared to colour and shape (Table 5.2 and Table 5.3; Figure 5.13). In each hemisphere a cluster in the posterior temporal lobe extending to lateral occipital areas was activated (Table 5.2) which is likely to be hMT/V5. This is consistent with the more anterior cluster that the functional localiser scan identified in the right hemisphere (Table 5.1). Other clusters were activated when participants attended to motion were found along the IPS and in lateral frontal areas. This is consistent with the results of Experiment 4 and the results of Peuskens et al. (2004). Further clusters that responded to motion are listed in Table 5.3.

Similar to the attention to motion condition, attention to colour compared to motion and shape activated a large set of clusters across the brain (Table 5.2 and Table

5.3; Figure 5.14). The colour sensitive CoS was activated bilaterally (Table 5.2), similar to the results of the functional localiser scan for colour sensitive areas (Table 5.1). Further activation in the occipital lobe was found in the cuneus bilaterally. In the left hemisphere the activated cluster was located at the posterior ventral parts of the cuneus (Table 5.2). This part of the cuneus had previously been found to be colour sensitive (Cant and Goodale, 2007). In the right hemisphere the anterior dorsal parts of the cuneus showed activation (Table 5.2 and Table 5.3). Consistent with Experiment 4 strong activation was found in medial brain areas when participants attended to colour.

In Experiment 4 we found that attention to motion and attention to shape activated largely overlapping clusters. At first glance, the results of the present study seem at odds with this. To further explore the results for attention to shape the initial threshold for significance was lowered to $\alpha = .01$ and clusters with significance at uncorrected p values were accepted (Table 5.4). These clusters were overlaid with the results for attention to motion (Figure 5.15). As shown in Figure 5.15, lowering the significance criteria for the attention to shape condition revealed similar areas as the ones found in Experiment 4. As in Experiment 4, both attention to shape and attention to motion activated posterior parietal and occipito-temporal brain areas. With the lower threshold attending to shape activated two clusters in the occipital cortex of the right hemisphere; one at the posterior part and one more ventrally at the occipito-temporal junction. In the left hemisphere only the ventral occipito-temporal cluster was activated. The activation found in areas along the IPS in Experiment 4 was replicated in the right hemisphere only. In contrast to Experiment 4 and the attention to motion condition of the present study no frontal activation was found for attention to shape. Comparing Figure 5.14 and Figure 5.15 reveals that colour activated a network of brain areas that was distinct from the network activated by shape and motion.

Table 5.2. Activity in brain areas that were expected to be sensitive to the attended feature.

brain area	MNI x	y	z	k	Z
attention shape > motion/colour					
MOG ~	33	-96	3	18	4.18+
attention motion > attention shape/colour					
MOG ~	48	-63	-9	198	4.80*
ITG ~	-42	-69	-6	90	4.10*
MFG	27	0	57	33	3.98*
MFG	-30	-3	51	37	3.85*
IFG	54	21	3	33	3.92*
IFG	45	33	-6	24	3.98+
IFG	39	6	36	45	4.13*
IFG	57	9	18	16	3.55+
IPL	39	-33	39	22	3.84+
IPS	33	-75	30	14	4.25#
attention colour > attention shape/motion					
CoS ~	24	-72	-6	39	3.68*
CoS ~	-18	-81	-9	35	4.34*
Cun ~	-3	-96	-6	21	3.75+

Note. * p < .0001, corrected at the cluster level. + p < .05, corrected at the cluster level. # p < .05, uncorrected. MNI coordinates are in mm. k: number of voxels in the cluster. ~ coordinates used for seed regions in the PPI. MOG: middle occipital gyrus, ITG: inferior temporal gyrus, MFG: middle frontal gyrus, IFG: inferior frontal gyrus, IPL: inferior parietal lobule, IPS: intraparietal sulcus, CoS: collateral sulcus, Cun: cuneus.

Table 5.3. Additional clusters that were activated in Experiment 5.

brain area	MNI x	y	z	k	Z
attention to motion > attention to shape and colour					
STG	66	-45	15	28	4.06+
PCG	63	-24	33	40	4.41*
cerebellum	33	-72	-27	123	4.57*
cerebellum	-36	-60	-30	167	4.69*
attention to colour > attention to shape and motion					
MeFG	3	63	9	307	4.64*
aCun	15	-90	27	21	3.40+
PreCun	-3	-72	33	19	3.84+
MTG	-48	-75	27	82	3.97*
ITG	-33	-30	-12	16	3.88+
MFG	-24	51	21	29	3.79+

Note. * p < .0001, corrected at the cluster level. + p < .05, corrected at the cluster level. MNI coordinates are in mm. k: number of voxels in the cluster. STG: superior temporal gyrus, PCG: postcentral gyrus, MeFG: medial frontal gyrus, aCun: anterior parts of the cuneus, PreCun: precuneus, MTG: middle temporal gyrus, ITG inferior temporal gyrus, MFG: middle frontal gyrus.

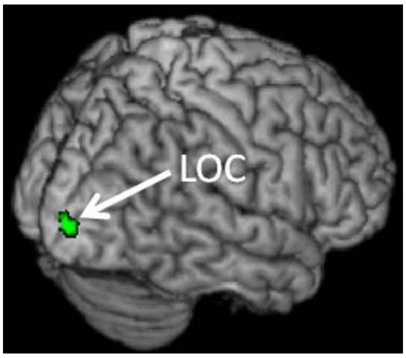


Figure 5.12. This cluster showed larger activation when participants attended to shape compared to attention to motion or colour. LOC: lateral occipital complex.

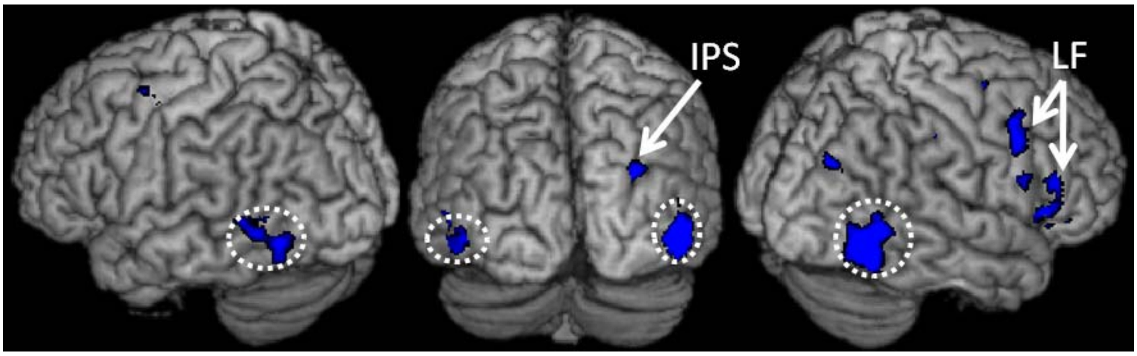


Figure 5.13. This network showed larger activation when participants attended to motion compared to attention to shape or colour. Dashed circles mark the hMT/V5. IPS: intraparietal sulcus; LF: lateral frontal.

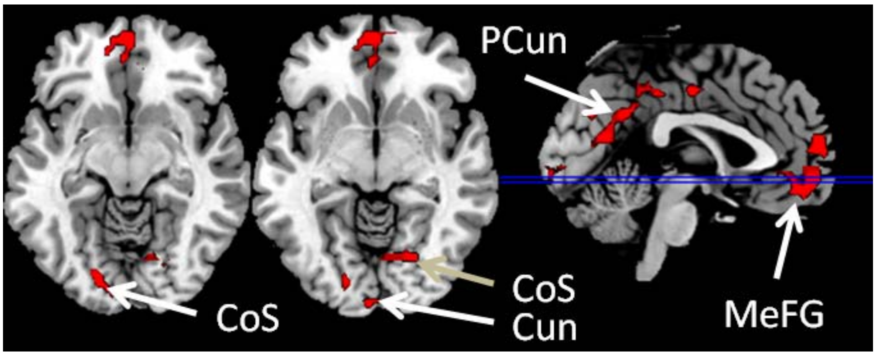


Figure 5.14. Brain areas that showed larger activation when participants attended to colour compared to attention to shape or motion. PCun: precuneus; CoS: collateral sulcus; Cun: cuneus; MeFG: medial frontal gyrus. Slices are MNI y-coordinates -9 and -6.

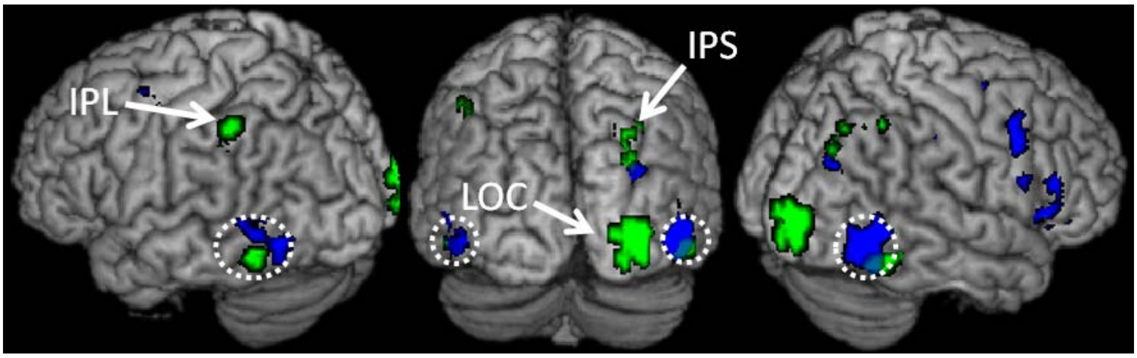


Figure 5.15. Overlay of the clusters activated when participants attended to shape (green) and when participants attended to motion (blue). Note that the initial threshold for the shape clusters was lowered to α = .01. Dashed circles mark the occipito-temporal junction where hMT/V5 and LOC overlapped. IPS: intraparietal sulcus; LOC: lateral occipital complex; IPL: inferior parietal lobule.

Table 5.4. Clusters that showed larger activation to shape as opposed to colour and motion. The initial threshold was lowered to α = .01.

	MNI				
brain area	x	y	z	k	Z
attention shape > attention colour/ motion					
MOG	33	-96	3	71	4.18*
PreCun	30	-51	48	48	3.42#
IOG	-33	-96	-9	21	3.36#
ITG	51	-63	-15	24	3.24#
IOG	-45	-57	-9	28	3.21#
IPL	-42	-42	39	27	2.87#

Note. * p < .05 FDR corrected for multiple comparisons at the cluster level, # p < .05 uncorrected. MOG: middle occipital gyrus, PreCun: precuneus, IOG: inferior occipital gyrus, ITG: inferior temporal gyrus, IPL: inferior parietal lobule, k: number of voxels. MNI coordinates in mm.

As a second step, we searched for brain areas with larger BOLD response during change blocks compared to no change blocks. It was assumed that on no change blocks BOLD responses in brain areas processing unattended features would adapt and therefore decrease their BOLD responses to the unattended features. No area reached significance at the α = .05 level. Finally, we tested for interactions between attended and unattended features. No significant interactions were identified.

5.5.4. Psychophysiological interactions (PPI)

As no differences between change and no change blocks were found in the GLM modulatory influences of regions sensitive to shape, colour, and motion on other brain areas were explored using PPI analyses.

The seed regions for the PPI were selected on basis of the following criteria. Firstly, the area had to be activated in the GLM at the group level. Secondly, the area had to correspond to anatomical locations that have previously been reported to be sensitive to shape, colour, or motion. Based on these criteria six brain areas were selected: bilateral hMT/V5, right LOC, bilateral CoS, and the left cuneus (clusters in Table 5.2 that are marked with ~). To control for unequal cluster sizes seed regions were built as spheres with a 6 mm radius around the voxel that showed the largest difference in BOLD signal change (Bingel et al., 2007). Each sphere consisted of 33 voxels apart from the ones that were fitted for the left hMT/V5 and the LOC (Table 5.2). These clusters only consisted of 25 and 26 voxels, respectively. This was due to the lateral location of the peak voxel which led to an overlap of the sphere and extra-cortical space. The BOLD time course of each seed region was submitted to a separate PPI as the physiological variable. Of main interested for this analyses was whether the brain areas that were identified by the GLM modified each other's activity. Therefore, we were particularly interested in areas that the PPI identified in occipital and occipito-temporal cortices. We set the initial threshold as for the GLM to $\alpha = .01$ (the minimal cluster size was 10 voxels) and considered clusters as showing significant functional connectivity with the seed region when their p value was less than .05 corrected at the cluster level or when their Z values were 3.50 or higher.

Table 5.5. Results of the PPI. Shown are clusters that were consistently modulated by different seed regions.

brain area	MNI				
	x	y	z	k	Z
seed region: right LOC					
SLO	-24	-69	27	24	3.53#
ITG	-54	-69	-3	59	3.50+
seed region: right hMT/V5					
LO	42	-84	0	22	3.63#
SLO	-30	-75	24	67	4.26+
LO	-45	-66	-9	92	4.16+
ITG	45	-66	-3	45	3.72#
seed region: left hMT/V5					
LO	45	-69	-9	21	4.13#
SLO	39	-87	0	39	3.76#
LO	-48	-75	6	21	3.88#
seed region: right CoS					
LO	48	-66	-15	90	3.64+
LO	27	-72	27	84	3.57+
LO	-39	-63	-9	19	3.52+
seed region: left CoS					
LO	-33	-81	-3	11	3.60#

Note. + $p < .05$, FDR correction for multiple comparisons at the cluster level. # $p < .05$, uncorrected. MNI coordinates are in mm. k: number of voxels in the cluster. LO: lateral occipital (includes inferior and middle occipital gyrus), ITG: inferior temporal gyrus, SLO: superior lateral occipital. Initial threshold: $\alpha = .01$.

Table 5.6. Additional clusters that showed PPIs with the seed regions.

brain area	MNI x	y	z	k	Z
seed region: right LOC					
ITG	-54	-69	-3	59	3.50+
CingC	9	9	30	21	3.56#
seed region: right hMT/V5					
ITG	45	-66	-3	45	3.72#
MeFG	-3	27	39	66	3.73+
MFG	-18	24	45	14	3.71#
FFG	-36	-78	-18	36	4.06#
SMG	54	-45	36	17	3.66#
Seed region: left hMT/V5					
MFG	-21	-6	48	19	3.95#
MFG	-36	9	33	15	3.79#
Cun	-3	-93	-9	10	3.84#
CingG	-6	21	42	124	3.76*
caudate	-15	18	3	37	3.52#
seed region: right CoS					
MeFG	12	18	48	120	4.05*
Cun	15	-102	0	20	4.02#
PG	21	-21	-12	36	3.96#
LG	-15	-81	0	32	3.79#
IPL	51	-30	30	34	3.79#
SMG	57	-48	33	58	3.64+
cerebellum	-30	-54	-45	14	3.95#
cerebellum	42	-75	-33	196	3.76*
seed region: left CoS					
MFG	-36	51	15	52	3.67#
IFG	-51	15	0	70	3.29+
caudate	12	15	0	19	4.33#
thalamus	-18	-33	-3	26	3.70#
insula	36	0	-3	12	3.62#
seed region: left Cun					
STG	-57	-21	9	40	4.23#
CingG	-27	30	21	27	3.95#
MFG	-36	39	15	40	3.75#
MFG	-24	15	45	12	3.61#
CingG	-6	9	45	81	3.64+
cerebellum	6	-45	-12	15	3.77#
cerebellum	18	-63	-33	61	3.65+

Note. * p < .0001, FDR correction for multiple comparisons at the cluster level. + p < .05, FDR correction for multiple comparisons at the cluster level. # p < .05, uncorrected. MNI coordinates are in mm. k: number of voxels in the cluster. ITG: inferior temporal gyrus, CingG: cingulate gyrus, MeFG: medial frontal gyrus, MFG: middle frontal gyrus, FFG: fusiform gyrus, SMG: supramarginal gyrus, Cun: cuneus, CoS: collateral sulcus, hMT/V5: middle temporal motion complex, PG: postcentral gyrus, IPL: inferior parietal lobule, IFG: inferior frontal gyrus, STG: superior temporal gyrus.

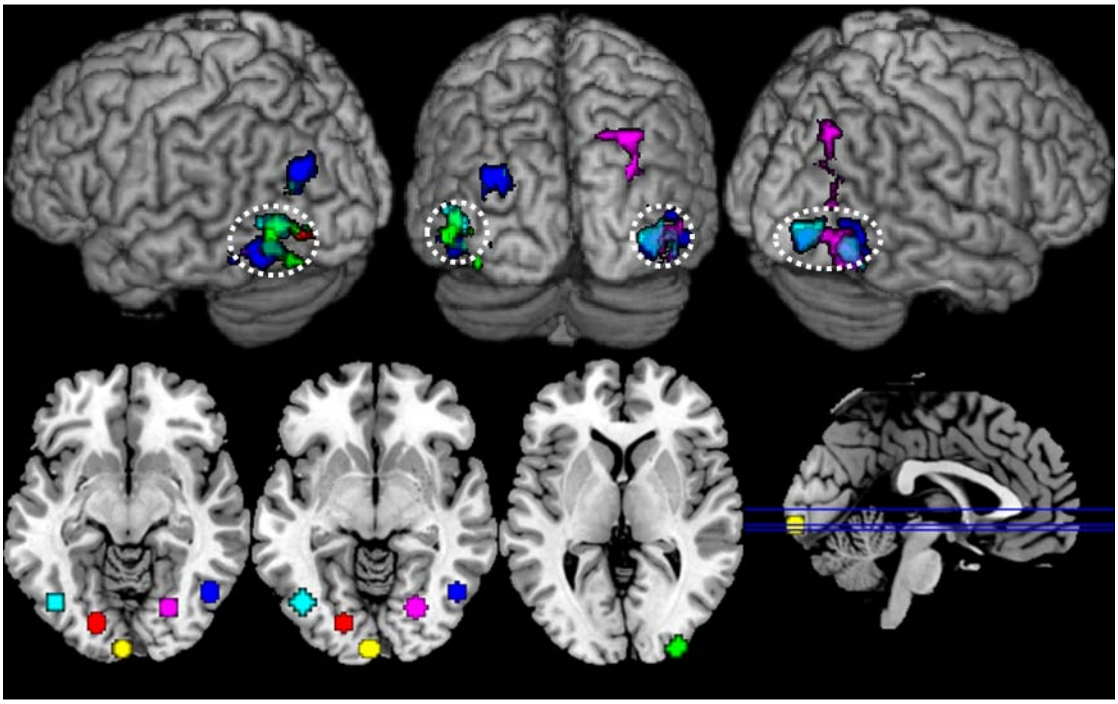


Figure 5.16. Top row: results of the PPI (Table 5.5). Colours indicate which brain area is modified by a certain seed region. Bottom row: seed regions. Colours of seed regions: blue: right hMT/V5; cyan: left hMT/V5; violet: right CoS; red: left CoS; yellow: left Cun; green: right LOC. Dashed circles mark lateral occipital areas that were modulated by the seed regions. Axial slices are MNI y-coordinates -9, -6, 3.

When using the right LOC as a seed region we found PPIs with contralateral occipital and temporal areas. These areas are likely to correspond to parts of the LOC in the left hemisphere (Table 5.5 and Table 5.6; Figure 5.16, green blobs). This means that the functional connectivity between the right LOC and the left LOC was stronger when change blocks were presented compared to when no change blocks are presented. Using the left or the right hMT/V5 as a seed region revealed PPIs with a large set of brain areas across the cortex (Table 5.5 and Table 5.6; Figure 5.16 blue and cyan blobs). Importantly, both the left and the right hMT/V5 showed PPIs with the MOG bilaterally (Table 5.5) which is a part of the LOC. Additionally, the left hMT/V5 showed a PPI with the left LG. This cluster is in proximity to the colour sensitive CoS. This might indicate communications between motion and colour sensitive areas. The right CoS showed PPIs with other colour sensitive areas such as the cuneus and the LG (Table 5.5). Both the left and the right CoS showed PPIs with frontal areas and the lateral occipital gyri (Table 5.5). The left cuneus showed PPIs with frontal and parietal regions. In contrast to the other seed regions the left cuneus did not modulate BOLD activity in lateral occipital or occipito-temporal areas. Overall, the most striking result of this PPI analysis is that five out of six seed regions modulated the activity in brain areas that are part of the anterior LOC. Figure 5.17 depicts a schematic summary of which of the seed regions modulated each other's activity patterns.

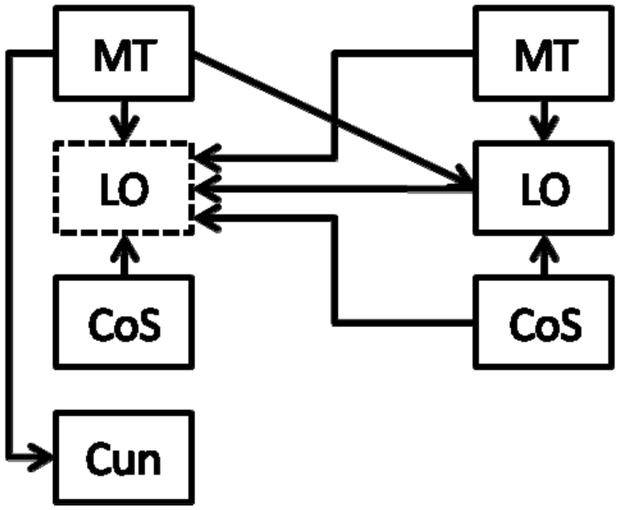


Figure 5.17. Seed regions used in the PPI analyses. The left column indicates seed regions in the left hemisphere; the right column indicates seed regions in the right hemisphere. Arrows show which seed regions were modulated by a seed region. The dashed box indicated an ROI that was not used as a seed region itself as it did not reach the significance criterion in the GLM. Note that the arrows merely reflect the results of the PPI analyses. No assumptions of underlying anatomical connections are made.

5.6. Discussion

This study investigated how unattended object features modulate the BOLD activity in cortical networks that are involved in object recognition. Participants attended to colour, shape, or motion of novel objects while the unattended object features changed or remained the same. Attending to certain features increased the BOLD activity in networks known to be involved in processing the attended feature. Further, we found that unattended object features did not affect the overall BOLD response on whole-brain comparisons. PPI analyses, however, revealed that the activity in the LOC was modulated by brain areas sensitive to the unattended features and these modulations depended on whether unattended features changed or not.

Distinct patterns of BOLD activity were found depending whether participants attended to object shape, motion, or colour. Attending to shape activated a cluster in the posterior part of the lateral occipital cortex in the right hemisphere. This cluster was inferred to be part of the LOC (see Grill-Spector et al., 2001, for a review). Attending to motion activated occipito-temporal clusters in both hemispheres. These are most likely clusters that belong to the motion sensitive hMT/V5 complex. Additionally, attending to motion activated parietal and frontal areas which is in accordance with previous work (Peuskens et al., 2004; Schultz et al., 2008).

Consistently, both Experiment 4 and 5 found that a common network consisting of occipito-temporal cortices, the areas along the IPS, and lateral frontal areas was activated when participants processed shape and motion of objects. This further supports the idea of a combined representation of shape and motion.

A different network of brain areas was activated when participants attended to colour. The bilateral activation along the CoS is consistent with research into colour processing in the brain (e.g., McKeefry and Zeki, 1997). However, an extensive network of other brain areas showed increased activity as well. Interestingly, the activity found in Experiment 4 in medial cortices was replicated. Participants responded more quickly when performing the task on the basis of colour compared to shape and motion. The finding that a different network of brain areas was activated when participants attended to colour might be due to participants not attending to the colour of the object but simply to the colour of a small patch of the surface of the centre of the object. This could have led to participants not comparing *object* colours but simply colours. This would mean that participants performed two different tasks: they attended to features for shape and motion and they attended to a spatial location for attention to colour (see Chapter 2.7). Eye movements were not measured in this experiment and therefore this possibility cannot fully be excluded. Moreover, apart from the expected colour sensitive brain areas at the ventral surface of the occipital cortex (Hadjikhani et al., 1998; McKeefry and Zeki, 1997; Peuskens et al., 2004) attending to colour activated extensive networks in medial brain areas, the CingC in particular, in both Experiment 4 and 5. The CingC is involved in performance monitoring (MacDonald et al., 2000). The finding that the CingC was activated when attending to colour might reflect differences in strategies when participants attended to colour compared to when they attended to shape and motion. However, this explanation is unlikely. Peuskens et al. (2004) controlled for different attention types in their study by introducing an additional task in which participants had to detect changes in luminance in the periphery or the central parts of the objects. They found that the differences between the two dimming conditions were small. The pattern of activation that was found in the present study is largely consistent with the findings by Peuskens et al. (2004). For attention to colour, the present study identified similar areas in ventral parts of the occipital cortex as

Peuskens et al.'s (2004) study. Based on their control study it is assumed that possible differences due to spatial and feature attention are not crucial for the interpretation of the present results.

In addition to the occipito-temporal areas, the lateral frontal areas, and posterior parietal areas along the IPS attending to motion elicited strong activation in the cerebellum bilaterally in both Experiment 4 and 5. This activation pattern was unexpected. The declive of the cerebellum in which the peak activity was found in Experiment 4 and 5 was previously found to respond to arrangements of objects (Gates and Yoon, 2005) which shows that the cerebellum responds to certain types of objects. The activated clusters extended to posterior parts of the cerebellum. O'Reilly et al. (2008) found posterior parts of the cerebellum were involved in predicting spatio-temporal properties of objects over time. The cerebellum activation found in Experiment 4 and 5 might reflect strategies of the participants in anticipating the motion of the object in order to identify which profile was applied to the present object.

Overall, unattended object features did not increase BOLD responses in areas sensitive to the unattended features. However, PPI analyses revealed systematic modulations of BOLD responses caused by unattended object features. Brain areas sensitive to shape, colour, and motion consistently modulated the activity in the anterior parts of the LOC. Even though attention to colour activated a different network compared to shape or motion, the colour sensitive areas along the CoS modulated lateral occipital regions. This means that these areas are functionally connected to the LOC.

That shape, colour, and motion sensitive areas modulated LOC activity suggests that the LOC is a region that collects information about different features from other areas that are sensitive to these features and integrates them. This is consistent with the findings of previous research that showed that LOC integrated different visual features. Self and Zeki (2005) found that when participants had to integrate colour and motion information for shape extraction the LOC was seemed to be involved in this process. The idea of the LOC as an area integrating object features fits nicely with the results of Experiment 4 that showed that when participants attended to whole objects as opposed to features of objects the LOC showed increased activation. The results of Experiment 4 and 5 indicate that the LOC is not only a brain area that responds to shapes but an area that combines information about different object features.

O’Craven et al. (1999) and Paradis et al. (2008) found increased activity in brain areas processing unattended object features when other features of the object were attended to. This indicated that the BOLD responses to the objects suggested object-based representation. Our results are in line with these findings. In contrast to previous research (O'Craven et al., 1999; Paradis et al., 2008) Experiment 5 did not find an overall increase in activation elicited by unattended features. Shape, motion, and colour sensitive brain areas modulated the activity of lateral occipital areas that are part of the LOC an area known to respond to objects. Although Experiments 1 to 3 suggested unattended colour affected recognition performance only to a small extend the activity in colour sensitive brain areas modulated the activity in the LOC as a function of whether unattended features changed or not. This result not only shows that unattended object features affect the BOLD response it also indicates how the brain integrates information about multiple object features.

6. General discussion

Motivated by the diverging assumptions of the two main theories of object recognition on how non-shape features are represented, this thesis investigated the response to colourful dynamic objects at the behavioural and the neural level. Object recognition tasks were combined with feature attention tasks to investigate the representations of shape, colour, and motion. Three behavioural, two fMRI, and one DTI study were conducted providing a comprehensive approach to extend the knowledge of how multi-featured objects are represented. The guiding hypothesis was that there are large scale cortical networks that are involved in object recognition with different network parts processing certain features of objects and top-down attentional mechanisms modulating the activation pattern in the network. In a series of psychophysical and neuroimaging experiments participants attended to one object feature while performing recognition tasks. At the same time, unattended features were manipulated to test how they affect recognition performance.

At the behavioural level, the results indicated that object shape and object motion were represented in an object-based manner which means that shape and motion were bound in a pre-attentive stage in the recognition process making it difficult for participants to process the two features independently. This result is consistent with previous research showing that task-irrelevant rotation direction of novel objects (e.g., Stone, 1998; 1999), facial and head motion (see O'Toole et al., 2002, for a review), and other types of motion (e.g., Pyles et al., 2007) affect shape recognition. Vuong and Tarr's (2004) work showed that the spatio-temporal signature of an object is important for object recognition. Unattended colour, however, had less influence on recognition performance. Based on diverging results of whether colour was a relevant feature for object recognition, Price and Humphreys (1989) suggested that colour and shape of objects are encoded in separate but highly interconnected representations. Our results are very much in line with this idea as colour information affected recognition performance but not to the same extent as shape and motion.

At the neural level, attending to shape and attending to motion activated a common network consisting of occipito-temporal areas, areas along the IPS in the posterior parietal cortex, and lateral frontal areas. It was expected that attending to shape would activate LOC (see Grill-Spector et al., 2001, for a review) and attending to motion would activate hMT/V5 (see Culham et al., 2001, for a review). The common clusters activated at the occipito-temporal junction in Experiment 4 and 5 possibly include both LOC and hMT/V5. Their co-activation might reflect the inability of the observers to process shape and motion independently. Future research in face perception, for example, can shed light on whether the co-activation is a specific effect caused by the objects used in the present experiments or whether it is an effect of non-rigid motion per se. If non-rigidly moving faces (such as speaking faces or faces expressing emotions) co-activate hMT/V5 in addition to face perception areas such as FFA it can be inferred that non-rigid motion causes co-activation of areas processing the shape and areas processing dynamic features of the object.

Both attention to shape and attention to motion activated posterior parietal cortices along the IPS, similar to the results of Peuskens et al. (2004). This is in line with previous research showing that the posterior parietal cortex was involved in processing shape (Paradis et al., 2001; Schultz et al., 2008) and high-level motion (Wu et al., 2008). Alternatively, the activation found in the posterior parietal cortex when participants attended to shape or motion might reflect higher load in visual working memory compared to when participants attended to colour (Kawasaki et al., 2008; Xu, 2007). If attending to shape or motion automatically led to co-selection of unattended motion or unattended shape, respectively, and if no such co-selection processes happen when participants attend to colour, the load in working visual memory is higher on attention to shape and attention to motion trials than on attention to colour trials. Consistent with this possibility, Xu et al. (2007) and Kawasaki et al. (2008) showed that the activity in parts of the posterior parietal cortex reflected the number of multiple object features in visual working memory.

The frontal activation found in Experiments 4 and 5 when participants attended to shape and motion of objects is in accordance with the results of Schultz et al. (2008) who found that lateral frontal clusters were activated when participants performed same-different tasks based on the shape of novel objects. Although other studies have not reported activation of frontal cortices in object recognition tasks (Paradis et al., 2008; Peuskens et al., 2004) it is likely that lateral frontal areas are an important part of the object recognition network. Schultz et al.'s (2008) results suggest a cortical network consisting of several areas each specialised for a sub-process that contributes to the complex process of object recognition. Lateral frontal areas are involved in attentional processes, cognitive control, and memory processes (Cabeza et al., 2003; Kanwisher and Wojciulik, 2000; Miller, 2000; Spencer et al., 2009; Yee et al., 2010) all of which are processes crucial to correctly identifying objects. Perhaps the non-rigid motion used in the present study and the (task-irrelevant) rotation reversals in Schultz et al.'s (2008) study placed higher load on the working memory resources compared to the stimuli and task used by Peuskens et al. (2004) and therefore activated frontal areas.

The network activated by both shape and motion included areas along both the ventral and the dorsal pathway. Traditionally, shape information was assumed to be processed along the ventral pathway and motion information was processed along the dorsal pathway. The present findings, however, add to the evidence (Konen and Kastner, 2008) that processing of object identity does not exclusively happen along the ventral pathway.

The parts of the network activated by shape and motion were not only co-activated when observers processed motion and shape they also showed connectivity at the functional and structural level. As found by Schultz et al. (2008), Experiment 4 showed that the areas in the network are functionally connected to each other as their BOLD signals were correlated, independent of the experimental paradigm (Haynes et al., 2005; Macaluso et al., 2000). Correlations in the activity between brain areas indicate that these areas communicate with each other at the functional level. While different parts of the network might be specialised to certain aspects of the object recognition process (Schultz et al., 2008), the functional connectivity between the parts of the network found in Experiments 4 and 5 might reflect crucial processes of co-ordinating specialised parts of the network. Moreover, Experiment 4 showed that the anatomical connections between occipito-temporal and lateral frontal parts of the network are relevant for object recognition performance. To our knowledge, this is the first study showing the relevance of white matter connections for the recognition of novel object.

In line with the behavioural results, attending to colour activated a network of brain areas that was distinct from the one activated by shape and motion. Dissociation of the representation of shape and colour in object recognition is in acccordance with previous studies showing that separate brain areas processed colour and shape of objects (Cant and Goodale, 2007; Peuskens et al., 2004). Recent research (Cavina-Pratesi et al., 2010) into patients with selective lesions in colour and texture or shape sensitive brain areas provided causal evidence that distinct areas in the brain enable processing of object colour and object shape.

Taken together, both the behavioural and the neuroimaging results of this thesis strongly support the idea by Price and Humphreys (1989) that colour and shape are represented separately but that both representations are highly interconnected. A neural correlate for this interconnection was found in Experiment 5. The colour sensitive areas at the ventral surface of the occipital cortex modulated the activation in shape-sensitive areas. This is an important finding for understanding how cortical areas involved in object recognition are organised. Features can be processed in specialised areas but the information is sent to other areas which combine information about different features to create a holistic mental image of a whole object.

Further evidence for how information about multiple object features is combined can be derived from Experiment 4. When participants attended to whole objects as opposed to object features brain areas that are part of the LOC showed increased activation. Combining these results with the finding that activation in anterior parts of the LOC were modulated by shape, motion, and colour sensitive brain areas (Experiment 5) extend the results by Self and Zeki (2005) which showed that the LOC integrated colour and motion in order to extract shape information. The current findings suggest that the LOC not only integrates features to define a shape, but that it actually integrates information about multiple features in order to create a holistic representation of an object.

All experiments conducted for this thesis indicate that attention plays an important role for object recognition. Participants were able to follow the task instructions and reliably attended to separate features of the objects. However, their behavioural performance was affected by unattended object features showing that attentional mechanisms are not sufficient to filter out task-irrelevant information from unattended features. Treisman and Gelade (1980) suggested that attention is necessary to bind object features together. The results of the present study, however, indicate object-based representations of object features suggesting that features are bound automatically in a pre-attentive state in the recognition process. Attention allows for enhanced processing of a feature of interest. However, it is not sufficient to suppress processing of other features. At the neural level, attention had a large influence on the BOLD response pattern. Depending on the attended feature brain areas involved in processing this feature showed increased activation indicating enhanced processing of the feature.

Understanding attentional mechanisms that are involved in object recognition is important for understanding object recognition in everyday life. Although attentional mechanisms allow for selective screening of the environment for certain features the experiments reported here show that other available features are not only perceived but actually represented and that they can influence recognition. Even though the brain activation is distinct for what we are screening the environment for (e.g., if we are looking for a friend wearing a blue shirt) other object features (e.g., the specific type of shirt such as long-sleeved) are probably still perceived and encoded.

The two main theories of object recognition (the structural description and the image-based theories) diverge in the assumptions on how non-shape object features are represented. Structural description models postulate that object recognition is an entirely shape-driven process (Biederman, 1987; Marr and Nishihara, 1978) suggesting feature-based processing of object features. Image-based theories of object recognition, in contrast, postulate that objects are represented as snapshots of the objects (Edelman and Bulthoff, 1992; Tarr, 1995; Tarr and Bulthoff, 1995; 1998), suggesting object-based processing of objects.

The psychophysical and neuroimaging results of the studies presented in this thesis found object-based representation of shape and motion which is consistent with the image-based theories (Edelman and Bulthoff, 1992; Tarr, 1995; Tarr and Bulthoff, 1995; 1998). Further, the results are consistent with the idea of Price and Humphreys (1989) that colour and shape are represented in separate but highly interconnected representations. Separate representation of colour is consistent with the assumptions of structural description theories (Biederman, 1987; Marr and Nishihara, 1978). This suggests that a synthesis of both object recognition theories to account for the special role of colour for object recognition might be appropriate.

Based on the findings of this thesis future theories of object recognition will only be able to describe object recognition completely if they take both motion and colour and their relations to shape into account. Tanaka et al. (2001) suggest that object recognition can be best described by structure plus surface models. However, it is necessary that future theories account for combined representations of shape and motion and separate but interconnected representations of shape and colour.

The influence of attention on object recognition found in the studies reported here shows that object recognition theories could benefit by taking into account attentional processes (e.g., Hummel and Stankiewicz, 1998; Walther and Koch, 2007). Our results suggest that object features (at least shape and motion) are bound in a pre-attentive manner but attention can still be used to enhance processing of a feature of interest. This means that attention enhances processing of the feature of interest but this is a graded process. Features are not selected in an all-or-none manner but rather all features are available and processing of a feature of interest can be enhanced relative to other object features.

7. General conclusion

The experiments conducted in the framework of this thesis suggest integrated (object-based) representations of shape and motion of objects. While object colour is likely to be represented separately; there are strong interconnections between the representation of colour and the representations of shape or motion (Price and Humphreys, 1989). By themselves, neither the image-based (Edelman and Bulthoff, 1992; Tarr, 1995; Tarr and Bulthoff, 1995; 1998) nor the structural description theories (Biederman, 1987; Marr and Nishihara, 1978) are sufficient to explain the current results. A synthesis of both theories seems appropriate to account for the qualitative differences in the representations of multiple object features. Importantly, the results of this thesis highlight the ability of the object recognition system to use object representations in a flexible way. Even though the representations of multiple object features are integrated or highly interconnected top-down attentional mechanisms allow for selecting relevant object features for the task at hand. Large scale networks were identified that consisted of functionally and anatomically connected brain regions specialised to process specific object features and regions integrating multiple features. Essentially, modulations in functional connectivity were revealed between brain areas that reflected processing of unattended object features. These modulations may be the neural correlate underlying the ability to efficiently select the important aspects of an object for the task at hand without discarding other information about object identity at the same time.

8. References

AGGARWAL, J. K., CAI, Q., LIAO, W. & SABATA, B. Year. Articulated and elastic non-rigid motion: a review. *In:* Motion of Non-Rigid and Articulated Objects, 1994., Proceedings of the 1994 IEEE Workshop on, 1994. 2-14.

BAIZER, J. S., ROBINSON, D. L. & DOW, B. M. 1977. Visual responses of area 18 neurons in awake, behaving monkey. *Journal of Neurophysiology,* 40**,** 1024-1037.

BEGRE, S., FROMMER, A., VON KANEL, R., KIEFER, C. & FEDERSPIEL, A. 2007. Relation of white matter anisotropy to visual memory in 17 healthy subjects. *Brain Res,* 1168**,** 60-6.

BEHRENS, T. E. J., BERG, H. J., JBABDI, S., RUSHWORTH, M. F. S. & WOOLRICH, M. W. 2007. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage,* 34**,** 144-155.

BEHRENS, T. E. J., JOHANSEN-BERG, H., WOOLRICH, M. W., SMITH, S. M., WHEELER-KINGSHOTT, C. A. M., BOULBY, P. A., BARKER, G. J., SILLERY, E. L., SHEEHAN, K., CICCARELLI, O., THOMPSON, A. J., BRADY, J. M. & MATTHEWS, P. M. 2003a. Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nature Neuroscience,* 6**,** 750-757.

BEHRENS, T. E. J., WOOLRICH, M. W., JENKINSON, M., JOHANSEN-BERG, H., NUNES, R. G., CLARE, S., MATTHEWS, P. M., BRADY, J. M. & SMITH, S. M. 2003b. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnetic Resonance in Medicine,* 50**,** 1077-1088.

BENSON, D. F. & GREENBERG, J. P. 1969. Visual form agnosia. A specific defect in visual discrimination. *Arch Neurol,* 20**,** 82-9.

BIEDERMAN, I. 1987. Recognition-by-components: a theory of human image understanding. *Psychol Rev,* 94**,** 115-47.

BIEDERMAN, I. & BAR, M. 1999. One-shot viewpoint invariance in matching novel objects. *Vision Res,* 39**,** 2885-99.

BIEDERMAN, I. & GERHARDSTEIN, P. C. 1993. Recognizing depth-rotated objects: evidence and conditions for three-dimensional viewpoint invariance. *J Exp Psychol Hum Percept Perform,* 19**,** 1162-82.

BIEDERMAN, I. & GERHARDSTEIN, P. C. 1995. Viewpoint-dependent mechanisms in visual object recognition - reply to Tarr and Bulthoff (1995). *Journal of Experimental Psychology-Human Perception and Performance,* 21**,** 1506-1514.

BIEDERMAN, I. & JU, G. 1988. Surface versus edge-based determinants of visual recognition. *Cognitive Psychology,* 20**,** 38-64.

BINGEL, U., ROSE, M., GLASCHER, J. & BUCHEL, C. 2007. fMRI reveals how pain modulates visual object processing in the ventral visual stream. *Neuron,* 55**,** 157-167.

BOHR, S., GULLMAR, D., KNAB, R., REICHENBACH, J. R., WITTE, O. W. & HAUEISEN, J. 2007. Fractional anisotropy correlates with auditory simple reaction time performance. *Brain Research,* 1186**,** 194-202.

BRAINARD, D. H. 1997. The psychophysics toolbox. *Spatial Vision,* 10**,** 433-436.

BROADBENT, D. E. (ed.) 1958. *Perception and communication,* New York: Pergamon Press.

CABEZA, R., DOLCOS, F., PRINCE, S. E., RICE, H. J., WEISSMAN, D. H. & NYBERG, L. 2003. Attention-related activity during episodic memory retrieval: a cross-function fMRI study. *Neuropsychologia,* 41**,** 390-9.

CANT, J. S. & GOODALE, M. A. 2007. Attention to form or surface properties modulates different regions of human occipitotemporal cortex. *Cereb Cortex,* 17**,** 713-31.

CANT, J. S., LARGE, M. E., MCCALL, L. & GOODALE, M. A. 2008. Independent processing of form, colour, and texture in object perception. *Perception,* 37**,** 57-78.

CAVINA-PRATESI, C., KENTRIDGE, R. W., HEYWOOD, C. A. & MILNER, A. D. 2010. Separate Processing of Texture and Form in the Ventral Stream: Evidence from fMRI and Visual Agnosia. *Cerebral Cortex,* 20**,** 433-446.

CORBETTA, M., MIEZIN, F. M., DOBMEYER, S., SHULMAN, G. L. & PETERSEN, S. E. 1990. Attentional modulation of neural processing of shape, color, and velocity in humans. *Science,* 248**,** 1556-9.

CORBETTA, M., MIEZIN, F. M., DOBMEYER, S., SHULMAN, G. L. & PETERSEN, S. E. 1991. Selective and divided attention during visual discriminations of shape, color, and speed: functional anatomy by positron emission tomography. *J Neurosci,* 11**,** 2383-402.

CULHAM, J., HE, S., DUKELOW, S. & VERSTRATEN, F. A. 2001. Visual motion and the human brain: what has neuroimaging told us? *Acta Psychol (Amst),* 107**,** 69-94.

DAVIDOFF, J. B. & OSTERGAARD, A. L. 1988. The role of colour in categorial judgements. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology,* 40**,** 533-544.

DEAN, P. 1976. Effects of inferotemporal lesions on the behavior of monkeys. *Psychol Bull,* 83**,** 41-71.

DOW, B. M. & GOURAS, P. 1973. Color and spatial specificity of single units in rhesus-monkey foveal striate cortex. *Journal of Neurophysiology,* 36**,** 79-100.

DUBNER, R. & ZEKI, S. M. 1971. Response properties and receptive fields of cells in an anatomically defined region of the superior temporal sulcus in the monkey. *Brain Res,* 35**,** 528-32.

DUMOULIN, S. O., BITTAR, R. G., KABANI, N. J., BAKER, C. L., JR., LE GOUALHER, G., BRUCE PIKE, G. & EVANS, A. C. 2000. A new anatomical landmark for reliable identification of human area V5/MT: a quantitative analysis of sulcal patterning. *Cereb Cortex,* 10**,** 454-63.

DUPONT, P., ORBAN, G. A., DE BRUYN, B., VERBRUGGEN, A. & MORTELMANS, L. 1994. Many areas in the human brain respond to visual motion. *J Neurophysiol,* 72**,** 1420-4.

EDELMAN, S. & BULTHOFF, H. H. 1992. Orientation dependence in the recognition of familiar and novel views of three-dimensional objects. *Vision Res,* 32**,** 2385-400.

ERIKSEN, C. W. & SCHULTZ, D. W. 1979. Information-processing in visual-search - continuous-flow conception and experimental results. *Perception & Psychophysics,* 25**,** 249-263.

FELLEMAN, D. J. & VAN ESSEN, D. C. 1987. Receptive-field properties of neurons in area V3 of macaque monkey extrastriate cortex. *Journal of Neurophysiology,* 57**,** 889-920.

FELLEMAN, D. J. & VAN ESSEN, D. C. 1991. Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex,* 1**,** 1-47.

FFYTCHE, D. H. & CATANI, M. 2005. Beyond localization: from hodology to function. *Philos Trans R Soc Lond B Biol Sci,* 360**,** 767-79.

FRISTON, K. J., BUECHEL, C., FINK, G. R., MORRIS, J., ROLLS, E. & DOLAN, R. J. 1997. Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage,* 6**,** 218-229.

FRISTON, K. J., HOLMES, A. P., WORSLEY, K. J., POLINE, J. P., FRITH, C. D. & FRACKOWIAK, R. S. J. 1994. Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping,* 2**,** 189-210.

FRISTON, K. J., STEPHAN, K. E., LUND, T. E., MORCOM, A. & KIEBEL, S. 2005. Mixed-effects and fMRI studies. *Neuroimage,* 24**,** 244-252.

GARNER, W. R. 1988. Facilitation and interference with a separable redundant dimension in stimulus comparison. *Percept Psychophys,* 44**,** 321-30.

GATES, L. & YOON, M. G. 2005. Distinct and shared cortical regions of the human brain activated by pictorial depictions versus verbal descriptions: an fMRI study. *Neuroimage,* 24**,** 473-486.

GEGENFURTNER, K. R. & KIPER, D. C. 2003. Color vision. *Annual Review of Neuroscience,* 26**,** 181-206.

GIESE, M. A. & POGGIO, T. 2003. Neural mechanisms for the recognition of biological movements. *Nat Rev Neurosci,* 4**,** 179-92.

GITELMAN, D. R., PENNY, W. D., ASHBURNER, J. & FRISTON, K. J. 2003. Modeling regional and psychophysiologic interactions in fMRI: the importance of hemodynamic deconvolution. *Neuroimage,* 19**,** 200-207.

GOTTWALD, R. L. & GARNER, W. R. 1975. Filtering and condensation tasks with integral and separable dimensions. *Perception & Psychophysics,* 18**,** 26-28.

GREFKES, C. & FINK, G. R. 2005. The functional organization of the intraparietal sulcus in humans and monkeys. *Journal of Anatomy,* 207**,** 3-17.

GREFKES, C., WEISS, P. H., ZILLES, K. & FINK, G. R. 2002. Crossmodal processing of object features in human anterior intraparietal cortex: an fMRI study implies equivalencies between humans and monkeys. *Neuron,* 35**,** 173-184.

GRILL-SPECTOR, K., KOURTZI, Z. & KANWISHER, N. 2001. The lateral occipital complex and its role in object recognition. *Vision Res,* 41**,** 1409-22.

GRILL-SPECTOR, K., KUSHNIR, T., EDELMAN, S., AVIDAN, G., ITZCHAK, Y. & MALACH, R. 1999. Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron,* 24**,** 187-203.

GRILL-SPECTOR, K. & MALACH, R. 2001. fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychol (Amst),* 107**,** 293-321.

GROSSMAN, E., DONNELLY, M., PRICE, R., PICKENS, D., MORGAN, V., NEIGHBOR, G. & BLAKE, R. 2000. Brain areas involved in perception of biological motion. *J Cogn Neurosci,* 12**,** 711-20.

HADJIKHANI, N., LIU, A. K., DALE, A. M., CAVANAGH, P. & TOOTELL, R. B. H. 1998. Retinotopy and color sensitivity in human visual cortical area V8. *Nature Neuroscience,* 1**,** 235-241.

HARADA, T., GODA, N., OGAWA, T., ITO, M., TOYODA, H., SADATO, N. & KOMATSU, H. 2009. Distribution of colour-selective activity in the monkey inferior temporal cortex revealed by functional magnetic resonance imaging. *European Journal of Neuroscience,* 30**,** 1960-1970.

HAYNES, J. D., DRIVER, J. & REES, G. 2005. Visibility reflects dynamic changes of effective connectivity between V1 and fusiform cortex. *Neuron,* 46**,** 811-21.

HAYWARD, W. G. & WILLIAMS, P. 2000. Viewpoint dependence and object discriminability. *Psychological Science,* 11**,** 7-12.

HAYWORTH, K. J. & BIEDERMAN, I. 2006. Neural evidence for intermediate representations in object recognition. *Vision Research,* 46**,** 4024-4031.

HEYWOOD, C. A. & COWEY, A. 1987. On the role of cortical area V4 in the discrimination of hue and pattern in macaque monkeys. *J Neurosci,* 7**,** 2601-17.

HUBEL, D. H. & WIESEL, T. N. 1977. Ferrier lecture. Functional architecture of macaque monkey visual cortex. *Proc R Soc Lond B Biol Sci,* 198**,** 1-59.

HUMMEL, J. E. & STANKIEWICZ, B. J. 1998. Two roles for attention in shape perception: A structural description model of visual scrutiny. *Visual Cognition,* 5**,** 49-79.

JASTORFF, J., KOURTZI, Z. & GIESE, M. A. 2006. Learning to discriminate complex movements: Biological versus artificial trajectories. *Journal of Vision,* 6**,** 791-804.

JASTORFF, J., KOURTZI, Z. & GIESE, M. A. 2009. Visual Learning Shapes the Processing of Complex Movement Stimuli in the Human Brain. *Journal of Neuroscience,* 29**,** 14026-14038.

JENKINSON, M., BANNISTER, P., BRADY, M. & SMITH, S. 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage,* 17**,** 825-841.

JENKINSON, M. & SMITH, S. 2001. A global optimisation method for robust affine registration of brain images. *Medical Image Analysis,* 5**,** 143-156.

JOHANSSON, G. 1973. Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics,* 14**,** 201-211.

KANWISHER, N., CHUN, M. M., MCDERMOTT, J. & LEDDEN, P. J. 1996. Functional imaging of human visual recognition. *Brain Res Cogn Brain Res,* 5**,** 55-67.

KANWISHER, N. & WOJCIULIK, E. 2000. Visual attention: insights from brain imaging. *Nat Rev Neurosci,* 1**,** 91-100.

KAWASAKI, M., WATANABE, M., OKUDA, J., SAKAGAMI, M. & AIHARA, K. 2008. Human posterior parietal cortex maintains color, shape and motion in visual short-term memory. *Brain Res,* 1213**,** 91-7.

KIM, D. S. & KIM, M. 2005. Combining functional and diffusion tensor MRI. *White Matter in Cognitive Neuroscience: Advances in Diffusion Tensor Imaging and Its Applications,* 1064**,** 1-+.

KNAPPMEYER, B., THORNTON, I. M. & BULTHOFF, H. H. 2003. The use of facial motion and facial form during the processing of identity. *Vision Research,* 43**,** 1921-1936.

KNIGHT, B. & JOHNSTON, A. 1997. The role of movement in face recognition. *Visual Cognition,* 4**,** 265-273.

KONEN, C. S. & KASTNER, S. 2008. Two hierarchically organized neural systems for object information in human visual cortex. *Nature Neuroscience,* 11**,** 224-231.

KOSSLYN, S. M., ALPERT, N. M., THOMPSON, W. L., CHABRIS, C. F., RAUCH, S. L. & ANDERSON, A. K. 1994. Identifying objects seen from different viewpoints. A PET investigation. *Brain,* 117 (Pt 5)**,** 1055-71.

KOURTZI, Z., BULTHOFF, H. H., ERB, M. & GRODD, W. 2002. Object-selective responses in the human motion area MT/MST. *Nat Neurosci,* 5**,** 17-8.

KOURTZI, Z., ERB, M., GRODD, W. & BULTHOFF, H. H. 2003. Representation of the perceived 3-D object shape in the human lateral occipital complex. *Cerebral Cortex,* 13**,** 911-920.

KRAUT, M., HART, J., JR., SOHER, B. J. & GORDON, B. 1997. Object shape processing in the visual system evaluated using functional MRI. *Neurology,* 48**,** 1416-20.

KRIEGESKORTE, N., SORGER, B., NAUMER, M., SCHWARZBACH, J., VAN DEN BOOGERT, E., HUSSY, W. & GOEBEL, R. 2003. Human cortical object recognition from a visual motion flowfield. *J Neurosci,* 23**,** 1451-63.

LANDER, K. & BRUCE, V. 2000. Recognizing famous faces: Exploring the benefits of facial motion. *Ecological Psychology,* 12**,** 259-272.

LANDER, K., CHRISTIE, F. & BRUCE, V. 1999. The role of movement in the recognition of famous faces. *Memory & Cognition,* 27**,** 974-985.

LEEMANS, A. & JONES, D. K. 2009. The B-Matrix Must Be Rotated When Correcting for Subject Motion in DTI Data. *Magnetic Resonance in Medicine,* 61**,** 1336-1349.

LERNER, Y., HENDLER, T., BEN-BASHAT, D., HAREL, M. & MALACH, R. 2001. A hierarchical axis of object processing stages in the human visual cortex. *Cereb Cortex,* 11**,** 287-97.

LING, Y. Z. & HURLBERT, A. 2004. Color and size interactions in a real 3D object similarity task. *Journal of Vision,* 4**,** 721-734.

LIU, T. & COOPER, L. A. 2003. Explicit and implicit memory for rotating objects. *J Exp Psychol Learn Mem Cogn,* 29**,** 554-62.

LOGOTHETIS, N. K., PAULS, J. & POGGIO, T. 1995. Shape representation in the inferior temporal cortex of monkeys. *Current Biology,* 5**,** 552-563.

LOGOTHETIS, N. K. & SHEINBERG, D. L. 1996. Visual object recognition. *Annu Rev Neurosci,* 19**,** 577-621.

LUCAS, B. D. & KANADE, T. 1981. An iterative image registration technique with an application to stereo vision. *Proceedings of the 7th international joint conference on Artificial intelligence - Volume 2.* Vancouver, BC, Canada: Morgan Kaufmann Publishers Inc.

MACALUSO, E., FRITH, C. D. & DRIVER, J. 2000. Modulation of human visual cortex by crossmodal spatial attention. *Science,* 289**,** 1206-8.

MACDONALD, A. W., COHEN, J. D., STENGER, V. A. & CARTER, C. S. 2000. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science,* 288**,** 1835-1838.

MALACH, R., REPPAS, J. B., BENSON, R. R., KWONG, K. K., JIANG, H., KENNEDY, W. A., LEDDEN, P. J., BRADY, T. J., ROSEN, B. R. & TOOTELL, R. B. 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc Natl Acad Sci U S A,* 92**,** 8135-9.

MALDJIAN, J. A., LAURIENTI, P. J. & BURDETTE, J. H. 2004. Precentral gyrus discrepancy in electronic versions of the Talairach atlas. *Neuroimage,* 21**,** 450-455.

MALDJIAN, J. A., LAURIENTI, P. J., KRAFT, R. A. & BURDETTE, J. H. 2003. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage,* 19**,** 1233-1239.

MARR, D. 1982. *Vision,* New York, W. H. Freeman and Company.

MARR, D. & NISHIHARA, H. K. 1978. Representation and recognition of the spatial organization of three-dimensional shapes. *Proc R Soc Lond B Biol Sci,* 200**,** 269-94.

MAUNSELL, J. H. & VAN ESSEN, D. C. 1983. The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *J Neurosci,* 3**,** 2563-86.

MCKEEFRY, D. J. & ZEKI, S. 1997. The position and topography of the human colour centre as revealed by functional magnetic resonance imaging. *Brain,* 120 (Pt 12)**,** 2229-42.

MEADOWS, J. C. 1974. DISTURBED PERCEPTION OF COLORS ASSOCIATED WITH LOCALIZED CEREBRAL LESIONS. *Brain,* 97**,** 615-632.

MILLER, E. K. 2000. The prefrontal cortex and cognitive control. *Nat Rev Neurosci,* 1**,** 59-65.

MILNER, A. D., PERRETT, D. I., JOHNSTON, R. S., BENSON, P. J., JORDAN, T. R., HEELEY, D. W., BETTUCCI, D., MORTARA, F., MUTANI, R., TERAZZI, E. & ET AL. 1991. Perception and action in 'visual form agnosia'. *Brain,* 114 (Pt 1B)**,** 405-28.

MISHKIN, M., UNGERLEIDER, L. G. & MACKO, K. A. 1983. Object vision and spatial vision - 2 cortical pathways. *Trends in Neurosciences,* 6**,** 414-417.

MOUNTS, J. R. & MELARA, R. D. 1999. Attentional selection of objects or features: evidence from a modified search task. *Percept Psychophys,* 61**,** 322-41.

MURATA, A., GALLESE, V., LUPPINO, G., KASEDA, M. & SAKATA, H. 2000. Selectivity for the shape, size, and orientation of objects for grasping in neurons of monkey parietal area AIP. *Journal of Neurophysiology,* 83**,** 2580-2601.

MURRAY, S. O., OLSHAUSEN, B. A. & WOODS, D. L. 2003. Processing shape, motion and three-dimensional shape-from-motion in the human cortex. *Cereb Cortex,* 13**,** 508-16.

MURRAY, S. O. & WOJCIULIK, E. 2004. Attention increases neural selectivity in the human lateral occipital complex. *Nat Neurosci,* 7**,** 70-4.

NAOR-RAZ, G., TARR, M. J. & KERSTEN, D. 2003. Is color an intrinsic property of object representation? *Perception,* 32**,** 667-80.

NEWELL, F. N., WALLRAVEN, C. & HUBER, S. 2004. The role of characteristic motion in object categorization. *Journal of Vision,* 4**,** 118-129.

O'CRAVEN, K. M., DOWNING, P. E. & KANWISHER, N. 1999. fMRI evidence for objects as the units of attentional selection. *Nature,* 401**,** 584-7.

O'REILLY, J. X., MESULAM, M. M. & NOBRE, A. C. 2008. The cerebellum predicts the timing of perceptual events. *Journal of Neuroscience,* 28**,** 2252-2260.

O'TOOLE, A. J., ROARK, D. A. & ABDI, H. 2002. Recognizing moving faces: a psychological and neural synthesis. *Trends Cogn Sci,* 6**,** 261-266.

OLIVA, A., TORRALBA, A., CASTELHANO, M. S. & HENDERSON, J. M. Year. Top-down control of visual attention in object detection. *In:* Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on, 2003. I-253-6 vol.1.

PARADIS, A. L., CORNILLEAU-PERES, V., DROULEZ, J., VAN DE MOORTELE, P. F., LOBEL, E., BERTHOZ, A., LE BIHAN, D. & POLINE, J. B. 2000. Visual perception of motion and 3-D structure from motion: an fMRI study. *Cerebral Cortex,* 10**,** 772-783.

PARADIS, A. L., DROULEZ, J., CORNILLEAU-PERES, V. & POLINE, J. B. 2001. Neural bases of the perception of 3-D structure-from-motion: insights from event-related fMRI and attention modulation. *Perception,* 30**,** 117-117.

PARADIS, A. L., DROULEZ, J., CORNILLEAU-PERES, V. & POLINE, J. B. 2008. Processing 3D form and 3D motion: respective contributions of attention-based and stimulus-driven activity. *Neuroimage,* 43**,** 736-47.

PELLI, D. G. 1997. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision,* 10**,** 437-442.

PEUSKENS, H., CLAEYS, K. G., TODD, J. T., NORMAN, J. F., VAN HECKE, P. & ORBAN, G. A. 2004. Attention to 3-D shape, 3-D motion, and texture in 3-D structure from motion displays. *J Cogn Neurosci,* 16**,** 665-82.

PILZ, K. S., THORNTON, I. M. & BULTHOFF, H. H. 2006. A search advantage for faces learned in motion. *Experimental Brain Research,* 171**,** 436-447.

PINKER, S. 1984. Visual cognition: an introduction. *Cognition,* 18**,** 1-63.

POGGIO, T. & EDELMAN, S. 1990. A network that learns to recognize three-dimensional objects. *Nature,* 343**,** 263-6.

POSNER, M. I. 1980. Orienting of attention. *Quarterly Journal of Experimental Psychology,* 32**,** 3-25.

POSNER, M. I., SNYDER, C. R. R. & DAVIDSON, B. J. 1980. Attention and the detection of signals. *Journal of Experimental Psychology-General,* 109**,** 160-174.

PRICE, C. J. & HUMPHREYS, G. W. 1989. The effects of surface detail on object categorization and naming. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology,* 41**,** 797-828.

PYLES, J. A., GARCIA, J. O., HOFFMAN, D. D. & GROSSMAN, E. D. 2007. Visual perception and neural correlates of novel 'biological motion'. *Vision Res,* 47**,** 2786-97.

PYLES, J. A. & GROSSMAN, E. D. 2009. Neural adaptation for novel objects during dynamic articulation. *Neuropsychologia,* 47**,** 1261-1268.

RIESENHUBER, M. & POGGIO, T. 1999. Hierarchical models of object recognition in cortex. *Nature Neuroscience,* 2**,** 1019-1025.

ROSCH, E., MERVIS, C. B., GRAY, W. D., JOHNSON, D. M. & BOYESBRAEM, P. 1976. Basic objects in natural categories. *Cognitive Psychology,* 8**,** 382-439.

ROSSION, B. & POURTOIS, G. 2004. Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception,* 33**,** 217-236.

SARKHEIL, P., VUONG, Q. C., BULTHOFF, H. H. & NOPPENEY, U. 2008. The integration of higher order form and motion by the human brain. *Neuroimage,* 42**,** 1529-36.

SASSON, E., DONIGER, G. M., PASTERNAK, O. & ASSAF, Y. 2010. Structural correlates of memory performance with diffusion tensor imaging. *Neuroimage,* 50**,** 1231-42.

SCHOLL, B. J. 2001. Objects and attention: the state of the art. *Cognition,* 80**,** 1-46.

SCHULTZ, J., CHUANG, L. & VUONG, Q. C. 2008. A dynamic object-processing network: metric shape discrimination of dynamic objects by activation of occipitotemporal, parietal, and frontal cortices. *Cereb Cortex,* 18**,** 1302-13.

SCHULTZ, J. & LENNERT, T. 2009. BOLD signal in intraparietal sulcus covaries with magnitude of implicitly driven attention shifts. *Neuroimage,* 45**,** 1314-28.

SELF, M. W. & ZEKI, S. 2005. The integration of colour and motion by the human visual brain. *Cereb Cortex,* 15**,** 1270-9.

SERGENT, J., OHTA, S. & MACDONALD, B. 1992. Functional neuroanatomy of face and object processing. A positron emission tomography study. *Brain,* 115 Pt 1**,** 15-36.

SETTI, A. & NEWELL, F. N. 2010. The effect of body and part-based motion on the recognition of unfamiliar objects. *Visual Cognition,* 18**,** 456-480.

SHINOURA, N., SUZUKI, Y., TSUKADA, M., KATSUKI, S., YAMADA, R., TABEI, Y., SAITO, K. & YAGI, K. 2007. Impairment of inferior longitudinal fasciculus plays a role in visual memory disturbance. *Neurocase,* 13**,** 127-30.

SMITH, S. M. 2002. Fast robust automated brain extraction. *Human Brain Mapping,* 17**,** 143-155.

SMITH, S. M., JENKINSON, M., WOOLRICH, M. W., BECKMANN, C. F., BEHRENS, T. E. J., JOHANSEN-BERG, H., BANNISTER, P. R., DE LUCA, M., DROBNJAK, I., FLITNEY, D. E., NIAZY, R. K., SAUNDERS, J., VICKERS, J., ZHANG, Y. Y., DE STEFANO, N., BRADY, J. M. & MATTHEWS, P. M. 2004. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage,* 23**,** S208-S219.

SPENCER, T. J., MONTALDI, D., GONG, Q. Y., ROBERTS, N. & MAYES, A. R. 2009. Object priming and recognition memory: dissociable effects in left frontal cortex at encoding. *Neuropsychologia,* 47**,** 2942-7.

STONE, J. V. 1998. Object recognition using spatiotemporal signatures. *Vision Res,* 38**,** 947-51.

STONE, J. V. 1999. Object recognition: view-specificity and motion-specificity. *Vision Res,* 39**,** 4032-44.

SUNAERT, S., VAN HECKE, P., MARCHAL, G. & ORBAN, G. A. 1999. Motion-responsive regions of the human brain. *Exp Brain Res,* 127**,** 355-70.

TANAKA, J., WEISKOPF, D. & WILLIAMS, P. 2001. The role of color in high-level vision. *Trends Cogn Sci,* 5**,** 211-215.

TANAKA, J. W. & PRESNELL, L. M. 1999. Color diagnosticity in object recognition. *Perception & Psychophysics,* 61**,** 1140-1153.

TARR, M. J. 1995. Rotating objects to recognize them - A case-study on the role of viewpoint dependency in the recognition of 3-dimensional objects. *Psychonomic Bulletin & Review,* 2**,** 55-82.

TARR, M. J. & BULTHOFF, H. H. 1995. Is human object recognition better described by geon structural descriptions or by multiple views? Comment on Biederman and Gerhardstein (1993). *J Exp Psychol Hum Percept Perform,* 21**,** 1494-505.

TARR, M. J. & BULTHOFF, H. H. 1998. Image-based object recognition in man, monkey and machine. *Cognition,* 67**,** 1-20.

TARR, M. J., KERSTEN, D. & BULTHOFF, H. H. 1998. Why the visual recognition system might encode the effects of illumination. *Vision Research,* 38**,** 2259-2275.

TARR, M. J. & PINKER, S. 1989. Mental rotation and orientation-dependence in shape recognition. *Cogn Psychol,* 21**,** 233-82.

THIELE, A., DOBKINS, K. R. & ALBRIGHT, T. D. 1999. The contribution of color to motion processing in macaque middle temporal area. *Journal of Neuroscience,* 19**,** 6571-6587.

THOMAS, C., AVIDAN, G., HUMPHREYS, K., JUNG, K. J., GAO, F. & BEHRMANN, M. 2009. Reduced structural connectivity in ventral visual cortex in congenital prosopagnosia. *Nat Neurosci,* 12**,** 29-31.

THOMAS, C., MOYA, L., AVIDAN, G., HUMPHREYS, K., JUNG, K. J., PETERSON, M. A. & BEHRMANN, M. 2008. Reduction in white matter connectivity, revealed by diffusion tensor imaging, may account for age-related changes in face perception. *J Cogn Neurosci,* 20**,** 268-84.

TODD, J. T., OOMES, A. H. J., KOENDERINK, J. J. & KAPPERS, A. M. L. 2001. On the affine structure of perceptual space. *Psychological Science,* 12**,** 191-196.

TOOTELL, R. B., REPPAS, J. B., KWONG, K. K., MALACH, R., BORN, R. T., BRADY, T. J., ROSEN, B. R. & BELLIVEAU, J. W. 1995. Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *J Neurosci,* 15**,** 3215-30.

TREISMAN, A. M. & GELADE, G. 1980. Feature-integration theory of attention. *Cognitive Psychology,* 12**,** 97-136.

TUCH, D. S., SALAT, D. H., WISCO, J. J., ZALETA, A. K., HEVELONE, N. D. & ROSAS, H. D. 2005. Choice reaction time performance correlates with diffusion anisotropy in white matter pathways supporting visuospatial attention. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 12212-12217.

ULLMAN, S. 1979. The interpretation of structure from motion. *Proc R Soc Lond B Biol Sci,* 203**,** 405-26.

VAN ESSEN, D. C. & GALLANT, J. L. 1994. Neural mechanisms of form and motion processing in the primate visual-system. *Neuron,* 13**,** 1-10.

VUONG, Q. C. & FRIEDMAN, A. submitted. Parameter-based motion morphing for novel dynamic objects. *Behaviour Research Methods*.

VUONG, Q. C., FRIEDMAN, A. & PLANTE, C. 2009. Modulation of viewpoint effects in object recognition by shape and motion cues. *Perception,* 38**,** 1628-1648.

VUONG, Q. C. & TARR, M. J. 2004. Rotation direction affects object recognition. *Vision Research,* 44**,** 1717-1730.

VUONG, Q. C. & TARR, M. J. 2006. Structural similarity and spatiotemporal noise effects on learning dynamic novel objects. *Perception,* 35**,** 497-510.

WADE, A. R., BREWER, A. A., RIEGER, J. W. & WANDELL, B. A. 2002. Functional measurements of human ventral occipital cortex: retinotopy and colour. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences,* 357**,** 963-973.

WAKANA, S., JIANG, H. Y., NAGAE-POETSCHER, L. M., VAN ZIJL, P. C. M. & MORI, S. 2004. Fiber tract-based atlas of human white matter anatomy. *Radiology,* 230**,** 77-87.

WALLIS, G. 1996. Using spatio-temporal correlations to learn invariant object recognition. *Neural Networks,* 9**,** 1513-1519.

WALLIS, G. 1998. Spatio-temporal influences at the neural level of object recognition. *Network-Computation in Neural Systems,* 9**,** 265-278.

WALLIS, G. 2002. The role of object motion in forging long-term representations of objects. *Visual Cognition,* 9**,** 233-247.

WALLIS, G. & BULTHOFF, H. H. 2001. Effects of temporal association on recognition memory. *Proceedings of the National Academy of Sciences of the United States of America,* 98**,** 4800-4804.

WALLRAVEN, C. & BULTHOFF, H. H. Year. Automatic acquisition of exemplar-based representations for recognition from image sequences. *In:* Computer vision and pattern recognition, 2001.

WALTHER, D. B. & KOCH, C. 2007. Attention in hierarchical models of object recognition. *Prog Brain Res,* 165**,** 57-78.

WATSON, J. D., MYERS, R., FRACKOWIAK, R. S., HAJNAL, J. V., WOODS, R. P., MAZZIOTTA, J. C., SHIPP, S. & ZEKI, S. 1993. Area V5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cereb Cortex,* 3**,** 79-94.

WEGENER, D., EHN, F., AURICH, M. K., GALASHAN, F. O. & KREITER, A. K. 2008. Feature-based attention and the suppression of non-relevant object features. *Vision Res,* 48**,** 2696-707.

WOOLRICH, M. W., JBABDI, S., PATENAUDE, B., CHAPPELL, M., MAKNI, S., BEHRENS, T., BECKMANN, C., JENKINSON, M. & SMITH, S. M. 2009. Bayesian analysis of neuroimaging data in FSL. *Neuroimage,* 45**,** S173-S186.

WU, D. H., MORGANTI, A. & CHATTERJEE, A. 2008. Neural substrates of processing path and manner information of a moving event. *Neuropsychologia,* 46**,** 704-13.

WURM, L. H., LEGGE, G. E., ISENBERG, L. M. & LUEBKER, A. 1993. COLOR IMPROVES OBJECT RECOGNITION IN NORMAL AND LOW-VISION. *Journal of Experimental Psychology-Human Perception and Performance,* 19**,** 899-911.

XU, Y. 2007. The role of the superior intraparietal sulcus in supporting visual short-term memory for multifeature objects. *J Neurosci,* 27**,** 11676-86.

YEE, L. T., ROE, K. & COURTNEY, S. M. 2010. Selective involvement of superior frontal cortex during working memory for shapes. *J Neurophysiol,* 103**,** 557-63.

ZEKI, S. 1980. The representation of colors in the cerebral-cortex. *Nature,* 284**,** 412-418.

ZEKI, S. 1990. A century of cerebral achromatopsia. *Brain,* 113 (Pt 6)**,** 1721-77.

ZEKI, S. & BARTELS, A. 1999. The clinical and functional measurement of cortical (in)activity in the visual brain, with special reference to the two subdivisions (V4 and V4 alpha) of the human colour centre. *Philosophical Transactions of the Royal Society B-Biological Sciences,* 354**,** 1371-1382.

ZEKI, S., WATSON, J. D., LUECK, C. J., FRISTON, K. J., KENNARD, C. & FRACKOWIAK, R. S. 1991. A direct demonstration of functional specialization in human visual cortex. *J Neurosci,* 11**,** 641-9.

ZEKI, S. M. 1973. Colour coding in rhesus monkey prestriate cortex. *Brain Res,* 53**,** 422-7.

ZIHL, J., VON CRAMON, D. & MAI, N. 1983. Selective disturbance of movement vision after bilateral brain damage. *Brain,* 106 (Pt 2)**,** 313-40.

ZIHL, J., VON CRAMON, D., MAI, N. & SCHMID, C. 1991. Disturbance of movement vision after bilateral posterior brain damage. Further evidence and follow up observations. *Brain,* 114 (Pt 5)**,** 2235-52.

9. Appendix

9.1. Appendix to Experiment 3

The remaining results of the omnibus ANOVAs computed for RTs and accuracy in the same-different comparisons experiment (Experiment 3) are shown in Table 9.1.

Table 9.1. Remaining results of the omnibus ANOVAs of Experiment 3. MSE: mean squared error. RT: recognition time.

factor	F(1, 22)	p	η_p^2	MSE
accuracy				
trial type	6.88	.016	.24	223.27
trial type x complement	18.37	< .0001	.46	61.97
trial type x complement x group	5.41	.03	.20	61.97
RT				
trial type	6.54	.018	.23	0.15
complement	6.09	.02	.22	0.02
trial type x complement	34.78	< .0001	.61	0.02
trial type x complement x group	14.34	< .0001	.40	0.02

Tukey post hoc tests were computed with formula (1):

$$(1) \qquad Q = \frac{M_1 - M_2}{\sqrt{\frac{MSE}{N}}}$$

With M_1 and M_2 being the means of interest; MSE the mean square error obtained from the ANOVA; and N the number of participants.

9.2. Appendix to Experiment 4

Table 9.2. Additional clusters that were activated in Experiment 4.

	MNI				
region	x	y	z	k	Z
attention colour > attention shape/motion					
extranuclear	-33	-10	25	24	3.65+
attention motion > attention colour/shape					
declive	-15	-76	-29	197	4.85*
pyramis	18	-79	-41	39	4.59+
declive	18	-76	-29	16	3.89+
declive	36	-58	-29	23	3.81+
PCG	-51	-31	40	22	4.32+
MeFG	6	26	43	45	4.28+

Note. * p < .0001, FDR correction for multiple comparisons at the cluster level. + p < .05, FDR correction for multiple comparisons at the cluster level. PCG: post central gyrus; MeFG: medial frontal gyrus. k: number of voxels per cluster.

Table 9.3. Residual correlations between the areas activated in Experiment 4 when attending to shape, motion, and colour. Shaded cells are significant at α = .05, uncorrected for multiple comparisons. Indices of labels are MNI coordinates (x_y_z_label).

		motion							shape							colour								
		45_-82_-11_IOG	-39_-46_55_IPl	-48_-76_-2_MOG	-57_14_31_IFG	27_-61_58_CS	33_23_-8_IFG	45_5_31_IFG	-27_-73_31_IPS	-45_-85_-5_MOG	-45_11_25_IFG	27_-100_-5_Cun	39_-61_43_IPL	45_-67_-14_FFG	48_38_10_IFG	0_50_-11_CingC	-12_-49_58_preCun	-21_-13_23_PG	-27_29_40_MFG	-30_41_25_MFG	-42_-67_25_MTG	-60_-16_-14_MTG	48_5_-2_insula	12_-64_-8_LG
motion	45_-82_-11_IOG	1.00	0.43	0.78	0.40	0.60	0.27	0.48	0.48	0.77	0.35	0.72	0.52	0.85	0.42	0.31	0.29	0.16	0.17	0.35	0.12	0.35	0.31	0.43
	-39_-46_55_IPL		1.00	0.46	0.46	0.69	0.28	0.47	0.57	0.43	0.41	0.33	0.61	0.39	0.43	0.17	0.32	0.05	0.13	0.30	0.02	0.18	0.17	0.21
	-48_-76_-2_MOG			1.00	0.41	0.59	0.30	0.48	0.50	0.92	0.36	0.61	0.52	0.66	0.44	0.27	0.28	0.15	0.15	0.36	0.07	0.31	0.29	0.42
	-57_14_31_IFG				1.00	0.55	0.37	0.63	0.58	0.40	0.88	0.32	0.53	0.36	0.59	0.18	0.19	0.05	0.14	0.31	0.10	0.20	0.19	0.20
	27_-61_58_CS					1.00	0.35	0.65	0.71	0.58	0.51	0.47	0.89	0.54	0.62	0.22	0.36	0.08	0.18	0.32	0.07	0.22	0.21	0.31
	33_23_-8_IFG						1.00	0.47	0.32	0.29	0.33	0.23	0.38	0.24	0.48	0.25	0.18	0.01	0.09	0.26	0.04	0.22	0.22	0.21
	45_5_31_IFG							1.00	0.57	0.47	0.59	0.38	0.66	0.42	0.80	0.25	0.26	0.06	0.16	0.34	0.13	0.22	0.22	0.26
shape	-27_-73_31_IPS								1.00	0.51	0.59	0.42	0.74	0.44	0.57	0.22	0.28	0.08	0.21	0.30	0.15	0.16	0.14	0.29
	-45_-85_-5_MOG									1.00	0.36	0.66	0.54	0.67	0.45	0.23	0.24	0.15	0.13	0.32	0.07	0.26	0.21	0.40
	-45_11_25_IFG										1.00	0.29	0.51	0.31	0.58	0.17	0.16	0.05	0.15	0.26	0.13	0.14	0.13	0.17
	27_-100_-5_Cun											1.00	0.45	0.56	0.38	0.25	0.20	0.13	0.14	0.28	0.13	0.26	0.22	0.38
	39_-61_43_IPL												1.00	0.49	0.70	0.21	0.29	0.08	0.18	0.27	0.12	0.16	0.12	0.29
	45_-67_-14_FFG													1.00	0.38	0.23	0.24	0.14	0.15	0.27	0.09	0.26	0.21	0.36
	48_38_10_IFG														1.00	0.23	0.19	0.06	0.17	0.29	0.19	0.19	0.18	0.25
colour	0_50_-11_CingC															1.00	0.39	0.30	0.50	0.47	0.55	0.52	0.44	0.46
	-12_-49_58_preCun																1.00	0.11	0.25	0.38	0.14	0.33	0.31	0.31
	-21_-13_23_PG																	1.00	0.18	0.11	0.22	0.20	0.11	0.23
	-27_29_40_MFG																		1.00	0.31	0.40	0.26	0.19	0.25
	-30_41_25_MFG																			1.00	0.14	0.45	0.43	0.40
	-42_-67_25_MTG																				1.00	0.21	0.15	0.22
	-60_-16_-14_MTG																					1.00	0.67	0.46
	48_5_-2_insula																						1.00	0.40
	12_-64_-8_LG																							1.00