

# Vision and audition in motion

Citation for published version (APA):

Alink, A. (2011). *Vision and audition in motion*. [Doctoral Thesis, Maastricht University]. Datawyse / Universitaire Pers Maastricht. <https://doi.org/10.26481/dis.20110211aa>

## Document status and date:

Published: 01/01/2011

## DOI:

[10.26481/dis.20110211aa](https://doi.org/10.26481/dis.20110211aa)

## Document Version:

Publisher's PDF, also known as Version of record

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

# **Vision and Audition in Motion**

**Arjen Alink**

**Colophon**

ISBN 978 94 6159 037 4

Copyright © Arjen Alink, Cambridge UK 2011

# Vision and Audition in Motion

## Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Maastricht,  
op gezag van Rector Magnificus, Prof. mr. G.P.M.F. Mols,  
volgens het besluit van het College van Decanen,  
in het openbaar te verdedigen  
op vrijdag 11 februari 2011 om 16:00 uur

door

**Arjen Alink**

Geboren op 24 september 1983 te Nijmegen



**Promotor**

Prof.dr. Rainer Goebel

**Copromotores**

Dr. Lars Folker Muckli

*University of Glasgow*

Dr. Axel Kohler

*Universitätsklinik Bern*

**Beoordelingscommissie**

Prof.dr. Bernadette Jansma (voorzitter)

Dr. Amanda Leonie Kaas

Prof.dr. Peter de Weerd

Prof.dr. Jochen Kaiser

*Goethe Universität, Frankfurt*

Dr.Nikolaus Kriegeskorte

*Medical Research Council, Cambridge UK*

# Contents

Chapter 1	Introduction	7
Chapter 2	Primary auditory cortex and lateral occipital cortex encode the direction of moving sounds	21
Chapter 3	Auditory, visual and audiovisual motion responses in hMT/V5+ and the planum temporale	37
Chapter 4	Auditory motion capturing ambiguous visual apparent motion	53
Chapter 5	Stimulus predictability reduces responses in primary visual cortex	65
	<i>Summary</i>	91
	<i>Samenvatting</i>	93
	<i>Acknowledgments</i>	97
	<i>Curriculum Vitae</i>	101



# **Chapter 1**

## Introduction

Being aware of the direction in which objects move is critical for safe navigation through our environment. This ability is referred to as motion perception, which is the main topic of this dissertation. Because object motion can be perceived most easily and with greatest precision by looking at the object (McKee and Nakayama, 1984) the term motion perception is commonly associated with the visual sensory modality. Motion can, however, also be perceived auditorily, albeit with a considerably lower precision (Perrott and Musicant, 1977).

Previous studies have shown that activation in a brain area in the extrastriate visual cortex, called hMT/V5+, is important for visual motion perception (Tootell et al., 1995; Beckers and Zeki, 1995) while auditory motion perception has been associated with activation in an area posterior from the primary auditory cortex called the planum temporale (Baumgart et al., 1999; Krumbholz et al., 2005). These findings imply some degree of modularity (Fodor, 1983) for visual and auditory motion processing in the human brain. However, few studies have tested directly whether the brain processes visual and auditory motion signals in an independent or interactive manner. In this dissertation we address this issue by presenting two functional magnetic resonance imaging (fMRI) studies, in chapter 2 and 3, indicating that the human brain processes auditory and visual motion in an interactive manner. Furthermore, in chapter 4 we describe a psychophysical study which shows that interactions between visual and auditory motion processing can also be observed at a perceptual level.

The experiment described in chapter 5 also relates to motion perception but is focused on visual motion perception and how neural responses in the visual cortex are affected by the predictability of a stimulus. The outcome of this experiment shows that a stimulus that is more predictable in the context of illusory visual motion induces a lower response in the primary visual cortex. This finding is in line with several models that suggest that the brain actively anticipates upcoming visual events (Mumford, 1992; Rao and Ballard, 1999; Erlhagen, 2003; Bar, 2007; Enns and Lleras 2008; Bar, 2009; Friston and Kiebel 2009; Grossberg, 2009).

The remaining part of this introduction serves to acquaint the reader with the main issues touched upon in the following chapters after being introduced to the method of fMRI and multi-voxel pattern analysis.

## **Functional Magnetic Resonance Imaging (fMRI) and the Blood Oxygenation Level Dependent (BOLD) response**

In the studies described in chapter 2, 3 and 5, brain function is investigated by measuring blood oxygenation level dependent (BOLD) responses with fMRI. This paragraph's purpose is to provide readers with limited knowledge about

## Introduction

fMRI with some basic information about how the BOLD response relates to neural activation and about how such responses are measured during an fMRI experiment.

When interpreting the results of a study employing fMRI, one has to keep in mind that this neuroimaging method only indirectly measures neural activation levels in the brain. The signal changes measured with fMRI depend on increases in local blood flow and blood volume nearby sites of enhanced neural activation. It is generally assumed in the field of human neuroimaging that such a hemodynamic response is related to an increase of action potentials fired by neurons and that such responses peak around six seconds after an increase of neural activation has taken place. Although convincing support has been presented for this assumption (Logothetis et al., 2001) the neurovascular coupling that drives these hemodynamic responses is still under intense investigation (Koehler et al., 2009).

In humans one can non-invasively measure hemodynamic responses using magnetic resonance imaging (MRI)- a Nobel-price-winning method with which one can obtain three-dimensional images of internal organs (Mansfield, 2005). The signal measured by an MRI scanner are resonance waves emitted by hydrogen atom nuclei as they realign their poles to the strong magnetic field that exists inside the bore of the scanner after this alignment has been disturbed by pulses of radio waves. With functional magnetic resonance imaging (fMRI), one measures changes of this MRI signal in neural tissue as a result of hemodynamic responses. This is possible because a hemodynamic response effectively washes out deoxygenated hemoglobin, a substance that reduces local MRI signal, from the venous capillaries of active neural tissue. For this reason hemodynamic responses induce a local MRI signal increase which one refers to as the blood oxygenation level dependent (BOLD) response.

By using specially designed scanner sequences, called echo-planar imaging (EPI) sequences, one can sample MRI signals across the entire human brain in less than two seconds. The resulting EPI images contain a three-dimensional map of the human brain consisting of around a hundred thousand voxels - three dimensional pixels with a size of around  $3 \times 3 \times 3$  millimetres which contain a value specifying the MRI signal intensity in this volume. A basic fMRI experiment involves presenting subjects with stimuli inside the scanner while continuously recording EPI images. This provides the experimenter with about a hundred thousand MRI intensity time courses (one for each voxel) and a log file that describes when stimuli were presented. Analyzing these data typically consists of testing for each voxel whether its time course is significantly affected by stimulus presentations. This is verified by testing whether the MRI signal in a voxel co-varies with a design matrix - a vector or matrix that describes the signal

increases expected for a responsive voxel based on the time of stimulation and a function specifying the typical time-amplitude relationship of the hemodynamic response (Boynton et al., 1996). The outcome of this analysis is a three-dimensional statistical map describing the responsiveness of each voxel. This map is typically projected onto a high resolution image of the brain (1x1x1 mm) which allows one to precisely determine which anatomical structures of the brain are active.

The great advantage that fMRI offers over all other neuroimaging methods is that it allows one to assess and precisely localize neural activation levels throughout the entire brain. A disadvantage of fMRI is that it provides little information about the behaviour of single neurons because the BOLD response measured inside a voxel results from the interaction of thousands of neurons with the local vasculature. However, a recent development in fMRI analysis, called multi-voxel pattern analysis, holds the promise that with fMRI one can obtain information about response properties of neural populations covering only a fraction of the volume of a single voxel.

### **Multi-voxel pattern analysis**

During an fMRI session one acquires one MRI signal time-course for each voxel. Conventional univariate fMRI analysis consists of sequentially testing whether voxels' MRI time courses contain stimulus-driven BOLD responses or whether BOLD responses differ in amplitude across stimulus conditions. Such a voxel-by-voxel analysis always involves setting a statistical threshold defining the minimal signal-to-noise ratio of the BOLD response, or BOLD response difference, that a voxel must exhibit to be classified as "responsive". By applying this threshold to all voxels, one ends up with a map of the brain in which only voxels are highlighted that exhibit a BOLD response exceeding the threshold. Only these voxels are considered as being involved in the cognitive function under investigation while all other voxels are regarded as "unresponsive" or as voxels whose time course only contains noise.

The adequacy of univariate analysis of fMRI data has been seriously questioned by the findings of a study by Haxby and colleagues (2001). This study investigated the responses of ventral visual cortex to pictures belonging to several object categories. A revolutionary finding presented in this study was that information about stimulus category was present in both the most active voxels as well as in mildly active voxels that conventional univariate analysis would classify as "unresponsive". Hence, this study showed that stimuli belonging to the same category induced response patterns in mildly active voxels that were more correlated than the response patterns induced by stimuli

## Introduction

belonging to different categories. The important implication of this study is that one can determine the category of a stimulus from the response pattern it induces in a population of voxels. Importantly, such a multi-voxel pattern analysis allows one to obtain information about stimuli from the response patterns of voxels that would be labeled as "unresponsive" by a univariate analysis.

The study of Haxby and colleagues (2001) inspired Kamitani and Tong (2005) to test whether the examination of multi-voxel response patterns allows one to classify the orientation of visual stimuli. From electrophysiological and optical imaging studies in animal models it is known that visual cortex contains cortical columns covering around 0.5mm<sup>2</sup> of cortex in which neurons are activated selectively by visual stimuli with a specific orientation (Hubel and Wiesel, 1963; Hubel and Wiesel, 1968; Mountcastle, 1997). Given the size and distribution of these orientation columns, one voxel covering 3x3x3 mm of visual cortex can be expected to sample responses from dozens of orientation columns preferring a variety of stimulus orientations. Nonetheless, Kamitani and Tong (2005) were able to decode the orientation of a visual stimulus from response patterns of voxels in early visual areas V1 to V4. Hence, a linear support vector machine algorithm (Suykens and Vandewall, 1999) was able to discriminate between the response patterns in these visual areas induced by stimuli with eight different orientations. The fact this was possible Kamitani and Tong (2005) attributed to voxels sampling a slightly greater amount of columns preferring one orientation than all the others by chance. As a consequence responses of most voxels can be expected to express a weak orientation selectivity. This orientation selectivity is too weak to be picked up with univariate analysis. However, pooling the information present in this weak orientation selectivity across a large population of voxels provides enough information for a linear support vector machine to classify the orientation of stimuli above chance level. In a following study it was shown that the motion direction of visual stimuli could be decoded in the same way (Kamitani and Tong, 2006).

In conclusion, multi-voxel pattern analysis allows for the assessment of information present in voxels which show only a weak response to a stimulus manipulation. An interesting feature of this analysis is that it enables one to measure response properties of neural populations covering only a fraction of the volume of a single voxel.

## **Auditory motion direction encoding in the human brain**

The discovery of motion direction sensitive neurons in primary visual cortex by Hubel and Wiesel (1963) sparked an interest in whether also neurons exist whose

responses encode auditory motion direction. Electrophysiological studies followed (Sovijärvi and Hyvärinen, 1974; Ahissar et al., 1992; Stumpf et al., 1992; Poirier et al., 1997) that identified neurons that respond selectively to a specific direction of auditory motion in the auditory cortex of cats and monkeys. Neuroimaging studies have also investigated auditory motion processing in the human brain and discovered that the planum temporale as well as regions in parietal and frontal cortex are activated more strongly when one perceives auditory motion (Griffiths et al., 1994; Griffiths et al., 1998; Baumgart et al., 1999; Griffiths et al., 2000; Warren et al., 2002; Pavani et al., 2002; Ducommun, 2002; Xiang et al., 2002; Zimmer et al., 2009). These findings indicate that there are human brain areas whose responses are enhanced by the presence of auditory motion. At present, however, neuroimaging studies have not been able to demonstrate auditory motion direction encoding in the human brain.

The main purpose of the experiment described in chapter 2 was to identify human brain areas whose responses encode the direction of moving sounds. Our experiment was inspired by the study of Kamitani and Tong (2006) demonstrating that one can determine the direction of movement of visual stimuli by analyzing multi-voxel response patterns in visual cortex. Because an electrophysiological study in the cortex of the cat indicates that auditory motion direction is encoded in a similar way as visual motion direction (Stumpf et al., 1992), we expected that sensitivity of brain areas to auditory motion direction would be expressed in multi-voxel response patterns to moving sounds.

In contrast to the study of Kamitani and Tong (2006) we had no a priori hypothesis about where in the brain motion direction would be encoded in response patterns. Therefore we used a “spherical searchlight” approach introduced by Kriegeskorte and colleagues (2006). This approach allows one to assess the information content of local response patterns across the entire brain. This is realized by sequentially selecting voxels inside a sphere centered on each position in the brain and assessing the information content of the response patterns inside these voxels based on the performance of a support vector machine (see figure 1c and the methods section of chapter 2 for more details).

With this analysis we were able to identify two regions in the human cerebral cortex in which activation patterns differed between left- and rightward auditory motion. One of these regions was the right primary auditory cortex. Therefore, this study indicates that the human primary auditory cortex, as its equivalent in cats and monkeys, processes the directionality of auditory motion. The second region affected by auditory motion direction was located in extrastriate visual cortex in the right occipital lobe. Interestingly, this region was located directly adjacent to the location typically reported for the visual area hMT/V5+ which is thought to be a cortical module supporting visual motion

perception (Tootell et al., 1995) This suggests that information about the direction of auditory motion is most likely available to hMT/V5+ which would imply that this area is sensitive to the direction of both visual and auditory motion.

### **The effect of motion direction congruency of audiovisual motion stimuli**

If information about auditory motion direction is passed on to the visual cortex, does this information have an effect on how visual motion signals are processed in the visual motion area hMT/V5+? This we have tested in an fMRI experiment described in chapter 3. During this experiment subjects were presented with moving visual stimuli and moving sounds simultaneously and we tested whether responses of hMT/V5+ to moving visual stimuli were affected by the direction in which sounds were moving.

Indeed, the direction of auditory motion was found to affect visual responses in hMT/V5+. When sounds moved in the same direction as visual stimuli, responses in hMT/V5+ were elevated as compared to the responses to visual stimuli and sounds moving in opposite directions. The design of this experiment excluded the possibility that this interaction was caused by within-sensory-modality effects.

Another interesting finding of this study was that subjects reported that sounds moved in the same direction as visual stimuli in a significant proportion of trials containing sounds and visual stimuli moving in opposite directions. This finding is in line with previous psychophysical studies showing that visual motion tends to capture the perceived direction of auditory motion (Soto-Faraco et al., 2004), which is an auditory illusion referred to as cross-modal dynamic capture. When this illusion occurred and our subjects reported perceiving coherent audiovisual motion although motion direction was conflicting across senses, hMT/V5+ responses increased. Interestingly, hMT/V5+ responses associated with this illusory percept were indistinguishable from responses to audiovisual stimuli whose motion direction was physically coherent. The planum temporale, on the other hand, was found to exhibit a reduced response when subjects experienced the cross-modal dynamic capture illusion.

These findings indicate that the processing of visual motion signals in hMT/V5+ is affected by the direction of simultaneously presented auditory motion and that the processing of auditory motion signals in the planum temporale can be altered by an auditory illusion that is induced by visual motion. Both these findings indicate that the human brain processes auditory and visual motion signals in an interactive manner.

## **Visual motion capturing auditory motion, and vice versa?**

In chapter 3 we show that auditory motion has an effect on responses in the visual motion area hMT/V5+. hMT/V5+ activation has been associated with our conscious experience of visual motion (Tootell et al., 1995; Kamitani and Tong, 2006). Does this suggest that auditory motion can affect our conscious experience of visual motion? Previous studies have shown that visual motion can affect the perceived direction of auditory motion (Soto-Faraco et al., 2002; Soto-Faraco et al., 2003; Soto-Faraco et al., 2004). However, it remains to be shown that auditory motion direction can affect the perceived direction of visual motion. In the experiment described in chapter 4 we try to find an answer to the question if auditory motion can affect visual motion perception.

In order to address this question it is important to consider the factors that determine how sensory input received in one sensory modality affects the interpretation of sensory signals received via another modality. At present, the most influential models in this regard describe such cross-modal interactions as resembling Bayesian integration (Pouget et al, 2002; Battaglia et al., 2003; Knill and Pouget, 2004). The gist of these models is that multisensory perception of a stimulus feature is based on a weighted average of sensory signals in proportion to each signal's relative reliability in conveying information about the feature. Because vision can more reliably convey information about motion direction than audition (McKee and Nakayama, 1984; Perrott and Musicant, 1977), it is not surprising that previous studies (Soto-Faraco et al., 2002; Soto-Faraco et al., 2003; Soto-Faraco et al., 2004) have not been able to demonstrate an effect of auditory motion on visual motion perception as their visual stimuli much more reliably conveyed information about motion direction than the auditory stimuli. For such an effect to occur, visual motion direction should be made ambiguous so that it provides information about motion direction less reliably than auditory motion.

To this end, we presented subjects with a variety of visual apparent motion stimuli (Kohlers, 1963; Braddick, 1974) and determined for each subject the stimulus configuration that induced a fully bistable visual motion percept - a configuration which was equally likely to induce the percept of leftwards as rightwards visual motion. Subsequently, we tested whether the direction of auditory apparent motion, when presented simultaneously with a fully bistable visual apparent motion stimulus, biases the perceived direction of visual motion. The outcome of this experiment shows that bistable visual apparent motion stimuli were more often perceived as moving in the same direction as auditory apparent motion than in the opposite direction. From this we conclude that auditory motion can affect our conscious experience of visual motion.

Therefore, chapter 4 shows that interactions between brain areas supporting auditory and visual motion perception are accompanied by audiovisual interactions on a perceptual level. In addition, this study, in conjunction with previous research (Soto-Faraco et al., 2002; Soto-Faraco et al., 2003; Soto-Faraco et al., 2004), indicates that auditory and visual motion perception interact with each other in a bidirectional manner.

### **Interaction between auditory and visual motion processing - concluding remarks**

The main aim of this dissertation is to determine whether the human brain processes auditory and visual motion signals in an independent or in an interactive manner. The two studies described in chapter 2 and 3 provide three indications for an interaction between auditory and visual motion processing in the human brain. First of all, it is shown that the direction in which sounds move can be decoded from neural response patterns in extrastriate visual cortex directly adjacent to the visual motion area hMT/V5+. Second, it is shown that the amplitude of neural responses in the visual motion area hMT/V5+ to moving visual stimuli is affected by the relative direction of simultaneously presented moving sounds. Third, it is shown that when visual motion affects the direction in which auditory motion is perceived hMT/V5+ responses to moving visual stimuli are enhanced while responses in the auditory motion area, the planum temporale, are reduced. Furthermore, in chapter 4 it becomes apparent that there is also an interaction between visual and auditory motion processing on a perceptual level. This is implied by the finding that moving sounds can affect the perceived directionality of visual motion.

These findings indicate that visual and auditory motion perception can not be considered as two strictly modular cognitive facilities as our findings indicate violations of the domain specificity criterion and the information encapsulation criterion for modularity (Fodor, 1983). It is tempting to speculate, based on these findings, that seeing and hearing motion might be facilitated by a common set of brain regions. Such a conclusion would fit well with the finding that in blind subjects the visual motion area hMT/V5+ exhibits motion direction specific responses to moving sounds (Poirier et al., 2006; Saenz et al, 2008; Wolbers et al., 2010) as this implies that hMT/V5+ can support both visual and auditory motion perception. However, it has not yet been shown that neural responses in hMT/V5+ are critical for auditory motion perception in blind or sighted subjects nor has it been shown that activation in auditory cortex is of importance for visual motion perception. Therefore, the final evidence for the existence of brain regions supporting both visual and auditory motion perception is still lacking.

Such evidence, however, could be obtained in a future study that assesses whether disrupting neural activation in hMT/V5+ with Transcranial Magnetic Stimulation (TMS) impairs one's ability to determine the direction in which sounds move.

## **Predictability of visual stimuli in the context of apparent motion**

The last chapter of this dissertation relates to the idea that cortical processing of visual stimuli strongly depends on visual context and visual expectations (Mumford, 1992; Rao and Ballard, 1999; Erlhagen, 2003; Bar, 2007; Enns 2008; Bar, 2009; Friston 2009; Grossberg, 2009). One model specifying how visual context affects visual processing in the cerebral cortex is the predictive coding model of Rao and Ballard (1999). The main implication of this model is that higher-level visual cortex, based on previous input and contextual cues, attempts to predict the input to lower-level visual cortex and sends feedback to lower-level visual cortex conveying these predictions. In chapter 5 we test one main claim made by this model that activation levels in early visual areas are reduced when visual input is predictable.

This we have tested by assessing whether responses in primary visual cortex (V1), the first relay for visual processing in the cerebral cortex, are reduced when a stimulus is predictable. To this end, we presented subjects with long-range apparent-motion-inducing stimuli - two consecutively presented spatially separated flashes that induce an illusory percept of motion in the space in between the flashes (Kohlers, 1963; Braddick, 1974). In one of the experiments described in chapter 5 we presented a brief flash at the location where subjects perceived illusory motion. Critically, this flash occurred either at the moment that illusory motion passed by this location or slightly later. As a result, the presentation time of the first type of flash is predicted better by the trajectory of illusory motion than that of the second type of flash. In concordance with Rao and Ballard's (1999) predictive coding model we observed a lower V1 response for the predictable flash.

We moreover assessed whether V1 responses parametrically decrease when stimulus predictability is parametrically increased. To this end we presented moving dots on the path of apparent motion and we varied the direction of the dot motion from being orthogonal to the direction of apparent motion (highly unpredictable) to being parallel to the direction of apparent motion (highly predictable) in four steps. As expected, V1 responses were found to systematically decrease as the direction of dot motion deviated less from the direction of surrounding illusory motion.

These findings indicate that a stimulus whose timing or motion direction can be predicted from its visual context induces less activation in primary visual cortex. This finding is in line with the predictive coding model of Rao and Ballard (1999) as well as with the older and more general idea of efficient coding put forward by Attneave (1953) as our findings indicate that the brain represents stimuli more economically when aspects of stimuli are redundant given the context in which they are presented.

## References

- Ahissar M, Ahissar E, Bergman H, Vaadia E (1992) Encoding of sound-source location and movement: activity of single neurons and interactions between adjacent neurons in the monkey auditory cortex. *Journal of neurophysiology* 67:203
- Attneave F (1954) Some informational aspects of visual perception. *Psychological review* 61:183–193
- Bar M (2007) The proactive brain: using analogies and associations to generate predictions. *Trends in Cognitive Sciences* 11:280–289
- Bar M (2009) The proactive brain: memory for predictions. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364:1235
- Battaglia PW, Jacobs RA, Aslin RN (2003) Bayesian integration of visual and auditory signals for spatial localization. *Journal of the Optical Society of America A* 20:1391–1397
- Baumgart F, Gaschler-Markefski B, Woldorff MG, Heinze HJ, Scheich H (1999) A movement-sensitive area in auditory cortex. *Nature* 400:724–726
- Beckers, G., Zeki, S. (1995). The consequences of inactivating areas V1 and V5 on visual motion perception. *Brain*, 118(1), 49-60
- Boynton GM, Engel SA, Glover GH, Heeger DJ (1996) Linear systems analysis of functional magnetic resonance imaging in human V1. *Journal of Neuroscience* 16:4207
- Braddick O (1974) A short-range process in apparent motion. *Vision Research* 14:519–527
- Ducommun CY, Murray MM, Thut G, Bellmann A, Viaud-Delmon I, Clarke S, Michel CM (2002) Segregated processing of auditory motion and auditory location: an ERP mapping study. *Neuroimage* 16:76–88
- Enns JT, Lleras A (2008) What's next? New evidence for prediction in human vision. *Trends in cognitive sciences* 12:327–333
- Erlhagen W (2003) Internal models for visual perception. *Biological cybernetics* 88:409–417
- Fodor JA (1983) *The modularity of mind*. MIT press Cambridge, MA.

## Chapter 1

- Friston K, Kiebel S (2009) Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364:1211
- Griffiths TD, Green GG, Rees A, Rees G (2000) Human brain areas involved in the analysis of auditory movement. *Human Brain Mapping* 9:72–80
- Griffiths TD, Rees G, Rees A, Green GG, Witton C, Rowe D, Büchel C, Turner R, Frackowiak RS (1998) Right parietal cortex is involved in the perception of sound movement in humans. *nature neuroscience* 1:74–79
- Griffiths TD, Bench CJ, Frackowiak RSJ (1994) Human cortical areas selectively activated by apparent sound movement. *Current Biology* 4:892–895
- Grossberg S (2009) Cortical and subcortical predictive dynamics and learning during perception, cognition, emotion and action. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364:1223
- Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293:2425
- Hubel DH, Wiesel TN (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology* 160:106
- Hubel DH, Wiesel TN (1968) Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology* 195:215
- Kamitani Y, Tong F (2005) Decoding the visual and subjective contents of the human brain. *Nature Neuroscience* 8:679–685
- Kamitani Y, Tong F (2006) Decoding seen and attended motion directions from activity in the human visual cortex. *Current Biology* 16:1096–1102
- Knill DC, Pouget A (2004) The Bayesian brain: the role of uncertainty in neural coding and computation. *TRENDS in Neurosciences* 27:712–719
- Koehler RC, Roman RJ, Harder DR (2009) Astrocytes and the regulation of cerebral blood flow. *Trends in neurosciences* 32:160–169
- Kolers PA (1963) A difference between real and apparent visual movement. *Nature* 197:271-272
- Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping. *Proceedings of the National Academy of Sciences* 103:3863
- Krumbholz, K., Schönwiesner, M., Rübsem, R., Zilles, K., Fink, G. R., von Cramon, D. Y. (2005). Hierarchical processing of sound location and motion in the human brainstem and planum temporale. *European Journal of Neuroscience*, 21(1), 230–238
- Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A (2001) Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412:150–157

## Introduction

- Mansfield P, Sc AA (2005) Medical imaging by NMR. *Classic papers in modern diagnostic radiology* 50:104
- McKee SP, Nakayama K (1984) The detection of motion in the peripheral visual field. *Vision Research* 24:25–32
- Mountcastle VB (1997) The columnar organization of the neocortex. *Brain* 120:701
- Mumford D (1992) On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biol Cybern* 66:241–251
- Pavani F, Macaluso E, Warren JD, Driver J, Griffiths TD (2002) A common cortical substrate activated by horizontal and vertical sound movement in the human brain. *Current Biology* 12:1584–1590
- Perrott DR, Musicant AD (1977) Minimum auditory movement angle: Binaural localization of moving sound sources. *The Journal of the Acoustical Society of America* 62:1463
- Poirier C, Collignon O, Scheiber C, Renier L, Vanlierde A, Tranduy D, Veraart C, De Volder AG (2006) Auditory motion perception activates visual motion areas in early blind subjects. *Neuroimage* 31:279–285
- Poirier P, Jiang H, Lepore F, Guillemot JP (1997) Positional, directional and speed selectivities in the primary auditory cortex of the cat. *Hearing research* 113:1–13
- Pouget A, Deneve S, Duhamel JR (2002) A computational perspective on the neural basis of multisensory spatial representations. *Nature Reviews Neuroscience* 3:741–747
- Rao RP, Ballard DH (1999) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *nature neuroscience* 2:79–87
- Saenz M, Lewis LB, Huth AG, Fine I, Koch C (2008) Visual motion area MT+/V5 responds to auditory motion in human sight-recovery subjects. *Journal of Neuroscience* 28:5141
- Soto-Faraco S, Kingstone A, Spence C (2003) Multisensory contributions to the perception of motion. *Neuropsychologia* 41:1847–1862
- Soto-Faraco S, Lyons J, Gazzaniga M, Spence C, Kingstone A (2002) The ventriloquist in motion: Illusory capture of dynamic information across sensory modalities. *Cognitive Brain Research* 14:139–146
- Soto-Faraco S, Spence C, Kingstone A (2004) Congruency effects between auditory and tactile motion: extending the phenomenon of cross-modal dynamic capture. *Cognitive, Affective, & Behavioural Neuroscience* 4:208
- Sovijärvi AR, Hyvärinen J (1974) Auditory cortical neurons in the cat sensitive to the direction of sound source movement. *Brain research* 73:455

## Chapter 1

- Stumpf E, Toronchuk JM, Cynader MS (1992) Neurons in cat primary auditory cortex sensitive to correlates of auditory motion in three-dimensional space. *Experimental Brain Research* 88:158–168
- Suykens JA, Vandewalle J (1999) Least squares support vector machine classifiers. *Neural processing letters* 9:293–300
- Tootell RB, Reppas JB, Kwong KK, Malach R, Born RT, Brady TJ, Rosen BR, Belliveau JW (1995) Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *Journal of Neuroscience* 15:3215
- Warren JD, Zielinski BA, Green GG, Rauschecker JP, Griffiths TD (2002) Perception of sound-source motion by the human brain. *Neuron* 34:139–148
- Wolbers T, Zahorik P, Giudice NA (2010) Decoding the direction of auditory motion in blind humans. *NeuroImage*
- Xiang J, Chuang S, Wilson D, Otsubo H, Pang E, Holowka S, Sharma R, Ochi A, Chitoku S (2002) Sound motion evoked magnetic fields. *Clin Neurophysiol* 113:1-9
- Zimmer U, Macaluso E (2009) Interaural temporal and coherence cues jointly contribute to successful sound movement perception and activation of parietal cortex. *NeuroImage* 46:1200–1208

## **Chapter 2**

# Auditory motion direction encoded in human primary auditory and high-level visual cortex

Based on the manuscript:

Auditory motion direction encoded in human primary auditory and high-level visual cortex. Arjen Alink, Felix Euler, Nikolaus Kriegeskorte, Wolf Singer and Axel Kohler

Submitted on 10-05-10 as a brief communication to the Journal of Neuroscience

**In the present functional magnetic resonance imaging (fMRI) study, we searched for human brain areas encoding the direction of auditory motion using multivoxel pattern analysis. Our results show that the direction of sound source movement can be decoded from fMRI response patterns in the right primary auditory cortex and right lateral occipital cortex. This suggests that directional selectivity for auditory motion is already present at the first cortical stage of auditory processing in humans and is also detectable in higher-level visual cortex.**

Psychophysical studies have demonstrated that human observers can determine the direction in which an object is moving relying solely on auditory cues (Harris et al., 1971; Perrott and Musicant, 1977; Grantham, 1989). Several previous human neuroimaging studies have investigated the neural mechanisms that support this perceptual ability (Griffiths et al., 1994, 1998, 1999, 2000; Baumgart et al., 1999; Lewis et al., 2000; Pavani et al., 2002; Warren et al., 2002; Zimmer and Macaluso, 2009) by identifying human brain areas responding more vigorously to moving than to static sounds. Such areas were found in the posterior planum temporale as well as in frontal and parietal cortices. Based on these findings it has been claimed that auditory motion perception is supported by activation within this network of cortical areas.

However, differences between activations to moving and static auditory stimuli can be confounded by attentional and other cognitive factors. In addition, the comparison between average activation levels might overlook areas that contain neurons responding selectively to a specific auditory motion direction which are intermingled with neurons expressing a different directional selectivity. Such an intermingled neural architecture is typical of primary sensory areas (Livingstone and Hubel, 1988) and electrophysiological findings in cats (Stumpf et al., 1992) suggest that auditory motion encoding in the primary auditory cortex is consistent with this architecture. For these reasons we decided to investigate auditory motion encoding in the human brain using multivoxel pattern analysis (Haxby et al., 2001; Kamitani and Tong, 2005; Haynes and Rees, 2005; Norman et al. 2006), an fMRI analysis technique that allows for the identification of areas encoding stimulus features in fine-grained neural response patterns. We applied this technique in a whole-brain searchlight approach (Kriegeskorte et al., 2006, 2007) to find areas where fMRI response patterns provide information about the direction in which sounds are moving.

## Materials and methods

### Participants

Nineteen healthy volunteers participated in this fMRI study (age range: 20–31 years; 13 females). All subjects had normal hearing and normal or corrected-to-normal vision. All subjects gave their informed consent after being introduced to the experimental procedure in accordance with the Declaration of Helsinki.

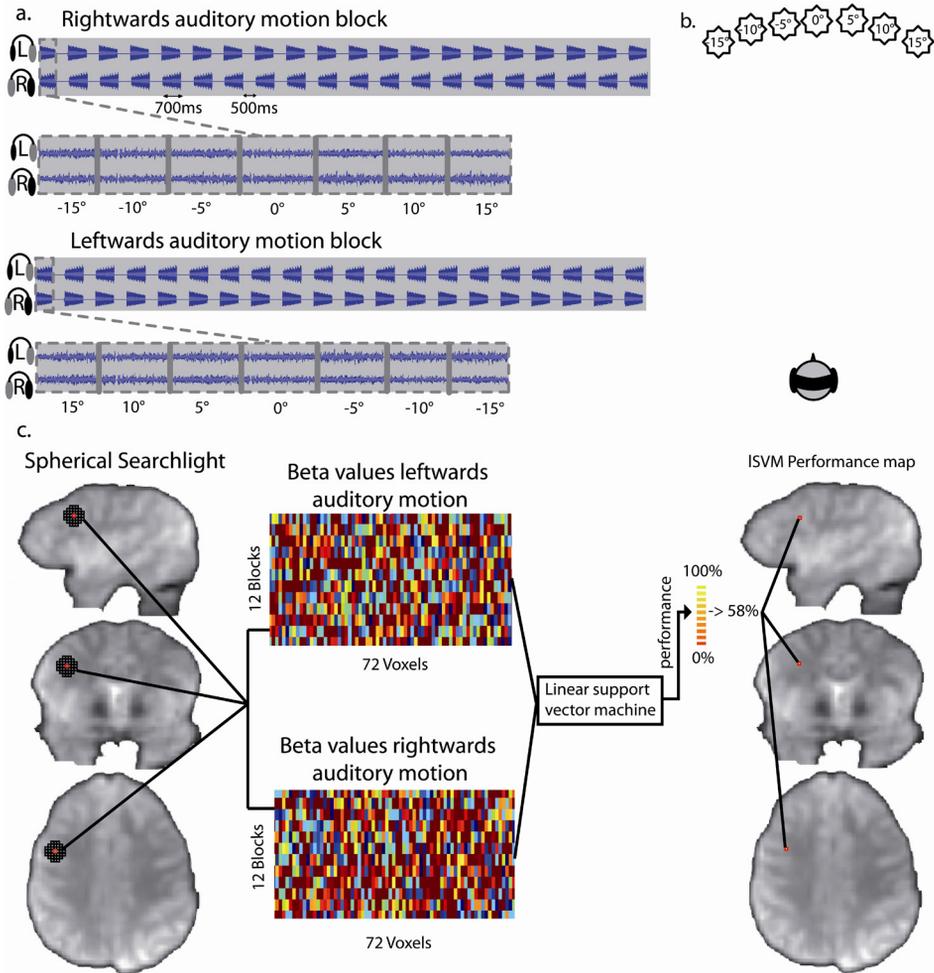
### Stimuli and Task

Visual stimuli were presented using an MR-compatible goggle system with two organic light-emitting diode displays (MR Vision 2000, Resonance Technology, Northridge, CA), and auditory stimulation was realized using an MR-compatible headphone system (Commander XG, Resonance Technology, Northridge, CA). The screen had a width of 30° and a height of 22.5°, and the luminance of the gray background was 24 cd/m<sup>2</sup>.

Auditory stimulation consisted of seven types of 100-ms-long pink-noise bursts. Subjects perceived these pink-noise bursts as originating 15°, 10° and 5° to the left or to the right or in front of the head in the horizontal plane with a distance of 1.4 m. This spatial perception was induced by convolving the pink-noise bursts with a generic head-related transfer function (Wightman and Kistler, 1989) derived from the KEMAR head model (Gardner and Martin, 1994) as previously implemented by Altmann and colleagues (2007) and in our lab (Alink et al., 2008). Leftwards auditory motion sweeps contained each noise burst once going from the outer right to the outer left position and vice versa for rightwards motion sweeps. This resulted in motion sweeps covering an arc of 30° traversed in 700 ms (speed = 43 degrees/second) containing exactly the same sounds for left- and rightwards motion sweeps. Hence, only the order in which the sounds were presented differed across conditions (see Figure 1A).

Motion sweeps of one directionality were presented in blocks containing twenty sweeps, which were presented with an inter-stimulus-interval of 500 ms. This interval was introduced to avoid that subjects would perceive vivid apparent motion in the direction opposite to the motion-sweep direction in between stimulus presentations. This resulted in blocks containing left- or rightwards motion sweeps lasting 24 seconds. During each of the four functional runs, we presented each type of auditory-motion block three times. Furthermore, each run contained three blocks of left- and three blocks of rightwards visual motion with an identical duration. The visual data are not reported here and are used for separate analyses. The order of all types of stimulation blocks was randomized for each run. Between stimulation blocks, there were 24-second periods during which no auditory stimuli were presented, which served to

## Chapter 2



**Figure 1. Experimental Design and Analysis**

(a) Sound stimuli used in the experiment that induced the percept of left- or rightwards auditory motion in blocks containing 20 motion sweeps covering an arc of 30°. (b) Schematic display of the sound locations covered during auditory motion sweeps. (c) Overview of the searchlight-based multivoxel pattern analysis employed in this study.

assess a baseline signal. The visual-attention control task, however, continued during these periods.

During stimulation blocks as well as during the baseline periods, subjects continuously performed a visual-attention control task. This involved fixating

5.6° below the center of the screen, where a stream of letters and numbers appeared at a rate of two symbols per second. The task of the subjects was to press a button with their right index finger as soon as possible when a number appeared. During each stimulation or baseline period, nine numbers appeared with at least one letter being presented between two consecutive numbers. We registered button presses as hits if they occurred within 100 ms and 1000 ms after the number onset.

### **MRI Data Acquisition**

Functional and anatomical MRI data were acquired with a 3T-MRI system (Siemens Allegra; Siemens, Erlangen, Germany) using a four-channel head coil. For each subject, we obtained 300 volumes containing 40 slices covering the entire brain during each of the four functional scans using a gradient-echo echo-planar-imaging (EPI) sequence [repetition time (TR), 2000 ms; echo time (TE), 25 ms; flip angle (FA), 70°; voxel size, 3.28\*3.28\*3.0 mm; field of view (FOV), 210 mm; gap thickness, 0.3 mm]. We corrected for spatial distortions in the EPI images using a point-spread-function (PSF) sequence (Zaitsev et al., 2004). We also obtained a T1-weighted anatomical scan for each of the subjects using a Siemens magnetization-prepared rapid-acquisition gradient echo (MPRAGE) sequence (1\*1\*1 mm).

### **Data Analysis**

Functional as well as anatomical MRI data were analyzed using the Brainvoyager QX software package (Brain Innovation, Maastricht, The Netherlands). The first four volumes of the functional runs were discarded to preclude T1 saturation effects. After pre-processing (motion correction, linear trend removal, temporal high pass filtering at 0.01 Hz and slice-scan-time correction), functional data for all subjects were aligned with the individual high-resolution anatomical MPRAGE image and transformed into Talairach space (Talairach and Tournoux, 1988) interpolating the data to a four-dimensional matrix (three for space, one for time) containing 3.0 mm isotropic voxels using trilinear interpolation.

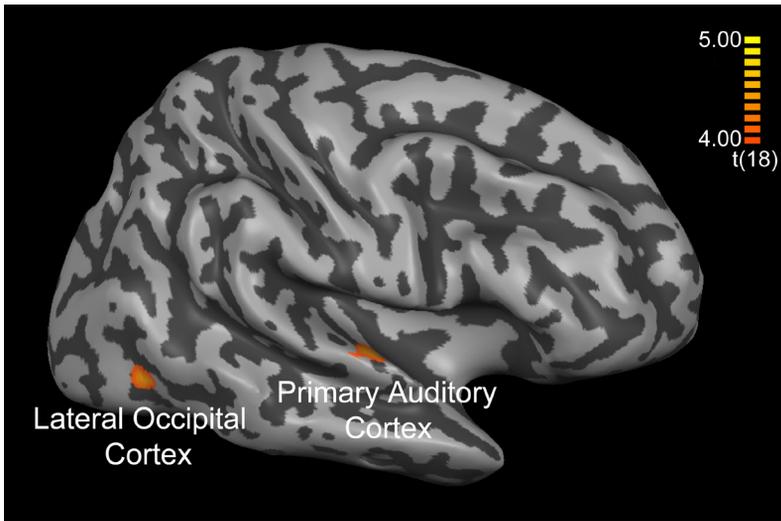
A multivoxel pattern analysis, using a spherical searchlight approach (Kriegeskorte et al., 2006, 2007), was performed over these data separately for each subject using custom-made code programmed in Matlab (The Mathworks, Inc, Natick, US). The first step of this analysis consisted of defining spherical searchlights centered on each single voxel in Talairach space. These searchlights contained the 72 voxels that were inside a radius of 1.05 cm around the center voxel. As a first step, we subtracted the average time course across voxels (searchlight mean) from each single voxel time course to ensure that

homogeneous univariate effects would not influence classification. Within each searchlight, we then determined the response patterns that were evoked by each single auditory-motion stimulus block using a general linear model (GLM). This resulted in twelve beta-value vectors (corresponding to the number of blocks) with a length of 72 units (corresponding to the number of searchlight voxels) for each auditory motion direction. We assessed whether a linear support-vector machine (LSVM) could classify motion direction based on pattern differences between auditory motion directions. To this end, we used an LSVM defined in LIBSVM (Chang et al., 2001). The LSVM was trained on beta-value patterns for eleven left- and eleven rightwards blocks, after which the LSVM attempted to classify the direction of the remaining two blocks. This procedure was performed twelve times using each pair of blocks once for testing. The output of this analysis was the average performance over these twelve classifications. Performance of the LSVM for each searchlight was stored in Talairach space with each searchlight projecting its performance to the position of its centre voxel (see Figure 1C). The outcome of this analysis were nineteen individual performance maps aligned in Talairach space.

To assess whether a region in Talairach space contained directional information we used a random-effects analysis to test whether performance for this region was consistently higher than chance level (50%) across subjects after spatially smoothing the individual performance maps using a Gaussian kernel (6 mm FWHM). A *t*-threshold of 4.0 was used in conjunction with a cluster threshold which required *t*-values of at least four adjacent voxels to exceed the *t*-threshold. This cluster threshold was computed using the method introduced by Forman and colleagues (1995) and implemented in BrainVoyager QX by Fabrizio Esposito and Rainer Goebel (University of Maastricht, Maastricht, The Netherlands) and corresponds to a *p*-value lower than 0.001 corrected for multiple comparisons. The cluster threshold was selected from a range of cluster thresholds after determining the false-positive rate for these thresholds over a thousand randomly generated statistical maps with the same spatial smoothness as the map acquired in our group analysis. We also performed a typical per-voxel GLM analysis, a region-of-interest-based analysis and generated event-related averages.

## Results

Auditory stimuli were presented in blocks containing twenty horizontal motion sweeps traversing an arc of 30° centred in front of the subject's head. For each motion direction, twelve blocks were presented and we determined for each spatial unit (voxel) the response amplitude evoked by each of the 24 blocks.



**Figure 2. Results of Searchlight-Based Multivoxel Pattern Analysis**

Group statistics projected on a cortical reconstruction of the right hemisphere of one of the subjects. T-values indicate the extent to which performance for a location in Talairach space was higher than chance level (50%) across our nineteen subjects. Significant areas displayed on this map are a region in the right auditory cortex (mean ISVM performance = 56.9%, SEM 1.4%) and a region in the lateral occipital cortex (mean ISVM performance 56.2%, SEM 1.1%).

Subsequently, we quantified the directional information contained in local activation patterns. This was achieved by assessing how well a linear support-vector machine (ISVM; Chang and Lin, 2001) could classify motion direction from local activation patterns within subjects (Figure 1, for more details on methods and analysis see Experimental Procedures). Training and classification was always performed over local activation clusters, which were selected using a “spherical searchlight” (Kriegeskorte et al., 2006; Kriegeskorte et al., 2007). This approach allowed us to analyze motion-direction information across the entire brain. The result of this analysis was a map in Talairach space (Talairach and Tournoux, 1988) for each subject containing performance values of the ISVM. For all positions in Talairach space, we assessed whether directional information was present by testing whether ISVM performance was consistently above chance level (50%) across subjects.

Our group analysis identified two cortical regions whose activation patterns contained directional information ( $p < .001$ , corrected, Figure 2). One of these regions was located near the right Heschl’s gyrus (mean ISVM performance

56.9%, SEM 1.4%, Talairach coordinates:  $x = 54$ ,  $y = -13$   $z = 7$ ). The coordinates of this region fall within the area previously reported for primary auditory cortex ( $x = 32-57$ ,  $y = -28-3$   $z = 0-16$ ; Rademacher et al., 2001). The second region was located in the right lateral occipital cortex (mean ISVM performance 56.2%, SEM 1.1, Talairach coordinates:  $x = 35$ ,  $y = -67$   $z = -8$ ).

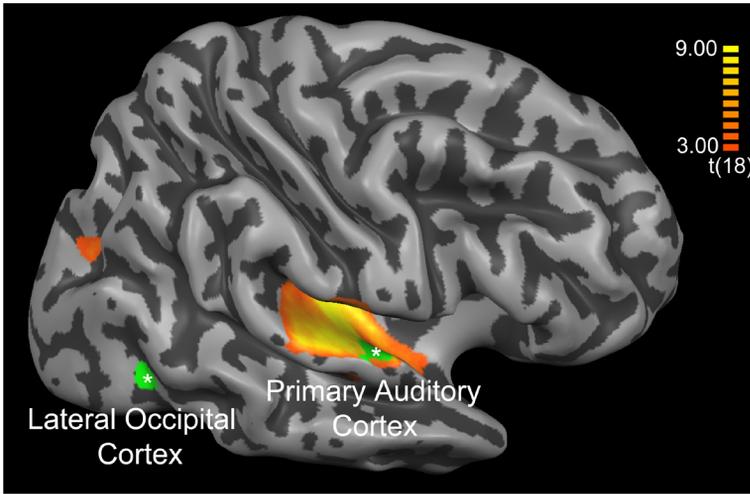
We also performed a conventional univariate analysis where we looked for single voxels in which BOLD responses differed between left- and rightwards motion blocks. No such voxels were found ( $p > .05$  corrected). A region-of-interest-based analysis over the two regions identified by the multivariate analysis also showed no univariate effect of motion direction ( $p > .05$  corrected). The region close to Heschl's gyrus responded robustly to the auditory motion stimuli ( $p < .001$ , corrected), but no BOLD response was detectable in the lateral occipital cortex ( $p > .05$ , uncorrected; Figure 3B). Thus, activation patterns in this area contained directional information about sound-source motion in the absence of a significant average BOLD response. Figure 3A shows a comparison between univariate activation and multivariate classification maps. Similar results of significant decoding performance for non-activated regions have been described previously (Serences and Boynton, 2007; Harrison and Tong, 2007).

To control for attention effects, subjects were asked to fixate on a stream of alphanumeric characters presented at a frequency of two Hz and to press a button when they saw a number appearing among a stream of letters. We then compared task performance between blocks of left- and rightward movements. There was no significant difference ( $p > .05$ ) in the percentage of correct number detection (rightwards: 94.2%, SEM = 2.7%; leftwards: 94.7%, SEM 2.4%) and reaction times ( $p > .05$ ; rightwards: 421ms, SEM = 6ms; leftwards: 418ms, SEM = 6ms).

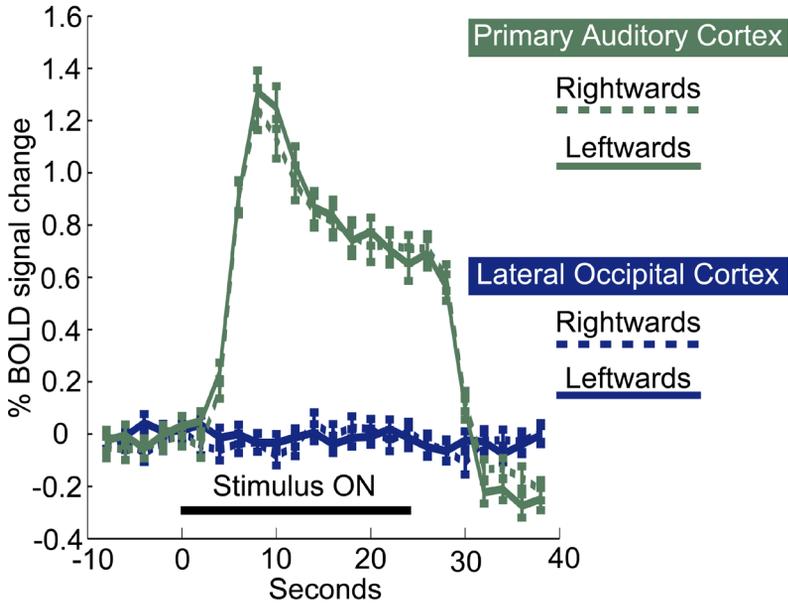
## Discussion

The present study indicates that activation patterns in the right primary auditory cortex and right lateral occipital cortex contain information about the direction of auditory motion. Auditory motion sensitivity appears to arise at an earlier stage of sound processing than suggested by previous human neuroimaging studies claiming that auditory motion sensitivity only arises at the level of the planum temporale or later (Griffiths et al., 1994, 1998, 1999, 2000; Baumgart et al., 1999; Lewis et al., 2000; Pavani et al., 2002; Warren et al., 2002; Zimmer and Macaluso, 2009). Our results are consistent with electrophysiological research in animal

A



B



**Figure 3. Comparison of Multivariate and Univariate Results**

(A) Overlay of areas with significant multi-voxel pattern classification from Figure 1 (green, marked with asterisk) on a random-effects map of univariate responses to auditory stimulation (average of rightwards and leftwards auditory motion). (B) Event-related average responses to rightwards and leftwards auditory motion in primary auditory cortex (dark green) and lateral occipital cortex (blue).

models showing that the location of auditory stimuli is already encoded in the brainstem (Goldberg and Brown, 1969) and that auditory-motion-direction sensitivity is a feature of neurons in the primary auditory cortex of cats and monkeys (Sovijärvi and Hyvärinen, 1974; Ahissar et al., 1992; Stumpf et al., 1992). Furthermore, this finding agrees with the recent observation that complex auditory stimulus features, like speech content and speaker identity, are represented already at early stages of sound processing (Formisano et al., 2008).

The finding of auditory-motion-direction sensitivity in the extrastriate visual cortex suggests that information about auditory motion direction is passed on to visual cortex. This finding is in line with other human neuroimaging studies that have demonstrated that extrastriate visual cortex, including the human motion complex hMT/V5+, is sensitive to the presence of auditory motion (Porier et al., 2005; Porier et al., 2006; Alink et al., 2008; Saenz et al., 2008). The study at hand, however, is the first to show that responses in extrastriate visual cortex are informative with regard to the direction of sound-source motion. Therefore, it seems likely that the human visual motion area hMT/V5+ has access to directional information about auditory stimuli which might account for the observation of previous psychophysical studies that auditory motion direction can bias the perceived direction of visual motion (Meyer et al., 2005; Meyer and Wuerger, 2001; Wuerger et al., 2003; Hidaka et al., 2009; Maeda et al. 2004; Sadaghiani et al., 2009). However, the auditory-motion-direction sensitivity was observed in a lateral occipital region in higher visual cortex, inferior and medial to the typical location of hMT/V5+ (Dumoulin et al., 2000).

Crossmodal convergence in lateral occipital cortex has been previously described in the visuotactile domain for object (Amedi et al., 2001; Lucan et al., 2010) and hand processing (Beauchamp et al., 2009). These regions are thought to support modality-independent representations (Lacey et al., 2009) and multisensory coordination (Beauchamp et al., 2009). In one study, the possible contribution of auditory input to object representations in lateral occipital cortex was explicitly tested, but no significant involvement could be detected with conventional analyses (Amedi et al., 2002). Using more sensitive multivariate techniques, we can show that information concerning the direction of auditory motion is encoded in lateral occipital cortex, despite the absence of univariate responses.

Another interesting outcome is that auditory-motion-direction sensitivity was only observed in the right hemisphere. This is in line with the findings of Baumgart and colleagues (Baumgart et al., 1999) and with neurophysiological evidence for a specific impairment of auditory motion perception after right-hemisphere lesions (Griffiths et al., 1996). Furthermore, this finding is in

agreement with the suggestion that spatial processing tends to occur more dominantly in the right hemisphere (Corballis, 1991).

Multivoxel pattern analysis allowed us to detect auditory-motion direction sensitivity in the human cerebral cortex that would not have been captured by a conventional univariate analysis because none of the motion-sensitive areas reported here responded more strongly to one of the two motion directions. Information being present in areas that do not show an overall BOLD response amplitude difference between stimulus conditions is in line with previous studies (Serences and Boynton, 2007; Harrison and Tong, 2007). This finding raises the question what kind of neural organization allowed us to decode auditory motion direction from multivoxel activation patterns. Kamitani and Tong (2005) suggested that patterns informative about visual grating orientation result from randomly biased sampling of orientation columns in early visual cortex by fMRI voxels. Accordingly, each voxel shows a slight preference for a specific orientation and these small preferences are summed up by multivariate analysis techniques. Results from an electrophysiological study by Stumpf and colleagues (1992) indicate that neurons preferring a specific auditory-motion direction in the cat auditory cortex tend to be clustered together. Our findings suggest that a similar columnar organization might exist in the human auditory cortex.

In conclusion, by analyzing the information content of multivoxel patterns, we showed that auditory motion direction is encoded at the first stage of auditory processing in the human cerebral cortex. We further showed that responses in higher-level visual cortex also contain information about the direction of sound-source motion, suggesting a convergence of motion signals from both modalities in visual cortex.

## References

- Ahissar M, Ahissar E, Bergman H, Vaadia E (1992) Encoding of sound-source location and movement: activity of single neurons and interactions between adjacent neurons in the monkey auditory cortex. *J Neurophysiol* 67:203–215
- Alink A, Singer W, Muckli L (2008) Capture of auditory motion by vision is represented by an activation shift from auditory to visual motion cortex. *J Neurosci* 28:2690-2697
- Altmann CF, Bledowski C, Wibral M, Kaiser J (2007) Processing of location and pattern changes of natural sounds in the human auditory cortex. *Neuroimage* 35:1192-1200

- Amedi A, Jacobson G, Hendler T, Malach R, Zohary E (2002) Convergence of visual and tactile shape processing in the human lateral occipital complex. *Cereb Cortex* 12:1202-1212
- Amedi A, Malach R, Hendler T, Peled S, Zohary E (2001) Visuo-haptic object-related activation in the ventral visual pathway. *Nat Neurosci* 4:324-330
- Baumgart F, Gaschler-Markefski B, Woldorff MG, Heinze HJ, Scheich H (1999) A movement-sensitive area in auditory cortex. *Nature* 400:724-726
- Beauchamp MS, LaConte S, Yasar N (2009) Distributed representation of single touches in somatosensory and visual cortex. *Hum Brain Mapp* 30:3163-3171.
- Chang C-C, Lin C-J (2001) LIBSVM: a library for support vector machines. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- Corballis MC (1991) *The lopsided ape: the evolution of the generative mind*. New York, NY: Oxford University Press
- Dumoulin SO, Bittar RG, Kabani NJ, Baker Jr CL, Le Goualher G, Pike GB, Evans AC (2000) A new anatomical landmark for reliable identification of human area V5/MT: a quantitative analysis of sulcal patterning. *Cereb Cortex* 10:454-463
- Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995) Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn Reson Med* 33:636-647
- Formisano E, De Martino F, Bonte M, Goebel R (2008) "Who" is saying "what"? Brain-based decoding of human voice and speech. *Science* 322:970-973
- Gardner WG, Martin KD (1994) HRTF measurements of a KEMAR dummy-head microphone. MIT Media Lab Perceptual Computing Technical Report 280
- Goldberg JM, Brown PB (1969) Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. *J Neurophysiol* 32:613-636
- Grantham DW (1989) Motion aftereffects with horizontally moving sound sources in the free field. *Percept Psychophys* 45:129-136
- Griffiths TD, Green GGR (1999) Cortical activation during perception of a rotating wide-field acoustic stimulus. *Neuroimage* 10:84-90
- Griffiths TD, Green GGR, Rees A, Rees G (2000) Human brain areas involved in the analysis of auditory movement. *Hum Brain Mapp* 9:72-80
- Griffiths TD, Bench CJ, Frackowiak RSJ (1994) Human cortical areas selectively activated by apparent sound movement. *Curr Biol* 4:892-895
- Griffiths TD, Rees A, Witton C, Shakir RA, Henning GB, Green GGR (1996) Evidence for a sound movement area in the human cerebral cortex. *Nature* 383:425-427

- Griffiths TD, Rees G, Rees A, Green GGR, Witton C, Rowe D, Büchel C, Turner R, Frackowiak RSJ (1998) Right parietal cortex is involved in the perception of sound movement in humans. *Nat Neurosci* 1:74–79
- Harris JD, Sergeant RL (1971) Monaural/binaural minimum audible angles for a moving sound source. *J Speech Lang Hear Res* 14:618–629
- Harrison SA, Tong F (2009) Decoding reveals the contents of visual working memory in early visual areas. *Nature* 458:632–635
- Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293:2425
- Haynes JD, Rees, G (2005). Predicting the orientation of invisible stimuli from activity in human primary visual cortex *Nat. Neurosci.* 8:686–691
- Hidaka S, Manaka Y, Teramoto W, Sugita Y, Miyauchi R, Gyoba J, Suzuki Y, Iwaya Y (2009). Alternation of sound location induces visual motion perception of a static object. *PLoS One* 4:e8188
- Kamitani Y, Tong F (2005) Decoding the visual and subjective contents of the human brain. *Nat Neurosci* 8:679–685
- Kriegeskorte N, Formisano E, Sorger B, Goebel R (2007) Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proc Natl Acad Sci USA* 104:20600–20605
- Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping. *Proc Natl Acad Sci USA* 103:3863–3868
- Lucan JN, Foxe JJ, Gomez-Ramirez M, Sathian K, Molholm S (2010) Tactile shape discrimination recruits human lateral occipital complex during early perceptual processing. *Hum Brain Mapp* epub ahead of print
- Lacey S, Tal N, Amedi A, Sathian K (2009) A putative model of multisensory object representation. *Brain Topogr* 21:269–274
- Lewis JW, Beauchamp MS, DeYoe EA (2000) A comparison of visual and auditory motion processing in human cerebral cortex. *Cereb Cortex* 10:873–888
- Livingstone M, Hubel D (1988) Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science* 240:740–749
- Maeda F, Kanai R, Shimojo S (2004) Changing pitch induced visual motion illusion. *Curr Biol* 14:R990–R991
- Meyer GF, Wuerger SM (2001) Cross-modal integration of auditory and visual motion signals. *Neuroreport* 12:2557–2560
- Meyer GF, Wuerger SM, Röhrbein F, Zetzsche C (2005) Low-level integration of auditory and visual motion signals requires spatial co-localisation. *Exp Brain Res* 166:538–547
- Norman KA, Polyn SM, Detre GJ, Haxby JV (2006) Beyond mindreading: multi-voxel pattern analysis of fMRI data. *Trends Cogn Sci* 10 (9):424–430

- Pavani F, Macaluso E, Warren JD, Driver J, Griffiths TD (2002) A common cortical substrate activated by horizontal and vertical sound movement in the human brain. *Curr Biol* 12:1584-1590
- Perrott DR, Musicant AD (1977) Minimum auditory movement angle: binaural localization of moving sound sources. *J Acoust Soc Am* 62:1463-1466
- Poirier C, Collignon O, DeVolder AG, Renier L, Vanlierde A, Tranduy D, Scheiber C (2005) Specific activation of the V5 brain area by auditory motion processing: an fMRI study. *Brain Res Cogn Brain Res* 25:650-658
- Poirier C, Collignon O, Scheiber C, Renier L, Vanlierde A, Tranduy D, Veraart C, De Volder AG (2006) Auditory motion perception activates visual motion areas in early blind subjects. *Neuroimage* 31:279-285
- Rademacher J, Morosan P, Schormann T, Schleicher A, Werner C, Freund H-J, Zilles K (2001) Probabilistic mapping and volume measurement of human primary auditory cortex. *Neuroimage* 13:669-683
- Sadaghiani S, Maier JX, Noppeney U (2009) Natural, metaphoric, and linguistic auditory direction signals have distinct influences on visual motion processing. *J Neurosci* 29:6490-6499
- Saenz M, Lewis LB, Huth AG, Fine I, Koch C (2008) Visual motion area MT+/V5 responds to auditory motion in human sight-recovery subjects. *J Neurosci* 28:5141-5148
- Serences JT, Boynton GM (2007) Feature-based attentional modulations in the absence of direct visual stimulation. *Neuron* 55:301-312
- Sovijärvi ARA, Hyvärinen J (1974) Auditory cortical neurons in the cat sensitive to the direction of sound source movement. *Brain Res* 73:455-471
- Stumpf E, Toronchuk JM, Cynader MS (1992) Neurons in cat primary auditory cortex sensitive to correlates of auditory motion in three-dimensional space. *Exp Brain Res* 88:158-168
- Talairach J, Tournoux P (1988) Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system – an approach to cerebral imaging. New York, NY: Thieme Medical Publishers
- Warren JD, Zielinski BA, Green GGR, Rauschecker JP, Griffiths TD (2002) Perception of sound-source motion by the human brain. *Neuron* 34:139-148
- Wightman FL, Kistler DJ (1989) Headphone simulation of free-field listening. I: Stimulus synthesis. *J Acoust Soc Am* 85:858-867
- Wuerger SM, Hofbauer M, Meyer GF (2003) The integration of auditory and visual motion signals at threshold. *Percept Psychophys* 65:1188-1196
- Zimmer U, Macaluso E (2009) Interaural temporal and coherence cues jointly contribute to successful sound movement perception and activation of parietal cortex. *NeuroImage* 46:1200-1208

Auditory motion direction encoding

Zaitsev M, Hennig J, Speck O (2004) Point spread function mapping with parallel imaging techniques and high acceleration factors: fast, robust, and flexible method for echo-planar imaging distortion correction. *Magn Reson Med* 52:1156-1166



## **Chapter 3**

# Auditory, visual and audiovisual motion responses in hMT/V5+ and the planum temporale

Based on the publication:

Alink A, Singer W, Muckli L (2008) Capture of auditory motion by vision is represented by an activation shift from auditory to visual motion cortex. *Journal of Neuroscience* 28:2690

**Visual and auditory motion perception has previously been associated with enhanced neural responses in the human visual motion complex (hMT/V5+) and the planum temporale (PT) respectively. The aim of the functional magnetic resonance imaging (fMRI) study at hand is to test whether responses in these two areas are affected by audiovisual interactions. To this end we tested whether responses in hMT/V5+ and PT are affected by audiovisual motion coherency and the cross-modal dynamic capture (CDC) illusion – perceiving sounds as moving coherently with visual motion while in reality sounds and visual stimuli move in opposite directions. In addition, we tested whether hMT/V5+ and PT responded both to auditory and visual motion presented in isolation. Our results show that audiovisual motion coherence enhances responses in hMT/V5+ while not affecting activation levels of PT. Furthermore, we observed an effect of CDC illusion on activation levels in hMT/V5+ and PT. This illusion was found to enhance responses in hMT/V5+ while reducing responses in PT. We also observed that hMT/V5+ and PT exhibited responses both when subjects heard and saw moving stimuli presented in isolation. From these findings we conclude that cortical areas supporting visual and auditory motion perception process moving stimuli in an interactive manner.**

Object motion can be perceived both by seeing motion and hearing motion. For the human brain it is known that visual motion perception correlates with enhanced responses in the extra striate visual area hMT/V5+ (Beckers and Homber, 1992; Watson et al., 1993; Tootell et al., 1995) while auditory motion perception has been associated with enhanced responses in an area posterior to the primary auditory cortex called the planum temporale (PT, Baumgart et al., 1999; Griffiths and Warren, 2002; Warren et al., 2002).

These findings seem to suggest that auditory and visual motion perception are two independent cognitive modules (Fodor, 1983) which are supported by two independent cortical modules. This view, however, is incompatible with the finding that visual motion can alter the directionality in which one hears a sound move as illustrated by the cross-modal capture (CDC) illusion. This illusion consists of hearing a sound moving in the same direction as a visual stimulus although in reality sounds and visual stimuli are moving in opposite directions (Soto-Faraco et al., 2001; Soto-Faraco et al., 2002; Soto-Faraco et al., 2003; Soto-Faraco et al., 2004; Soto-Faraco et al., 2005; Sanabria et al., 2007). Furthermore, it has been shown that auditory motion presented in conjunction with visual motion can enhance the detection rate of visual motion when motion direction is coherent across senses (Wuerger et al., 2003) and that visual motion can cause an auditory motion after-effect (Kitagawa & Ichihara, 2002). These findings indicate

that audiovisual motion interactions exist on a perceptual level. Whether such interactions also exist on a neurophysiological level we assess in the current functional magnetic resonance imaging (fMRI) study by assessing whether responses in hMT/V5+ and PT are affected by both visual and auditory motion.

To this end we presented subjects with auditory apparent motion and visual motion stimuli that moved simultaneously in the same (coherent) or in the opposite (conflicting) direction. Subsequently, we tested whether blood oxygenation level depended (BOLD) responses in functionally defined regions of interest (ROIs) for hMT/V5+ and PT differentiate between coherent and conflicting audiovisual motion. Furthermore, we tested whether the CDC illusion affects responses in hMT/V5+ and PT and we assessed whether hMT/V5+ and PT responded both to visual and auditory motion presented in isolation.

## Materials and Methods

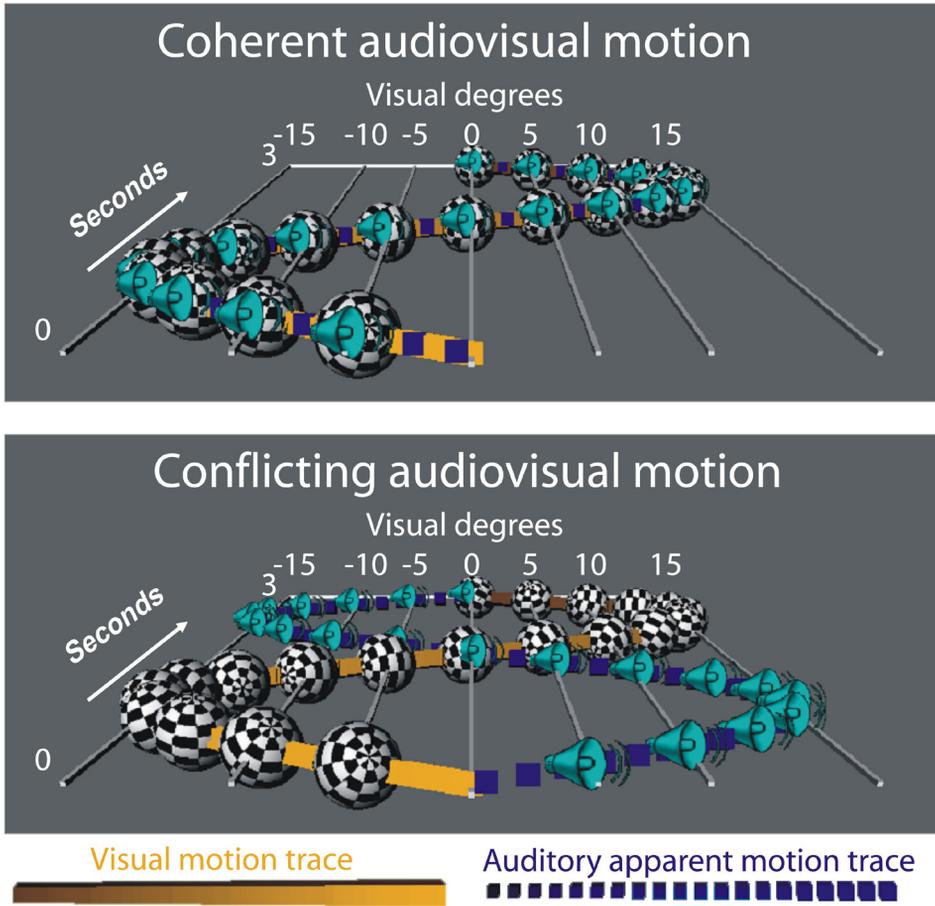
### Subjects

Ten healthy volunteers participated in the main fMRI study (age range 23-36 years, five females). Seven of these subjects (three females) participated in the localizer experiment. All subjects had normal hearing and normal or corrected-to-normal vision. All subjects gave their informed consent after being introduced to the experimental procedure in accordance with the declaration of Helsinki.

### Stimuli and task-Main experiment

Visual stimuli were presented using an MR compatible goggle system with two OLEDs (organic light-emitting diodes) displays (MR Vision 2000, Resonance Technology, Northridge, CA) and auditory stimulation was performed using an MR-compatible head phone system (Comander XG, Resonance Technology, Northridge, CA). The screen had a width of 30° and a height of 22.5° and the luminance of the grey background was 24.0 cd/m<sup>2</sup>.

During the audiovisual trials subjects were exposed to a moving sphere (1.5° radius) with a black and white checkerboard texture (luminance black = 1.2 cd/m<sup>2</sup>, luminance white = 43.9 cd/m<sup>2</sup>). At the same time they heard a stream of 20 base drum sounds (67.5-79.5 dB) with a duration of 100ms were presented with a 50 ms inter-stimulus interval. During these 20 periods of auditory stimulation the sphere pulsated (radius increase of .45°) which resulted in strong perceptual binding between visual and auditory stimuli. The sounds induced an apparent motion percept which was realized by a transformation of the sounds by a head related transfer function (HRTF) created by the MIT Media Lab Machine Listening Group (Gardner and Martin, 1994, <http://sound.media.mit.edu/KEMAR.html>) with a precision of 5° (implemented



**Figure 1**

A graphical representation of the audiovisual stimuli used in the experiment. Stimulation consisted of continuous visual motion and auditory apparent motion. The spheres represent the location of the visual stimuli and the speakers represent the perceived location of the auditory stimuli over time during a trial. Auditory stimuli were presented with a spatial resolution of five degrees. Stimuli locations are depicted for the 20 intervals during which an auditory stimulus was presented for 100ms (inter-stimulus-interval was 50ms). An example is shown of a coherent trial in which audiovisual stimuli move coherently (above) and of a conflicting trial (below) in which motion direction was opposite across senses. During the experiment these trials were presented together with an equal amount of their mirrored equivalents.

with Matlab). Both visual and auditory stimuli moved sinusoidally on the horizontal midline of the screen from the centre to 15° eccentricity of both sides and back to the centre within three seconds while subjects fixated at a white fixation cross 3.75° below the centre of the screen. During coherent audiovisual trials the auditory apparent motion had the same direction as the visual stimulus while the direction was opposite across senses during conflicting trials (see figure 1). The task during both audiovisual conditions was to report on the initial motion direction of the auditory stimulus. Subjects had to respond with their right hand before the stimuli disappeared. An index finger press indicated initial leftwards motion and a middle finger press initial rightward motion. After the stimuli disappeared the fixation cross turned green in the case of a correct response and red in the case of an incorrect response, no response or multiple responses. The fixation cross remained in this colour until the end of the trial which had a total length of four seconds.

During unimodal trials either visual or auditory stimuli were presented – the unimodal stimuli were essentially the same as those presented in the audiovisual trials. During the unimodal auditory trials subjects had to respond to the initial motion direction of the auditory stimulus in the same way as in the audiovisual trials. During the visual trials, however, subjects had to respond to the initial visual motion direction. Before the experiment began each subject completed 10 practice trials for each condition outside the scanner.

Within the scanner, subjects completed four runs that each contained 25 trials per condition. Within each run trials were intermixed with 25 fixation periods (four second duration each) which served to assess the baseline signal. Thus, during an fMRI session each subject was presented with a hundred trials per condition over all runs. We employed a rapid event related paradigm. In order to allow for correct deconvolution of the BOLD responses for each condition the history of trials was balanced in each run to control for an equal occurrence of the proceeding two trials. This ‘two-back’ balancing was achieved by drawing randomly a start sequence of three trials (triplet) and subsequently drawing a fitting triplet in which the first two trials match the last two trials of the preceding triplet. The drawing procedure was repeated until a valid solution was found that used all triplets (sequences were automatically generated using Matlab). The first triplet in each run was returned to the bowl and was drawn again at a random later position within the run. This resulted in a two-back balanced sequence per run that contained 25 trials for all five conditions (including fixation) plus the initial triplet that was disregarded in the data analysis. Furthermore, we ensured that the frequency of the initial motion direction of the visual and auditory motion stimuli was identical within and between conditions.

### **Stimuli and task-Localizer experiment**

For localization of the human visual motion complex (hMT/V5+) we used a standard block design mapping procedure that we have used and described previously (Muckli et al. 2002). In short, we used expanding low contrast random-dot-flowfield patterns covering a visual field of roughly  $20^\circ \times 30^\circ$  visual angle. Moving flowfields of random dot patterns were compared to static random dot patterns (RDP). Visual projection during the localizer experiment: the visual stimuli were back-projected onto a frosted screen attached to the end of the head coil which subjects could see through a mirror mounted to the inside of the head coil.

For the localization of motion sensitive auditory cortex the same HRTF function as in the main experiment was used to create a sound that appeared to rotate around the subjects head. For this we used the first 18 seconds of guitar music by Jeff Wahl (title: Groove) because it contained only small amplitude changes over time and we considered it a relaxing and enjoyable sound for the subjects to focus on during the measurement. The sound rotated on the horizontal plane (0 azimuth) with a speed of  $5^\circ$  per second and was presented using the same headphones as in the main experiment. Using the left and right channel of this rotating stimulus we created two stationary control stimuli. One of the stationary control stimuli consisted of the binaurally presented left channel and the other of the binaurally presented right channel of the rotating sound stimulus. Both of these stimuli gave rise to the perception of amplitude changes over time but did not induce any motion percept. For the analysis we pooled the two stationary sounds into one static control condition that contained the same stimulation across ears over all trials as the rotating stimulus. In total we presented the subjects with 30 samples of the moving sound and 30 samples of the static control sounds (15 of each type). All samples had lasted 18 seconds. The order of the trials was randomized and sound samples were separated from one another by an 18 second fixation period. During the entire experiment subjects were instructed to fixate at a central white fixation cross on a black screen.

### **fMRI procedure-Main experiment**

Functional and anatomical MRI data were acquired with a 3T-MRI system (Siemens Allegra; Siemens, Erlangen, Germany) using a 4-channel head coil. For each subject we obtained 516 volumes containing 20 slices covering the entire brain during each of the four functional scans using a gradient echo– echo planar imaging (EPI) sequence (repetition time (TR), 1000 ms; echo time (TE), 25 ms; flip angle,  $70^\circ$ ; voxel size,  $3.4 \times 3.4 \times 5.0$  mm; Field Of View (FOV), 220mm; gap-thickness, 0.7 mm). We corrected for spatial distortions in the EPI images using a

Point Spread Function (PSF) (Zaitsev et al., 2004). We also obtained a T1-weighted anatomical scan for each of the subjects using a Siemens MPRAGE sequence (1 X 1 X 2mm).

fMRI procedure- Localizer experiment. For this experiment we used the standard Siemens 3T head coil. Functional data for visual motion mapping was acquired in one functional scan containing 216 volumes. For the auditory motion mapping we measured three times 720 volumes in three scans. The same EPI sequence was used for visual and auditory motion mapping (TR=1000 ms, TE, 30 ms; flip angle, 77°; voxel size, 3.4 X 3.4 X 3.5 mm; FOV, 220mm; gap-thickness, 0.35 mm) and PSF was applied to correct for spatial distortions. Slices were orientated parallel to the planum temporale covering the lower part of the parietal, the lower part of the frontal, the upper part of the temporal and the entire occipital lobe. We also obtained a T1-weighted anatomical scan for each of the subjects using a Siemens MPRAGE sequence (1 X 1 X 1mm).

### **Data analysis-Main experiment**

Cross-modal dynamic capture trials were defined as conflicting trials in which the opposite of true motion direction was reported. All other trial types (coherent, visual and auditory trials) with false responses were excluded from further analysis. Furthermore, the first three trials in each scan were excluded to preclude T1 saturation effects and to ensure a balanced 2-back history across conditions. fMRI data was analysed using the BrainVoyager QX software package (Brain Innovation, Maastricht, The Netherlands). Data were pre-processed using the default settings of BrainVoyager QX. After alignment with the anatomical reference scan we spatially smoothed the functional data using a Gaussian kernel with a full width at half maximum (FWHM) of 8mm. All individual datasets were transformed into Talairach space (Talairach and Tournoux, 1988). For each subject a General Linear Model (GLM) was computed using a deconvolution design (Glover, 1999).

### **Data analysis-Localizer experiment and ROI analysis**

All data were preprocessed and normalized as in the main experiment although the data was not spatially smoothed. GLMs were calculated for auditory and visual motion mapping using a single factor design that used regression to the hemodynamic response function (Boynton et al., 1996). T-value maps were computed for the contrast moving auditory stimulus > stationary auditory stimulus for localization of PT and flowfield RDP > static RDP for the localization of hMT/V5+. Regions of interests (ROIs) for hMT/V5+ were identified individually for the seven subjects that participated in the localizer experiment. These consisted of the 1000 +/- 44 mm<sup>3</sup> that was most significantly activated for

the flowfield RDP > static RDP contrast near the posterior part of the inferior temporal sulcus. For the three subjects that could not participate in the localizer experiment we defined hMT/V5+ ROIS as the 1000 +/- 44 mm<sup>3</sup> close to the posterior part of the inferior temporal sulcus that was most significantly activated by the visual motion condition in the main experiment.

ROIs for PT were defined on a group level of all seven subjects in the localizer experiment. We preferred group level ROIs to individual ROIs because PT mapping was not robust enough at an individual level but was sound at a group level. Note that although this ROI was defined based on the data of seven subjects we used it for the ROI analysis over all ten subjects. Defining these regions of interest allowed us to generate event related time courses for hMT/V5+ and PT for the data in the main experiment using deconvolution.

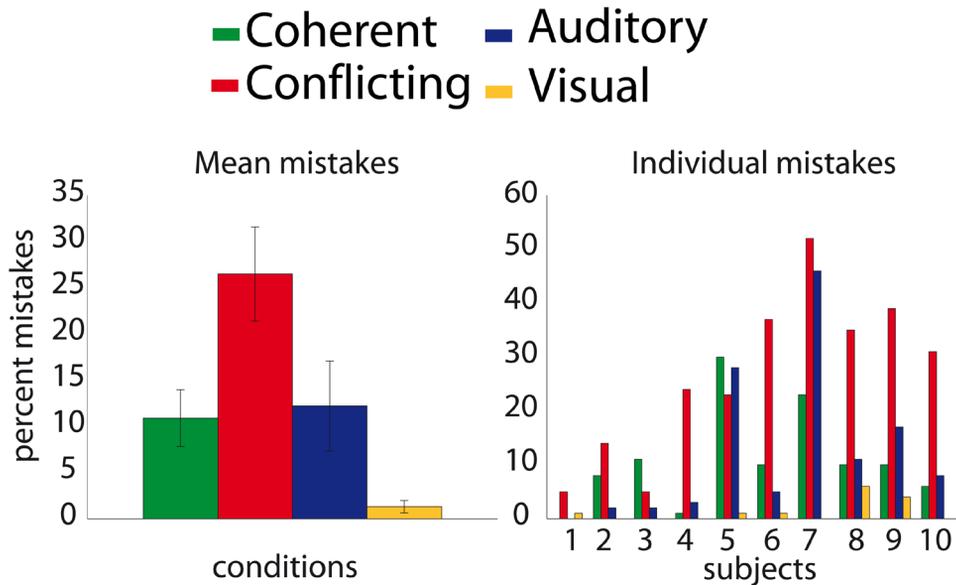
In order to test whether audiovisual integration took place in these motion areas we tested the following contrasts for significance: coherent > conflicting, conflicting > coherent, cross-modal dynamic capture > conflicting and conflicting > cross-modal dynamic capture. Results indicated that responses in hMT/V5+ for audiovisual stimulation were actually lower than those for visual stimulation alone. Furthermore, both motion areas turned out to respond to both unimodal auditory as well as unimodal visual stimulation. Therefore, we also tested the following contrasts for significance (visual > average (coherent, conflicting), visual > 0 and auditory > 0. All contrasts were calculated over the onset and the peak of the BOLD responses which lasted from two to seven seconds after stimulus onset. This corresponded to a contrast over the data points three to eight as shown in figure 3 (one data point/volume was recorded each second the first being recorded at stimulus onset).

In order to assess how consistent multisensory effects in the auditory and visual motion areas were over subjects we plotted the effect sizes for the contrasts for CDC minus conflicting and coherent minus conflicting for each individual separately. In order to make sure that multisensory effects on a group level were not driven by a single subject we tested whether there were extreme outliers in the individual data by testing for the presence of subjects with effect sizes that were beyond three times the inter-quartile range.

## Results

### Behavioural data

During the fMRI experiment Cross-modal dynamic capture occurred in 26.5% of the conflicting trials. This was significantly higher ( $p < 0.0005$ ) than the misclassifications in the congruent trials (10.9%) and the auditory trials (12.2%)



**Figure 2**

Left, the group average of the behavioural data showing the amount of misclassification during the different conditions (including standard errors across subjects). During all types of trials in which auditory stimuli were presented subjects reported the direction of auditory motion while during the visual trials subjects responded to the direction if visual motion. Right, the misclassification rate for each single subject.

which did not significantly differ from each other. The visual trials showed significantly lower misclassification than any other type of trial (1.3%,  $p < 0.0007$ ). Individual mistake rates varied widely over subjects. However, higher mistake rates for the conflicting condition turned out to be consistent with eight out of ten subjects exhibiting this trend. See Figure 2 for more details.

## fMRI data

### Localization of hMT/V5+

In our study we focussed on early motion specific areas in the visual and auditory cortex. We therefore used standard mapping procedures to define motion specific ROIs in visual and auditory motion cortex. For the visual modality responses to moving low contrast RDPs were compared to responses to static low contrast RDPs in seven out of the ten subjects. We localized the left and right human visual motion complex (hMT/V5+) for each of these subjects by

selecting 1000 +/- 44 mm<sup>3</sup> of brain volume near the posterior part of the inferior temporal sulcus that was most motion sensitive. On the group average the left and right hMT/V5+ was located at the Tal coordinates: left hMT/V5+, x = -44.1 (SE 1.5) y = -67.0 (SE .9) z = 0.7 (SE 1.6), right hMT/V5+, x = 45.3 (SE 2.1) y = -64.3 (SE 3.0) z = -0.3 (SE 1.9). For the three subjects that could not participate in the localizer experiment we used the activation for the visual motion condition in the main experiment to define a ROI for hMT/V5+. This area was always an island of activation in the vicinity of the posterior part of the inferior temporal sulcus. For this group the ROI for the left hMT/V5+ was located on average at x = -49,3 (SE 3,3) y = -65,0 (SE 3,2) z = 2,3 (SE 1,5) and the location of the right hMT/V5+ ROI was x = 42,3 (SE 2,3) y = -62,0 (SE ,7) z = 1,7 (SE 2,2).

### **Localization of PT**

For the auditory modality, responses to moving sounds, that appeared to rotate around the subjects head, were compared to static control conditions. We defined auditory motion sensitive areas based on a group analysis. An area covering the planum temporale in both hemispheres was found to respond more strongly to moving sounds than to static control sounds. The location of this area (Tal coordinates: left PT- x = -52, y = -33, z = 15; right PT- x = 53, y = -26 z = 17) is in agreement with the motion sensitive area found in previous fMRI studies (Baumgart et al., 1999; Griffiths and Warren, 2002; Warren et al., 2002). We refer to this area as PT.

### **ROI-based hMT/V5+ analysis**

One of the main questions of our study was whether motion sensitive areas, as defined by the ROI mapping procedures, are affected by the cross-modal dynamic capture (CDC) illusion. Therefore, we compared the ROI-activity for hMT/V5+ in trials in which the CDC illusion was present with those – physically identical – trials in which no such illusion was perceived. Both the left and right hMT/V5+ responded more strongly to conditions in which the illusion occurred (left:  $p < 0.0004$ , right:  $p < .04$  resp.). In individual subjects responses were consistently higher for CDC trials in left hMT/V5+ while this difference was less consistent in the right hMT/V5+ (see figure 3b).

Furthermore, we were interested in whether hMT/V5+ responses differentiate between audiovisual motion with a coherent and a conflicting direction of motion across senses. Left and right hMT/V5+ both show stronger responses to coherent audio-visual stimulation as compared to conflicting stimulation (left hMT/V5+,  $p < 0.00003$ ; right hMT/V5+,  $p < 0.002$ ). This effect of motion coherency was highly consistent across subjects with all subjects having a higher BOLD response for the coherent condition in left hMT/V5+ and nine out of ten in the

right hMT/V5+. This is a strong indication that this visual motion area is affected by the direction of auditory motion relative to visual motion. Testing for outliers regarding the individual CDC and coherency effect sizes showed that there were no extreme outliers that drive the observed effects on a group level.

If hMT/V5+ is indeed involved in auditory motion processing this area can also be expected to respond to auditory motion presented in isolation. Concordantly, we found such a unimodal response in left and right hMT/V5+ to auditory motion (left hMT/V5+,  $p < 0.00001$ ; right hMT/V5+,  $p < 0.00001$ ). The strongest response, however, occurred for pure visual stimulation (visual > audiovisual stimulation: left hMT/V5+,  $p < 0.00001$ ; right hMT/V5+,  $p < 0.00001$ ). This is most likely due to the fact that this was the only condition during which subjects attended visual motion direction.

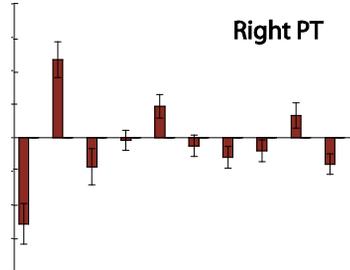
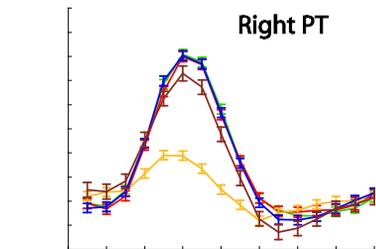
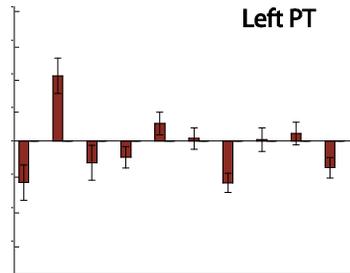
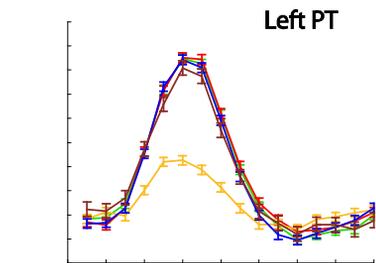
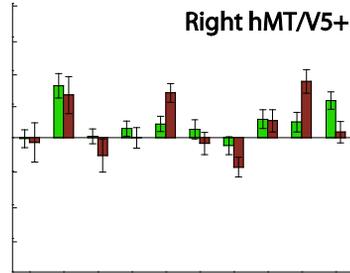
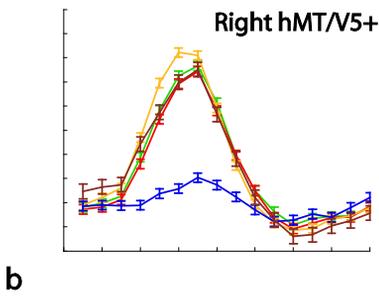
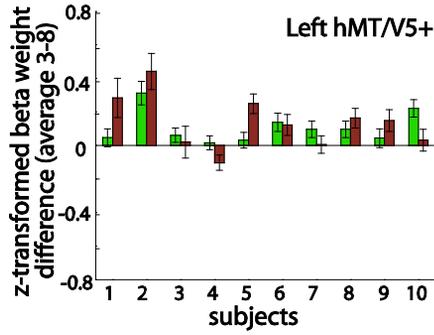
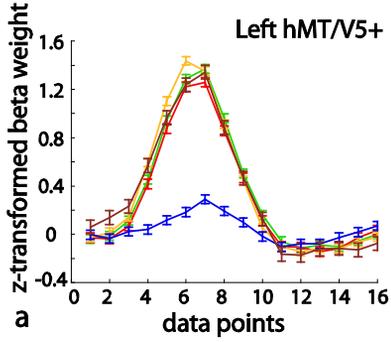
### **ROI-based PT analysis**

Like hMT/V5+, also PT was affected by cross-modal dynamic capture. Hence, both the left and the right PT showed a decrease in activation when cross-modal dynamic capture took place (left PT,  $p < 0.005$ ; right PT,  $p < 0.006$ ). Thus, when the visual motion percept dominated the auditory motion percept there was a bilateral reduction in activation in auditory motion cortices. In the individual data this effect was expressed in the majority of the subjects in the right PT while being less consistent within the left PT. In contrast to hMT/V5+ responses, PT responses did not differentiate coherent and conflicting audiovisual motion trials.

As for hMT/V5+, we tested whether there was a different response in PT to audiovisual motion as compared to the preferred unimodal motion stimuli which is auditory motion for PT. We did not find evidence for such a difference. Hence, in PT responses were close to identical for auditory motion, coherent audiovisual motion and conflicting audiovisual motion. Surprisingly, both left and right PT responded robustly to the unimodal visual motion stimulus (left PT,  $p < 0.00001$ ; right PT,  $p < 0.00001$ ).

### **Discussion**

In this fMRI study we tested whether responses in hMT/V5+ and PT to moving stimuli are subject to audiovisual interactions. To this end we investigated whether audiovisual motion coherency and the cross-modal dynamic capture (CDC) illusion affect responses in these areas. Furthermore, we have assessed whether hMT/V5+ and PT respond both to visual and auditory motion presented in isolation.



**Figure 3**

Left, the grand mean event related responses generated using deconvolution for the regions of interest (ROIs) hMT/V5+ and the Auditory motion complex (AMT) as defined in the localizer experiment. The ROI for the hMT/V5+ was defined on an individual level and the PT on a group level. Event related responses are shown for Coherent audiovisual trials (green), Conflicting audiovisual trials with correct responses (red), and conflicting audiovisual trials with wrong responses i.e. inducing the illusion of cross-modal dynamic capture (brown), Auditory (blue) and Visual (yellow) trials. Right, individual differences of the peak of the BOLD response (mean of data points 3-8) between coherent and conflicting trials (green) and the difference between cross-modal dynamic capture and conflicting trials (brown). Data points were recorded each second the first starting at stimulus onset.

One of the main findings of this study is that moving sounds affect responses in the visual motion area hMT/V5+. hMT/V5+ is here shown to be sensitive to the direction of auditory motion relative to the direction of visual motion as it exhibits an enhanced response for coherent versus conflicting audiovisual motion. Such a response enhancement might explain the fact that auditory motion can aid visual motion direction sensitivity (Wuerger et al., 2003). Moreover, in our study we observed that hMT/V5+ responses are elevated when the perceived direction of auditory motion changes without physical changes of the auditory stimulus in the CDC trials. On top of this, hMT/V5+ was also found to be activated by auditory motion stimuli presented in isolation which is in line with the findings of previous human neuroimaging studies (Porrier et al., 2005; Porrier et al., 2006). From this we conclude that hMT/V5+ responses are affected by the presence of auditory motion as well as the physical and the perceived direction of auditory motion

PT, on the other hand was found to exhibit a reduced response during CDC trials while not being sensitive to physical coherence of the audiovisual motion stimuli. Furthermore, this area was found to respond to visual motion presented in isolation. The finding that responses in PT are reduced when the CDC illusion appears compatible with the decrease in  $d'$  for auditory motion direction perception observed by Sanabria and colleagues (2007) assuming that lower responses in PT worsen the quality of auditory motion perception. When such a decrease takes place in PT, hMT/V5+ responses are here shown to increase which could explain why during the CDC illusion subjects tend to report that sounds move in the same direction as visual stimuli instead of exhibiting guessing behaviour (Soto-Faraco et al., 2004).

There is some disagreement between the results presented here and the results of the study by Baumann and Greenlee (2008) which reported clusters near the

planum temporale exhibiting enhanced responses for coherent versus incoherent motion while observing no effect of motion coherency in hMT/V5+. We here list some differences in methodology and analysis which most likely are responsible for this discrepancy. First of all, during the experiment of Baumann and Greenlee (2008) subjects were instructed to report the direction of visual motion while our subjects reported the direction of auditory motion. Secondly, we investigated responses in hMT/V5+ and PT using a region of interest based while Baumann and Greenlee (2008) performed a full brain analysis. Finally, it should be mentioned that the areas reported as responding stronger to coherent audiovisual motion in the study of Baumann and Greenlee were areas in which BOLD responses exceeded the statistical threshold ( $p < .05$ , corrected for multiple comparisons) for coherent but not for conflicting audiovisual stimulation. So, in contrast to our study they did not perform a direct statistical analysis of BOLD response differences which is the more conservative type of analysis.

To summarize, we here provide evidence for motion induced responses in the visual motion area hMT/V5+ and PT being subject to audiovisual interactions and that these areas respond both to visual and auditory motion presented in isolation. Therefore, we conclude that visual and auditory motion perception are processes that interact with each other both on a perceptual (Soto-Faraco et al., 2001; Soto-Faraco et al., 2002; Kitagawa & Ichihara, 2002; Wuerger et al., 2003; Soto-Faraco et al., 2003; Soto-Faraco et al., 2004; Soto-Faraco et al., 2005; Sanabria et al., 2007) and a neurophysiological level.

## References

- Baumann O, Greenlee MW (1999) Neural correlates of coherent auditory motion perception. *Cerebral Cortex* 17:1433
- Baumgart F, Gaschler-Markefski B, Woldorff MG, Heinze HJ, Scheich H (1999) A movement-sensitive area in auditory cortex. *Nature* 400:724–726
- Beckers G, Homberg V (1992) Cerebral visual motion blindness: Transitory akinetopsia induced by transcranial magnetic stimulation of human area V5. *Proceedings: Biological Sciences*:173–178
- Boynton GM, Engel SA, Glover GH, Heeger DJ (1996) Linear systems analysis of functional magnetic resonance imaging in human V1. *Journal of Neuroscience* 16:4207
- Fodor JA (1983) *The modularity of mind*. MIT press Cambridge, MA.
- Griffiths TD, Warren JD (2002) The planum temporale as a computational hub. *Trends in Neurosciences* 25:348-353
- Kitagawa N, Ichihara S (2002) Hearing visual motion in depth. *Nature* 416:172–174

- Muckli L, Kriegeskorte N, Lanfermann H, Zanella FE, Singer W, Goebel R (2002) Apparent motion: event-related functional magnetic resonance imaging of perceptual switches and states. *Journal of Neuroscience*: Vol. 22 RC219
- Poirier C, Collignon O, DeVolder AG, Renier L, Vanlierde A, Tranduy D, Scheiber C (2005) Specific activation of the V5 brain area by auditory motion processing: an fMRI study. *Cognitive Brain Research* 25:650–658
- Poirier C, Collignon O, Scheiber C, Renier L, Vanlierde A, Tranduy D, Veraart C, De Volder AG (2006) Auditory motion perception activates visual motion areas in early blind subjects. *Neuroimage* 31:279–285
- Sanabria D, Spence C, Soto-Faraco S (2007) Perceptual and decisional contributions to audiovisual interactions in the perception of apparent motion: a signal detection study. *Cognition* 102:299–310
- Soto-Faraco S, Kingstone A, Spence C (2003) Multisensory contributions to the perception of motion. *Neuropsychologia* 41:1847–1862
- Soto-Faraco S, Lyons J, Gazzaniga M, Spence C, Kingstone A (2002) The ventriloquist in motion: Illusory capture of dynamic information across sensory modalities. *Cognitive brain research* 14:139–146
- Soto-Faraco S, Spence C, Kingstone A (2004) Cross-modal dynamic capture: congruency effects in the perception of motion across sensory modalities. *Journal of Experimental Psychology Human Perception and Performance* 30:330–345
- Soto-Faraco S, Spence C, Kingstone A (2005) Assessing automaticity in the audiovisual integration of motion. *Acta psychologica* 118:71–92
- Talairach J, Tournoux P (1988) Co-planar stereotaxic atlas of the human brain. Thieme New York.
- Tootell RB, Reppas JB, Kwong KK, Malach R, Born RT, Brady TJ, Rosen BR, Belliveau JW (1995) Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *Journal of Neuroscience* 15:3215
- Warren JD, Zielinski BA, Green GG, Rauschecker JP, Griffiths TD (2002) Perception of sound-source motion by the human brain. *Neuron* 34:139–148
- Watson JDG, Myers R, Frackowiak RSJ, Hajnal JV, Woods RP, Mazziotta JC, Shipp S, Zeki S (1993) Area V5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cerebral Cortex* 3:79
- Wuerger SM, Hofbauer M, Meyer GF (2003) The integration of auditory and visual motion signals at threshold. *Perception and Psychophysics* 65:1188–1196
- Zaitsev M, Hennig J, Speck O (2004) Point spread function mapping with parallel imaging techniques and high acceleration factors: fast, robust, and

## Chapter 3

flexible method for echo-planar imaging distortion correction. *Magnetic Resonance in Medicine* 52:1156-1

## **Chapter 4**

# Auditory motion capturing ambiguous visual motion

Based on the manuscript:

Auditory motion capturing ambiguous visual motion. Arjen Alink, Felix Euler, Elena Galeano, Wolf Singer and Axel Kohler.

In preparation for a submission to Perception

**In this study we assessed whether the direction of auditory motion can affect the directionality of perceived visual motion. In contrast to previous studies we employed visual stimuli with a fully ambiguous motion direction. We suspected that the use of such stimuli would enhance the cross-modal influence of audition on visual motion perception as multisensory integration has been proposed to resemble Bayesian integration. In a psychophysical experiment we determined for each subject a visual stimulus configuration that was equally likely to induce the percept of left- and rightwards visual motion. Subsequently, it was tested whether such a bistable visual stimulus was perceived more often as moving leftwards when presented in conjunction with a leftwards moving sound and vice versa when the stimulus was presented in conjunction with a rightwards moving sound. Indeed such an effect of auditory motion direction on visual motion perception was observed which indicates that auditory and visual motion perception are two interdependent processes.**

Each of our senses provides us with qualitatively different impressions about the objects and events that surround us. Frequently, events elicit impressions in more than one sensory modality. For example, the event of someone walking towards you can be observed by seeing the person moving towards you as well by hearing the sound of footsteps getting louder. An interesting question that arises here is how an observer combines visual and auditory signals into a unified multisensory percept. One model on how such multisensory integration takes place is the modality appropriateness model by Welch and Warren (1980). In this model it is proposed that sensory modalities are differentially suited to detect stimulus features and that multisensory perception of a feature is dominated by the most “appropriate” sense. Therefore, spatial perception is dominated by vision because it has the highest spatial acuity (Howard and Templeton, 1966; Welch and Warren, 1980) while temporal perception is dominated by audition as this sense has the highest temporal acuity (Gebhard and Mowbray, 1959; Shams et al., 2002; Morein-Zamir et al., 2003; Spence and Squire, 2003). Recently, this model has been refined as it was shown that whether a sensory modality will dominate the multisensory percept of a stimulus feature depends on how reliably its sensory input provides information about this feature relative to the other sensory modalities (Wada et al, 2003; Alais and Burr, 2004). Such a dynamic stimulus dependent integration of multisensory signals has inspired the idea that multisensory perception resembles a process of Bayesian integration (Pouget et al, 2003; Battaglia et al., 2003; Knill and Pouget, 2004).

In the psychophysical study at hand, our main aim is to find out whether the direction of auditory motion can affect the perceived direction of visual motion. Previous studies have shown that visual stimuli tend to dominate the perceived direction of auditory motion (Soto-Faraco et al., 2002; Soto-Faraco et al., 2003; Soto-Faraco et al., 2004; Soto-Faraco et al., 2005; Sanabria et al., 2007). However, these studies did not observe an effect of auditory motion on visual motion perception. We suspect that previous studies did not observe such an effect because they employed visual stimuli that conveyed information about the direction of motion more reliably than simultaneously presented auditory stimuli. If audiovisual motion perception resembles Bayesian integration this should strongly reduce the impact of audition on vision. In this psychophysical study we overcame this shortcoming by employing visual stimuli with a fully ambiguous motion direction which should maximize the effect of auditory stimuli on visual motion perception. To achieve this we determined for each subject a visual stimulus configuration that was equally likely to induce the percept of leftwards and rightwards visual motion. In order to test whether the direction of auditory motion can affect the perceived direction of visual motion we assessed whether such a bistable visual stimulus was perceived more often as moving rightwards when presented simultaneously with a sound moving rightwards and vice versa when presented simultaneously with a leftwards moving sound.

## Methods

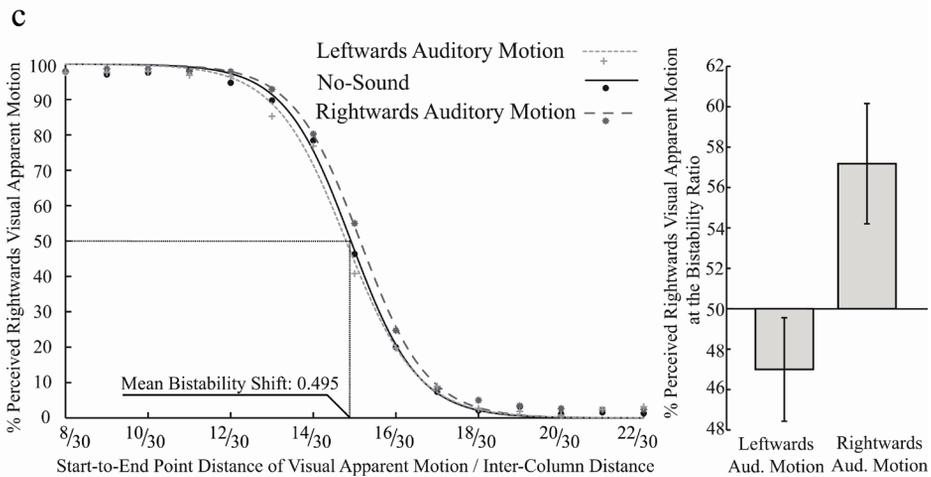
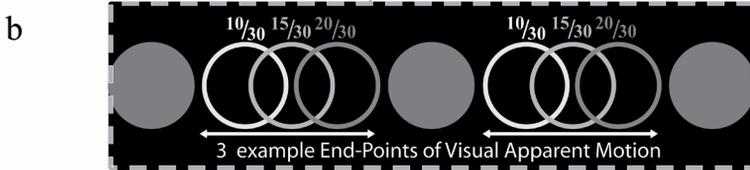
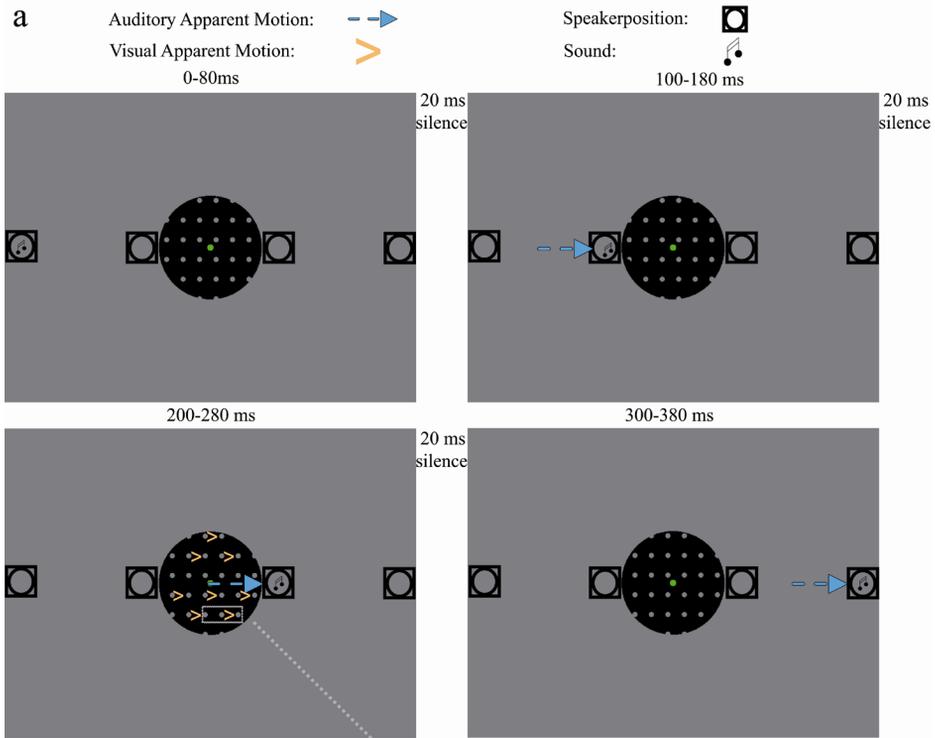
### Subjects

Nineteen healthy volunteers participated in the experiment (age range, 20–31 years; 11 females). All subjects had normal hearing and normal or corrected-to-normal vision and gave their informed consent after being introduced to the experimental procedure in accordance with the Declaration of Helsinki.

### Stimuli and procedure

Subjects were seated in a darkened room 2.2 meters away from a white wall on which a beamer (Acer, PD322) projected an image at 60 Hz with a width of 44.5° and a height of 33.4°. The background colour of this image was gray (luminance = 6.2 cd/m<sup>2</sup>) and the image contained a black aperture (luminance = 0.8 cd/m<sup>2</sup>) in the centre of the image, which had a radius of 5.6°. On the horizontal midline of the projected image, four black speakers (Trust 5.1 Surround Speaker Set SP-6210) were attached to the wall. The far left speaker was located 20.7° left from

Chapter 4



**Figure 1**

*a* A schematic depiction of the stimuli employed in this study illustrating a rightwards auditory-motion trial during which the subject perceived rightwards visual apparent motion. *b* An illustration of three possible endpoints of visual apparent motion whose position is expressed as the distance from the nearest left start point divided by the inter-column distance. The 10/30 endpoints most likely induce rightwards visual apparent motion and the 20/30 endpoints most likely induce leftwards visual apparent motion while the 15/30 endpoints are as likely to induce left- and rightwards visual apparent motion. *c* left The grand mean of percentage perceived rightwards visual apparent motion for each visual apparent motion configuration plotted separately for each auditory condition and their respective fitted Boltzmann functions. right Grand mean of the estimated percent perceived rightwards visual apparent motion for the left- and rightwards auditory-motion conditions for bistable visual apparent motion. The error bars indicate the standard error of mean (SEM) across subjects.

the image centre and the mid-left speaker  $6.9^\circ$  left from the image centre. The far right speaker was located  $20.7^\circ$  right from the image centre and the mid-left speaker  $6.9^\circ$  right from the image centre. The inter-speaker distances were  $13.8^\circ$  between all adjacent speakers. All images and sounds were generated using Presentation software (version 12.2, Neurobehavioral Systems, Inc.).

During a trial we presented visual apparent-motion stimuli in the aperture that were similar to those employed by Williams and colleagues (2003). Hence, in the aperture we presented visual stimuli consisting of columns of gray discs (radius =  $0.3^\circ$ , luminance =  $6.2 \text{ cd/m}^2$ ) with a fixed inter-column distance of  $1.8^\circ$  and a fixed inter-row distance of  $2.1^\circ$ . A visual apparent-motion stimulus consisted of two images. The first image was presented for 200 ms with a randomized horizontal position of the disc columns and was followed by a second image that had the disc columns shifted to the right. The degree of the rightward shift varied from 8/30 of the inter-column distance to 22/30 of the inter-column distance in steps of 1/30. This resulted in fifteen types of visual apparent motion. Based on the nearest-neighbour principle for the binding occurring during visual apparent motion (Shechter et al., 1988; Dawson, 1991), we expected subjects to consistently perceive rightwards visual apparent motion for an 8/30 shift while consistently perceiving leftwards visual apparent motion for a 22/30 shift. A 15/30 shift was expected to equally often induce a left- and rightwards visual apparent-motion percept (see figure 1b).

Trials contained one visual apparent-motion stimulus and at the end of each trial subjects indicated in a two-alternative forced choice regime which visual motion direction they had seen. Responses were made using a conventional PC mouse and subjects pressed the left and right mouse button when they perceived

left- and rightwards visual motion respectively. During the response period the gray discs disappeared and the next trial would start 500ms after a response was made. Subjects were instructed to fixate the green disc that persisted throughout the entire session at the centre of the image (radius =  $0.4^\circ$ , luminance =  $11.3 \text{ cd/m}^2$ ).

Visual apparent motion was presented under three auditory conditions. During the no-sound condition visual apparent motion was presented without auditory stimulation. During the left- and rightwards auditory-motion conditions visual apparent motion was presented in conjunction with auditory apparent-motion stimuli. Leftwards auditory-motion trials contained white noise bursts (82-85 dB) which were presented for 80ms with an inter-stimulus interval of 20 ms at the far-right, mid-right, mid-left and far-left speaker consecutively and visa versa for rightwards auditory motion trials. Critically, the third sound was presented simultaneously with the onset of the second visual apparent motion component which leads to an overlap of perceived visual and auditory motion in space and time (see figure 1a).

In total, each subject was presented ten times with each of the fifteen visual apparent-motion stimuli for each trial type in each of two sessions (900 trials overall). The order of trials was fully randomized. On average each session took around 15 minutes and subjects were allowed to take a short break in between the two sessions. Subjects were told before the experiment that sounds would be presented during the experiment but that they were irrelevant for the visual motion task.

## Analysis

For each subject, we calculated the percentage of rightwards visual apparent-motion reports for each of the rightwards shifts separately for the three main conditions. We fitted a sigmoidal psychometric function (Boltzmann function) to these data for each subject and condition separately (Draper and Smith, 1981). The tails of the Boltzmann function were fixed to 0 and 100 percent, respectively. An exemplary subject's data and its fitted curves are shown in the supplementary figure. The fitted Boltzmann function for the no-sound condition was used to estimate the shift that induces left- and rightwards visual apparent motion with an equal probability without auditory influences. This was realized by determining the shift at which the no-sound curve intersected with a 50% level of rightwards visual apparent motion perception (V50 of the Boltzmann function). This shift we refer to as the bistability shift and we determined this shift for each subject separately.

The main aim of this study is to test whether the direction of auditory motion affects the ratio of perceived left- and rightwards visual apparent motion when visual apparent motion is bistable. To this end, we estimated the proportion at which subjects perceived rightwards visual motion for the two main auditory conditions when a visual stimulus is presented using the individually defined bistability shift. This was realized by fitting two Boltzmann functions to the data from left- and rightwards auditory motion trials for each subject separately and comparing the amplitude of these curves for the bistability shift (see supplementary figure for an example). If our hypothesis is correct then these values should be higher than 50% for the rightwards auditory motion condition and lower than 50% for the leftwards auditory motion condition. We assessed the effect of auditory motion direction on the percentage reported rightwards visual apparent motion during bistable visual stimulation in a repeated measures analysis of variance (ANOVA) and subsequent within-subject t-tests. We furthermore assessed, in another ANOVA the effect of auditory motion direction on the threshold and slope of the fitted Boltzmann functions.

## Results

The average rightwards shift at which the visual apparent motion stimulus was estimated to be bistable was .495 (SD = .0126) of the inter-column distance. For an overview of the average percentage reported rightwards visual motion for all conditions and the full range of rightward shifts see figure 1c.

Our ANOVA shows that the direction of auditory motion significantly affects the percentage of trials in which a bistable visual stimulus is seen as moving rightwards ( $F = 7.8$ ,  $p < .003$  Greenhouse–Geisser). Subjects were estimated to perceive the bistable visual stimulus as moving rightwards in 57.2% (SD = 13.0%) of the rightwards auditory-motion trials while perceiving the stimulus as moving rightwards only in 47.0% (SD = 11.2%) of the leftwards auditory motion trials (see figure 1c). Paired t-tests show that bistable visual apparent motion induces a rightwards visual motion percept more often during rightwards auditory motion trials as compared to leftwards auditory motion trials ( $t = 4.3$ ,  $p < .001$  two-sided) and no-sound trials ( $t = 2.4$ ,  $p < .03$  two-sided). The difference between the leftwards auditory motion condition and the no-sound condition, however, did not reach significance ( $t = 1.2$ ,  $p = .26$ ).

Auditory stimulation was found to have an effect on the threshold of the fitted Boltzmann function ( $F = 7.6$ ,  $p < .004$  Greenhouse–Geisser) while not affecting the Boltzmann function's slope ( $F = .431$ ,  $p = .65$  Greenhouse–Geisser). The V50 of the Boltzmann function was .492 (SD = .0155), .495 (SD = .0126) and .505 (SD = .0146) of the inter-column distance for the leftwards auditory-motion condition, the no-

sound condition and the rightwards auditory-motion condition, respectively. The differences in threshold between the left- and rightwards auditory-motion condition and between the rightwards auditory-motion condition and the no-sound condition reached significance (respectively:  $t=3.5$ ,  $p<.004$  two sided;  $t=2.5$ ,  $p<.03$  two sided) while the difference between the leftwards auditory motion condition and the no-sound condition was not significant ( $t = 1.2$ ,  $p = .24$  two sided). These findings indicate that the effects of auditory motion found in our main analysis are due to a shift of the sigmoidal psychometric function without a change of slope.

## Discussion

To the best of our knowledge, this study is the first to show that the direction of auditory motion can bias the perceived direction of visual motion. Our results show that the presence of rightwards auditory motion causes visual stimuli to be perceived more often as moving rightwards when the visual stimulus' direction is ambiguous. Furthermore, in this study we observed an increased likelihood of perceiving leftwards visual motion when comparing trials containing leftwards and rightwards auditory motion. However, the ratio of left- and rightwards visual apparent motion perception was not found to differ between trials during which leftwards auditory motion was presented and trials without auditory stimulation.

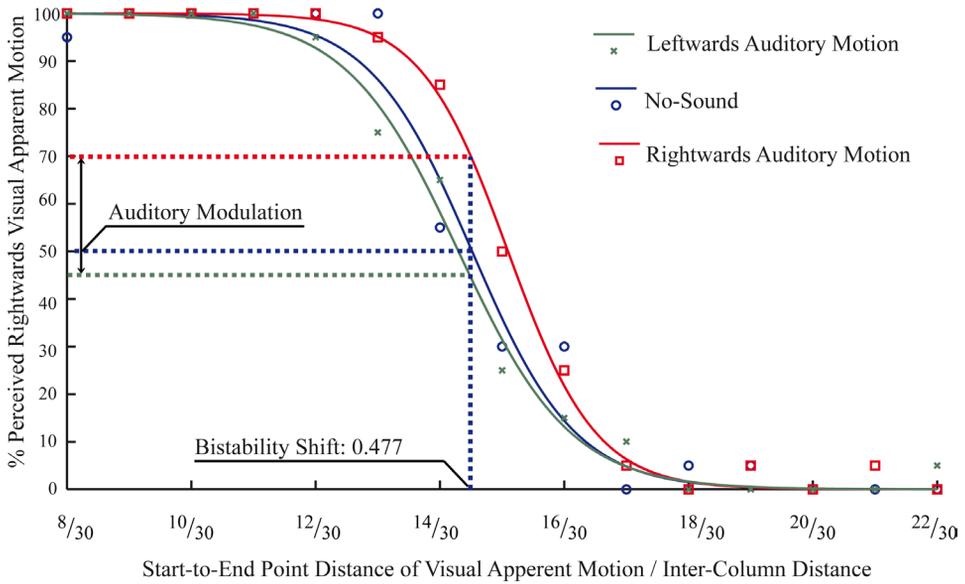
Our findings invalidate the claim made by previous studies that visual motion affects auditory motion perception but not the other way around (Soto-Faraco et al., 2002; Soto-Faraco et al., 2003; Soto-Faraco et al., 2004; Soto-Faraco et al., 2005; Sanabria et al., 2007). The lack of evidence in previous studies for an effect of auditory motion on visual motion perception is most likely due to the fact that those studies, in contrast to the present study, employed visual stimuli that more reliably conveyed information about the direction of motion than the auditory stimuli. Hence, if audiovisual integration resembles Bayesian integration (Pouget et al, 2003; Battaglia et al., 2003; Knill and Pouget, 2004) then bistable visual motion stimuli can be expected to be more susceptible to auditory influences than visual stimuli which motion direction is unmistakable.

Recently, Freeman and Driver (2008) demonstrated that the salience of visual apparent motion can be enhanced in a specific direction by presenting auditory stimuli. This effect mainly relies on the established temporal ventriloquism effect of audition on vision, which causes visual stimuli to be perceived as being temporally synchronous with asynchronously presented sounds (Gebhard and Mowbray, 1959; Shams et al., 2002; Morein-Zamir et al., 2003; Spence and Squire, 2003). The effect described here, however, cannot be attributed to temporal

ventriloquism as in our paradigm there were no differences in the relative timing of auditory and visual stimuli across the two auditory conditions. A spatial ventriloquism effect of audition on vision, as described by Alais and Burr (2004), also cannot explain our results because the visual stimuli employed in this experiment were not spatially ambiguous. Therefore, it appears that the effect of auditory stimuli on the perceived direction of visual motion reported in this study can only be attributed to the direction of auditory motion.

Our results taken together with previous studies (Soto-Faraco et al., 2002; Soto-Faraco et al., 2003; Soto-Faraco et al., 2004; Soto-Faraco et al., 2005; Sanabria et al., 2007), indicate that visual and auditory motion perception interact with each other bidirectionally. Such an interaction suggests that visual and auditory motion perception are supported by two interacting neural systems rather than by two independent cognitive modules (Fodor, 1983) as our results indicate a violation of the information encapsulation criterion for modularity. Interestingly, recent neuroimaging studies have shown that neural activation in hMT/V5+, a cortical area known to support visual motion perception (Tootell et al., 1995), is affected by auditory motion stimuli (Poirier et al., 2005; Poirier et al., 2006; Alink et al., 2008, Saenz et al., 2008). These findings are suggestive for hMT/V5+ supporting both visual and auditory motion perception which might clarify why visual and auditory motion perception are found to interact with each other in this study. However, further evidence for hMT/V5+ supporting both visual and auditory motion perception still needs to be put forward. For example, it still needs to be shown that disrupting activation in hMT/V5+, e.g. by means of transcranial magnetic stimulation (TMS), reduces one's visual and auditory ability to determine the direction of motion.

To summarize, the study at hand shows that the direction of auditory motion can bias the directionality of perceived visual motion when the direction of visual motion is ambiguous. Taken together with previous studies, this finding suggests that visual and auditory motion perception interact with each other bidirectionally. Based on this perceptual interdependency and previous neurophysiological findings we speculate that visual and auditory motion perception might rely on a common set of cortical areas including hMT/V5+.



**Supplementary figure**

This figure contains the plotted data for an exemplary subject and the fitted Boltzmann functions for all three conditions. The visual apparent motion percept of this subject was estimated to be bistable for a shift of .477 of the inter-column distance. For this bistability shift the subject was estimated to perceive the visual apparent motion stimulus as moving rightwards in 70% of the rightwards auditory-motion trials and in 46% of the leftwards auditory motion trials.

**References**

Alais D, Burr D (2004) The ventriloquist effect results from near-optimal bimodal integration. *Current Biology* 14:257–262

Alink A, Singer W, Muckli L (2008) Capture of auditory motion by vision is represented by an activation shift from auditory to visual motion cortex. *Journal of Neuroscience* 28:2690

Battaglia PW, Jacobs RA, Aslin RN (2003) Bayesian integration of visual and auditory signals for spatial localization. *Journal of the Optical Society of America A* 20:1391–1397

Beauchamp MS, Yasar NE, Kishan N, Ro T (2007) Human MST but not MT responds to tactile stimulation. *Journal of Neuroscience* 27:8261

- Blake R, Sobel KV, James TW (2004) Neural synergy between kinetic vision and touch. *Psychological Science* 15:397–402
- Dawson MR (1991) The how and why of what went where in apparent motion: Modeling solutions to the motion correspondence problem. *Psychological Review* 98:569–603
- Draper NR, Smith H (1981) *Applied regression analysis*.
- Freeman E, Driver J (2008) Direction of visual apparent motion driven solely by timing of a static sound. *Current Biology* 18:1262–1266
- Fodor JA (1983) *The modularity of mind*. MIT press Cambridge, MA.
- Gebhard JW, Mowbray GH (1959) On discriminating the rate of visual flicker and auditory flutter. *The American journal of psychology* 72:521–529
- Hagen MC, Franzén O, McGlone F, Essick G, Dancer C, Pardo JV (2002) Tactile motion activates the human middle temporal/V5 (MT/V5) complex. *European Journal of Neuroscience* 16:957–964
- Howard IP, Templeton WB (1966) *Human spatial orientation*. Wiley New York:
- Knill DC, Pouget A (2004) The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends in Neurosciences* 27:712–719
- Morein-Zamir S, Soto-Faraco S, Kingstone A (2003) Auditory capture of vision: examining temporal ventriloquism. *Cognitive Brain Research* 17:154–163
- Poirier C, Collignon O, DeVolder AG, Renier L, Vanlierde A, Tranduy D, Scheiber C (2005) Specific activation of the V5 brain area by auditory motion processing: an fMRI study. *Cognitive Brain Research* 25:650–658
- Poirier C, Collignon O, Scheiber C, Renier L, Vanlierde A, Tranduy D, Veraart C, De Volder AG (2006) Auditory motion perception activates visual motion areas in early blind subjects. *Neuroimage* 31:279–285
- Pouget A, Deneve S, Duhamel JR (2002) A computational perspective on the neural basis of multisensory spatial representations. *Nature Reviews Neuroscience* 3:741–747
- Sanz M, Lewis LB, Huth AG, Fine I, Koch C (2008) Visual motion area MT+/V5 responds to auditory motion in human sight-recovery subjects. *Journal of Neuroscience* 28:5141
- Sanabria D, Spence C, Soto-Faraco S (2007) Perceptual and decisional contributions to audiovisual interactions in the perception of apparent motion: a signal detection study. *Cognition* 102:299–310
- Shams L, Kamitani Y, Shimojo S (2002) Visual illusion induced by sound. *Cognitive Brain Research* 14:147–152
- Shechter S, Hochsteinn S, Hillman P (1988) Shape similarity and distance disparity as apparent motion correspondence cues. *Vision Research* 28:1013–1021

## Chapter 4

- Soto-Faraco S, Kingstone A, Spence C (2003) Multisensory contributions to the perception of motion. *Neuropsychologia* 41:1847–1862
- Soto-Faraco S, Lyons J, Gazzaniga M, Spence C, Kingstone A (2002) The ventriloquist in motion: Illusory capture of dynamic information across sensory modalities. *Cognitive Brain Research* 14:139–146
- Soto-Faraco S, Spence C, Kingstone A (2004) Cross-modal dynamic capture: congruency effects in the perception of motion across sensory modalities. *Journal of Experimental Psychology Human Perception and Performance* 30:330–345
- Soto-Faraco S, Spence C, Kingstone A (2005) Assessing automaticity in the audiovisual integration of motion. *Acta psychologica* 118:71–92
- Spence C, Squire S (2003) Multisensory Integration: Maintaining the Perception of Synchrony. *Current Biology* 13:R519-R521
- Tootell RB, Reppas JB, Kwong KK, Malach R, Born RT, Brady TJ, Rosen BR, Belliveau JW (1995) Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *Journal of Neuroscience* 15:3215
- Wada Y, Kitagawa N, Noguchi K (2003) Audio-visual integration in temporal perception. *International Journal of Psychophysiology* 50:117-124
- Welch RB, Warren DH (1980) Immediate perceptual response to intersensory discrepancy. *Psychological bulletin* 88:638–667
- Williams ZM, Elfar JC, Eskandar EN, Toth LJ, Assad JA (2003) Parietal activity and the perceived direction of ambiguous apparent motion. *Nature neuroscience* 6:616–623

## Chapter 5

# Stimulus predictability reduces responses in primary visual cortex

Based on the publication:

Alink, A., Schwiedrzik, C. M., Kohler, A., Singer, W., & Muckli, L. (2010). Stimulus predictability reduces responses in primary visual cortex. *Journal of Neuroscience*, 30(8), 2960-2966

**In this functional magnetic resonance imaging study we tested whether the predictability of stimuli affects responses in primary visual cortex (V1). The results of this study indicate that visual stimuli evoke smaller responses in V1 when their onset or motion direction can be predicted from the dynamics of surrounding illusory motion. We conclude from this finding that the human brain anticipates forthcoming sensory input which allows predictable visual stimuli to be processed with less neural activation at early stages of cortical processing.**

Vision can be regarded as a continuous cascade of neural reactions to the light that enters our eyes. Several theoretical models, however, elaborate this view of vision by claiming that the brain is not merely reactive but also “proactive” or “predictive” (Mumford, 1992; Rao & Ballard, 1999; Erkhagen, 2003; Bar, 2007; Enns 2008; Bar, 2009; Friston 2009; Grossberg, 2009). By “predictive” we refer to the idea that the brain generates predictions that estimate the visual input it will most likely receive given the contextual information from the recent past. In their theoretical model, Rao and Ballard (1999) have put forward the idea that such predictions play a central role in vision. They propose that visual cortices learn statistical regularities of the natural world and only signal the unpredictable components of their sensory input to higher visual areas. As a result, predictable stimuli are thought to require less neural activation to be conveyed from lower to higher visual cortex.

The aim of this functional magnetic resonance imaging (fMRI) study is to test whether predictability of stimuli reduces responses in the human visual cortex as proposed in the above model of predictive coding. If this model holds, then we expect predictability to reduce visual responses in primary visual cortex (V1), which is the earliest stage of visual processing in the human cerebral cortex. In order to test this hypothesis, we assessed whether visual stimuli induce smaller V1 responses when their onset or motion direction can be predicted from the trajectory of surrounding illusory motion. In addition to assessing blood oxygenation level dependent (BOLD) responses in V1, we also measured responses in the human visual motion area hMT/V5+, due to its known involvement in the processing of dynamic visual stimuli (Zeki et al., 1991, Tootell et al., 1995).

## **Materials and Methods**

### **Subjects fMRI**

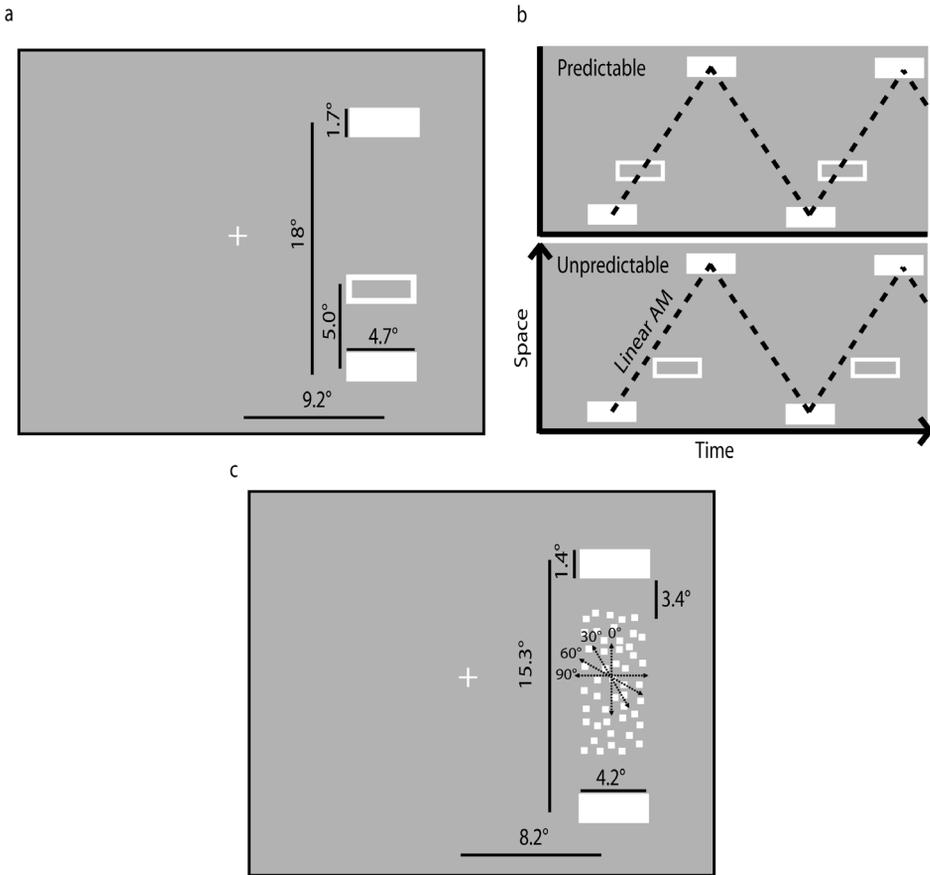
Twelve healthy subjects (6 male, 6 female) with normal or corrected-to-normal vision participated in the first fMRI experiment and five (4 male, 1 female) in the

second. All subjects gave their informed consent after being introduced to the experimental procedure in accordance with the Declaration of Helsinki.

### **Stimuli and task fMRI experiment 1**

Stimuli were generated using Presentation software (version 10.3., Neurobehavioral Systems, Inc.) and presented at a 60 Hz refresh rate using a projector (Sanyo Pro xtraX PLC-XP41 multiverse projector) with a zoom lens projecting from an adjacent room through a wave guide. Subjects viewed the stimuli through a tilted adjustable mirror (inside the head coil) on a screen which was attached to the back of the head coil. The viewable screen size subtended  $33.7^{\circ} \times 26.6^{\circ}$  of visual angle.

Subjects were instructed to fixate on a central fixation cross throughout the entire experiment, during which bars were consecutively presented  $9.0^{\circ}$  above and below the fixation cross and with a horizontal offset of  $9.2^{\circ}$  to the right (figure 1a). The screen had a gray background colour (luminance:  $28.3 \text{ cd/m}^2$ ), and the bars (height:  $1.7^{\circ}$ , width:  $4.7^{\circ}$ , luminance:  $139.0 \text{ cd/m}^2$ ) induced the impression of up- and downward long-range apparent motion (Ekroll et al., 2008; Beck et al., 1977; Kolars, 1963; Wertheimer, 1912; Exner, 1875) with a full-cycle frequency of 1.43 Hz. Each bar was presented for 200 ms with an inter-stimulus interval (ISI) between bar presentations of 150 ms. Critically, for predictable and unpredictable trials we briefly ( $16.7 \text{ ms}$ ) presented a test bar  $5.0^{\circ}$  above the lower bar during each of the ISIs in which upward apparent motion was perceived. During predictable trials, we presented this stimulus during the third frame after the offset of the lower bar stimulus, which corresponds to a presentation delay of  $41.7 \text{ ms}$  (assuming that the actual presentation occurred half-way through the third frame). This timing was chosen because it is exactly  $2.5/9$  of the ISI which corresponds to the ratio of the distance between the lower and the test stimulus to the total length of the apparent motion path ( $5.0^{\circ}/18^{\circ}$ ). Therefore, this stimulus is positioned and timed exactly on the motion trajectory of linear apparent motion between the lower and upper bar stimuli. For unpredictable trials, the test bar was presented at the same position during the seventh frame after the offset of the lower bar, corresponding to a  $108 \text{ ms}$  delay. This caused the test bar to appear at a time at which linear apparent motion had already passed the position of the test bar in unpredictable trials (for a schematic overview see figure 1b). The third trial contained the apparent-motion stimuli but no test bar and served to assess a baseline signal. Note here that this baseline included all other stimulus components except the test bar. Therefore, deconvolved BOLD responses for predictable and unpredictable trials can only be attributed to the presentation of the test bar and not to the presentation of the upper and lower bars because responses to these stimuli were subtracted by



**Figure 1**

Stimuli presented during experiments 1 and 2. a Schematic overview of the spatial layout of the stimuli presented during experiment 1. The upper and lower solid white bars represent the apparent-motion-inducing stimuli, which were presented for 200 ms with an inter-stimulus interval of 150 ms. The empty bar represents the test stimulus that was presented for 16 ms during upwards apparent motion, which occurred during the inter-stimulus interval following the presentation of the lower bar. b A schematic space-time plot that illustrates the time of presentation of the test bar relative to linear apparent motion during experiment 1. The dotted line represents the trajectory of linear apparent motion between the lower and the upper bar. For the predictable condition (upper), the test stimulus is presented at the time at which linear apparent motion passes the location of the test bar (41.7 ms after the offset of the lower bar). For the unpredictable condition (lower), the test bar is presented at the same location but with a greater delay than the predictable test bar (108 ms after the offset of the lower bar), which corresponds to the time at which linear apparent motion already passed the location of the test-bar stimulus. c A schematic depiction of the stimuli presented during experiment 2. Apparent-motion

stimuli were identical to those presented in experiment 1 although they were slightly smaller. During the inter-stimulus intervals, random-dot motion was presented on the path of apparent motion. The motion direction of these dots was either parallel to the apparent motion or 30°, 60° or 90° anticlockwise from the apparent-motion direction.

the baseline. All three types of trials lasted for seven seconds, during which ten apparent-motion cycles were presented. Subjects were presented with 81 trials of each type, distributed over three runs of fMRI measurements. We employed a rapid event-related design and, to ensure a correct deconvolution of the BOLD responses, used a sequence for the presentation of the trials within each run that was pseudo-randomized such that it ensured a two-back balanced trial history (Alink et al., 2008).

In order to localize the cortical representation of the test-bar stimulus in V1 and hMT/V5+, we presented inverting black-and-white checkerboards (spatial frequency: 1.2 cycles/deg, inversion frequency: 16 Hz, luminance white: 139.0 cd/m<sup>2</sup>, luminance black: 2.1 cd/m<sup>2</sup>) with the same location and extent as the test bar and the lower bar on a black background (2.1 cd/m<sup>2</sup>). These stimuli were presented in blocks of 16 seconds with 16-second fixation intervals which served as baseline. Throughout the entire run subjects fixated on a central white (139.0 cd/m<sup>2</sup>) fixation cross identical to the one in the main experiment.

### **Stimuli and task fMRI experiment 2**

During the second experiment we presented the same apparent-motion stimuli as during experiment 1. However, these were presented using an MR-compatible goggle system with two organic light-emitting diode displays (MR Vision 2000; Resonance Technology, Northridge, CA) which resulted in an 11% decrease in stimulus width and a 15% decrease in stimulus height due to the different screen size of the goggle system (30.0°×22.5°). The luminance of the gray background on this screen was 23.8 cd/m<sup>2</sup>, and the apparent-motion stimuli and fixation cross had a luminance of 44.0 cd/m<sup>2</sup>. During the entire experiment, subjects were instructed to maintain fixation on the fixation cross.

During the 150 ms ISI between the apparent-motion stimuli, we presented 150 dots (size = 0.1°, luminance = 44.0 cd/m<sup>2</sup>) randomly placed in an area with a width of 4.2° and a height of 8.25° centred on the apparent-motion path (see figure 1c). During these 150ms ISIs, the dots moved with a velocity of 9° per second in four different directions: parallel to apparent motion direction or 30°, 60° or 90° anti-clockwise from the apparent-motion direction. These random-dot motion configurations are referred to as 0°, 30°, 60° and 90° angle offset, respectively. Random-dot motion was presented during both upwards and

downwards apparent motion in opposite directions such that the 0°, 30° and 60° angle offset stimuli moved upwards during upwards apparent motion and *visa versa*. The 90° angle offset condition contained no vertical motion component and moved leftwards during upwards apparent motion and rightwards during downwards apparent motion. Moving dots exiting the motion area reappeared at the opposite side of the motion area. As in experiment 1, we used apparent motion without stimulation on the apparent-motion trace to assess the baseline signal and we presented the stimulus conditions in trials containing 10 apparent motion cycles, lasting seven seconds. In total, each angle offset trial type was presented 40 times and baseline trials were presented 120 times to all subjects, divided over four runs in a randomized order. Again, as in experiment 1 we used inverting checkerboards to localize the cortical representation in V1 of the area in which the random dots were presented.

### **fMRI procedure experiment 1**

Functional and anatomical MRI data were acquired with a 3T-MRI system (Siemens Trio; Siemens, Erlangen, Germany) using a standard CT head coil. During the presentation of the apparent motion stimuli we obtained three runs of 588 volumes containing 17 slices covering the occipital lobe as well as inferior parietal, inferior frontal and superior temporal regions for each subject using an echo planar imaging (EPI) sequence (repetition time (TR), 1000 ms; echo time (TE), 30 ms; flip angle, 62°; voxel size, 3.4 x 3.4 x 3.0 mm; Field Of View (FOV), 220 mm; gap thickness, 0.3 mm). Checkerboard stimuli were presented in a separate run during which 638 volumes were acquired using identical scanning parameters. All EPI images were corrected for spatial distortions using a point spread function sequence (Zaitsev et al., 2004). For each subject we also obtained a high resolution T1-weighted anatomical image using a Siemens MPRAGE sequence (1 x 1 x 1mm). For six of the subjects we also performed standard polar-angle retinotopic mapping using the same parameters employed routinely in our lab (Weigelt et al., 2007; Muckli et al., 2009). Furthermore, we measured eye movements during the fMRI experiment for eleven subjects using an infrared camera system placed outside the scanner room which measured the position of the right eye's pupil and cornea reflex at a rate of 60 Hz through a mirror system (Applied Science Laboratories, Bedford, MA, USA).

### **fMRI procedure experiment 2**

Functional and anatomical MRI data were acquired with a 3T-MRI system (Siemens Allegra; Siemens, Erlangen, Germany) using a four-channel head coil. During the presentation of the apparent-motion stimuli, we obtained four runs of 700 volumes containing 18 slices covering the occipital lobe as well as inferior

parietal, inferior frontal and superior temporal regions for each subject using an echo planar imaging (EPI) sequence (repetition time (TR), 1000 ms; echo time (TE), 30 ms; flip angle, 77°; voxel size, 3.3 x 3.3 x 3.5 mm; field of view (FOV), 210 mm; gap thickness, 0.35 mm). Checkerboard stimuli were presented in a separate run during which 484 volumes were acquired using identical scanning parameters. All EPI images were corrected for spatial distortions using a point-spread function (Zaitsev et al., 2004).

### **Analysis fMRI experiment 1**

Functional as well as anatomical MRI data were analyzed using the Brainvoyager QX software package (Brain Innovation, Maastricht, The Netherlands). The first four volumes of the functional runs were discarded to preclude T1 saturation effects. After pre-processing (motion correction, linear trend removal, temporal high pass filtering at 0.01 Hz and slice-scan-time correction), functional data for all subjects were aligned with the individual high-resolution anatomical MPRAGE image and transformed into Talairach space (Talairach & Tournoux, 1988). After manual correction for inhomogeneities, we created an inflated cortex reconstruction for all 12 subjects. For the six subjects for whom a polar angle map was acquired, we defined the V1-V2 borders on this cortex reconstruction as shown in supplementary figure 1. Regions of interest (ROIs) for the cortical representation of the location of the test-bar stimulus were defined individually in V1 as well as in hMT/V5+. The ROI in V1 consisted of the 500 mm<sup>3</sup> of cortex within the calcarine sulcus that responded most strongly when a checkerboard was presented at the location of the test bar but that showed no response when a checkerboard was presented at the lower bar location. The t-threshold that defined the minimum t-value of this area was different for each subject (see results section). This area was clearly within the borders of V1 for all six subjects for whom we mapped the V1-V2 border. This finding is in line with studies on human retinotopic organization of human primary visual cortex (Vanni et al., 2005) showing that stimulation both on and close to the horizontal meridian elicits activation within close proximity of the calcarine sulcus, which, when the eccentricity of the stimulus is sufficient, can easily be separated from parallel activation in V2. Because the six mapped subjects' data quality was sufficiently high to reproduce these stereotypical findings, we assumed that the ROIs defined in the calcarine sulcus for the other six subjects without defined V1-V2 borders should also be sound. For ten of the twelve subjects we were also able to select 200 mm<sup>3</sup> of cortex within V1 that was responsive to checkerboards presented at the lower bar location but not to checkerboards presented at the target location (for three exemplary subjects see supplementary figure 2a).

It has been shown that checkerboards with frequencies close to the one used here elicit BOLD responses in hMT/V5+ (Tootell et al. 1995). This was also apparent in our data, which allowed us to individually define ROIs consisting of 500 mm<sup>3</sup> of cortex in hMT/V5+ that was activated by the checkerboard presented at the test-bar location. Defining these ROIs allowed us to evaluate whether predictability in the context of apparent motion affected BOLD responses within the cortical representation of the test and lower bar in V1, and the representation of the test bar in hMT/V5+. This was tested on a group level by pooling the individually defined data from the main experiment originating from the ROIs in V1 and hMT/V5+ over all 12 subjects for the test-bar ROIs, and over ten subjects for the lower-bar ROIs. Over the pooled data, we computed a General Linear Model (GLM) for all ROIs using a deconvolution design (Glover, 2005) and tested whether the beta values for time points 4 to 12, which correspond to the peak of the BOLD response (4-12 seconds post stimulus), were significantly different for predictable as compared to unpredictable trials. P-values were corrected for multiple comparisons using a Bonferroni correction for the number of ROIs in which we compared BOLD responses between conditions. For ROIs that showed an effect, we assessed whether this effect was consistent across subjects by computing the direction of the difference for each subject individually and testing whether more subjects showed an effect in one direction than in the other than expected by chance using a sign test.

In addition to the ROI analysis, we also performed a group analysis over the entire brain volume to see whether we could find regions other than V1 and hMT/V5+ in which visual responses are affected by the predictability of the test stimulus. To this end we smoothed the functional data of each subject with a Gaussian kernel (8mm FWHM) and computed a GLM over the smoothed data across subjects. The effect of predictability was assessed by contrasting beta values across conditions for the time points 4 to 12 in conjunction with contrasts that tested whether both types of stimuli induced a significant signal increase as compared to baseline.

### **Analysis fMRI experiment 2**

The responses in V1 and hMT/V5+ to the different types of random-dot motion were analyzed using the same ROI approach as employed in experiment 1. For each subject we defined a ROI for V1 defined as a volume of 500 mm<sup>3</sup> close to the calcarine sulcus that responds to a checkerboard stimulus presented at the location of the random-dot motion area. For hMT/V5+ we defined the ROI as a volume of 500 mm<sup>3</sup> close to the posterior part of the inferior temporal sulcus that responded to all moving-dot configurations. Based on the group ROI data, we calculated an average BOLD response for V1 and hMT/V5+ for each of the angle

offsets. To test whether our hypothesis that predictability of the random-dot motion direction reduces visual responses was correct, we tested whether the most predictable motion type (angle offset  $0^\circ$ ) induced a significantly lower BOLD response than the least predictable motion type (angle offset  $90^\circ$ ). To this end, we tested whether beta values for time points 4 to 12 were significantly lower for the  $0^\circ$  angle offset condition. Furthermore, we tested whether angle offset linearly increases visual responses in V1 and hMT/V5+ by assessing the Pearson correlation between the mean beta value from time-points 4 to 12 and the angle offset.

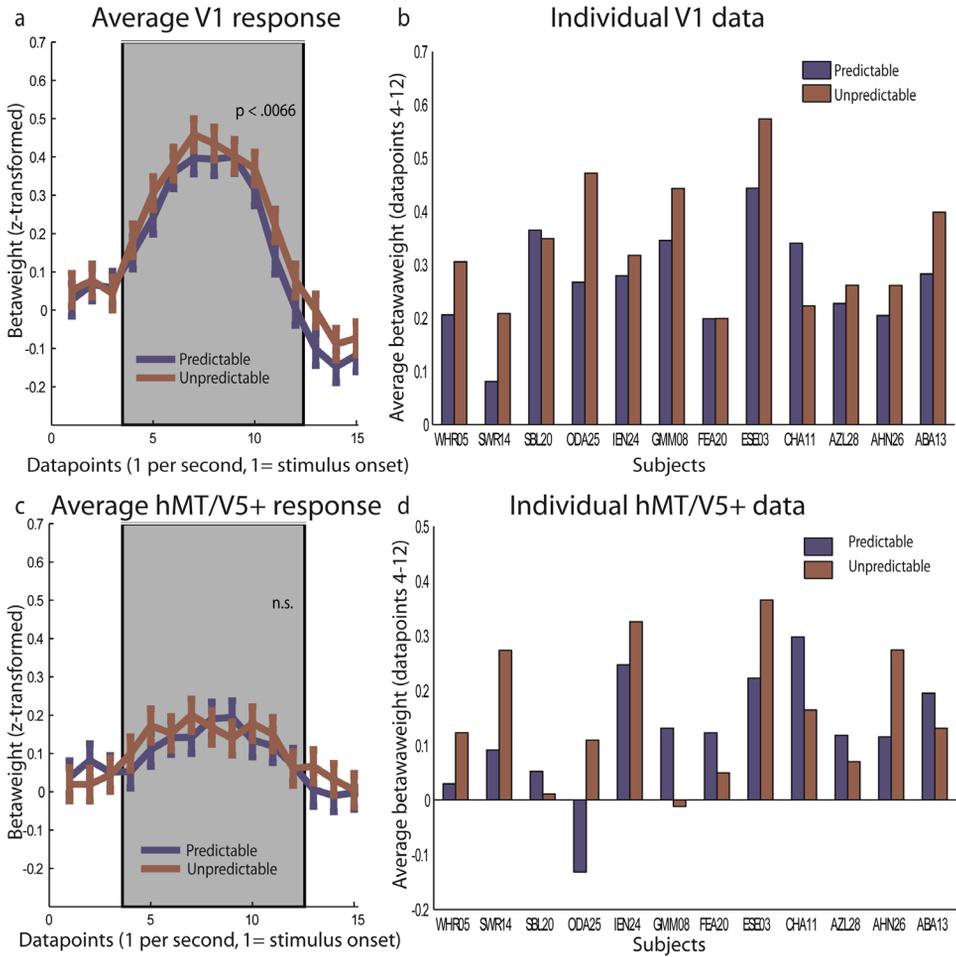
### **Analysis of eye movements**

For eleven subjects we calculated the mean and standard deviation of the horizontal and vertical position of fixation for the predictable and the unpredictable conditions over all data points that were outside a  $\pm 200$  ms interval of eye blinks (time points at which the pupil diameter was zero). We tested whether there were differences in mean and variance across conditions using a repeated-measures test over subjects. Furthermore, we created a density plot of eye position for both conditions using all eye-tracking data across all subjects (supplementary figure 3).

## **Results**

### **fMRI experiment 1**

We defined cortical ROIs within V1 and hMT/V5+ representing the position and extent of the predictable and unpredictable stimuli for all twelve subjects using individualized t-thresholds (mean [standard deviation] for t-thresholds in V1 = 5.15 [3.12], and in hMT/V5+ = 3.39 [1.99]; mean [standard deviation] of Talairach coordinates for V1:  $x = -6.7$  [2.9],  $y = -85.0$  [4.2],  $z = -2.1$  [2.9], and for hMT/V5+:  $x = -41.0$  [5.2],  $y = -75.0$  [5.2],  $z = 1.6$  [5.3]; for details see supplementary figure 1). From the data within these ROIs, pooled across all 12 subjects, we computed deconvolved BOLD responses for the predictable and unpredictable stimuli in V1 and hMT/V5+ (figure 2a and 2c). Within these ROIs, we analyzed BOLD responses to stimuli for which the onset could or could not be predicted from the trajectory of apparent motion. These stimuli are referred to as predictable and unpredictable stimuli, respectively, and were identical in all aspects besides the onset relative to the apparent-motion trajectory (see figure 1 for more details). We found that predictable stimuli gave rise to a significantly lower BOLD response in V1 than unpredictable stimuli ( $p < .0066$ , Bonferroni-corrected for the number of ROIs), while there appeared to be no effect of predictability within area hMT/V5+ ( $p > .05$ , Bonferroni-corrected). Individual responses in V1 turned



**Figure 2**

Left The grand-mean event-related BOLD responses for experiment 1 generated using deconvolution for the regions of interest representing the test bar in V1 and hMT/V5+. Event-related responses are shown for predictable trials (blue) and unpredictable trials (red). Right Individual BOLD-response peak amplitudes for the regions of interest V1 and hMT/V5+ expressed by the mean of the data points 4 to 12 for predictable (blue) and unpredictable (red) trials.

out to be reduced for predictable stimuli for ten out of twelve subjects (sign test:  $p < .05$ ) while only half of the subjects showed this effect in hMT/V5+, as expected by chance (sign test:  $p > .05$ ). Thus, our results indicate that the

predictability of the onset of a stimulus presented on the apparent-motion path reduces responses in V1 while not affecting hMT/V5+ responses.

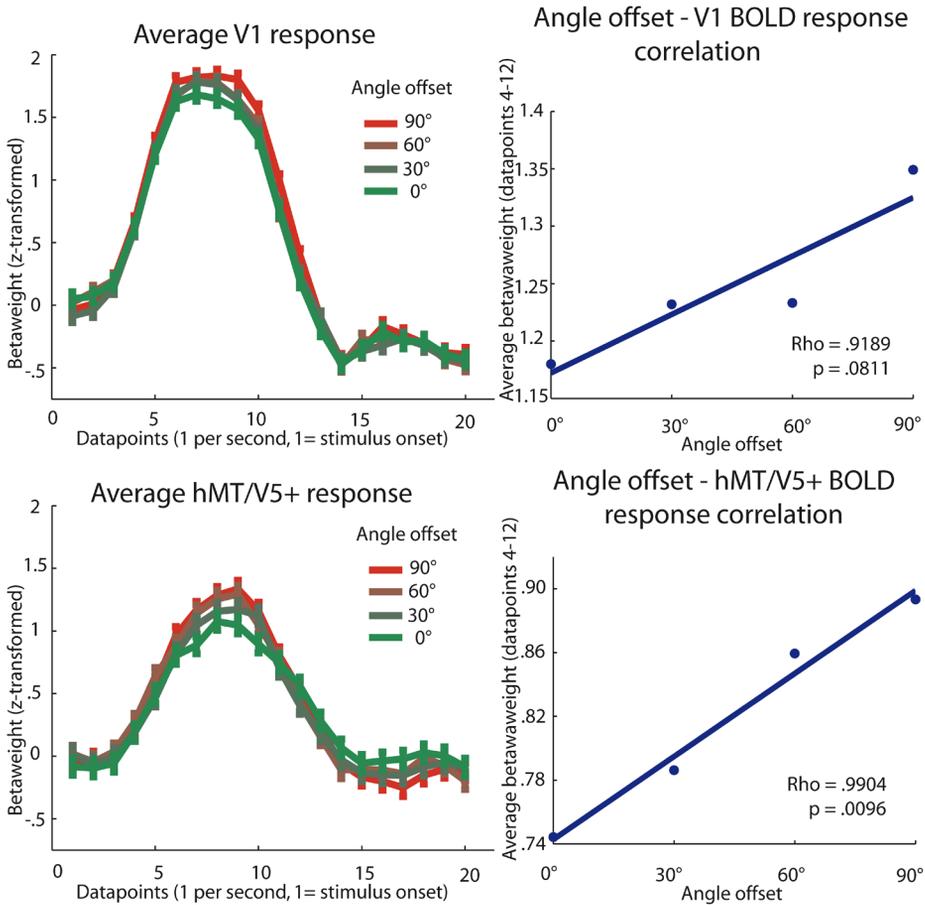
In order to test whether the effect of predictability in V1 was retinotopically specific we also analyzed BOLD responses of ROIs in V1 of ten subjects that represent the position and extent of the lower apparent-motion-inducing stimulus (mean [standard deviation] for t-thresholds: 5.06 [2.69]; mean [standard deviation] of Talairach coordinates:  $x = -2$  [3.0],  $y = -83.0$  [3.6],  $z = -1.2$  [3.4]). In this region, we observed no differences between BOLD responses to predictable and unpredictable stimuli. Thus, the effect that we observe in V1 for predictability is retinotopically specific to the V1 representation of the test stimulus presented on the apparent-motion path (for more details see supplementary figure 2).

We also performed a group analysis over the entire brain volume measured in experiment 1 in order to assess whether other regions beside V1 show an effect of stimulus predictability. This analysis did not identify any region that was significantly affected by stimulus predictability ( $p > .05$ , corrected using false discovery rate). Supplementary figure 5 shows a statistical map ( $p < .05$ , uncorrected) for this group analysis and demonstrates that the only activation clusters showing an effect of stimulus predictability (albeit not significant after the correction for multiple comparisons) are those inside or nearby the individual ROI volumes for V1.

To ensure that our effects did not result from differential fixation performance across conditions, we measured eye movements of our subjects inside the scanner. Differences between mean horizontal and vertical position of fixation over all subjects differed less than 0.1 degrees of visual angle between predictable and unpredictable trials ( $p > .05$  - repeated measure ANOVA). Also, the standard deviations for both dimensions did not differ across trial types ( $p > .05$ , repeated-measures ANOVA). Density plots of eye position show no gross differences in the distribution of fixation accuracy in space across conditions (supplementary figure 3).

## Results fMRI experiment 2

As in experiment 1, we analyzed BOLD responses in individual ROIs for V1 and hMT/V5+ (mean [standard deviation] for t-thresholds in V1 = 6.2 [2.0], and in hMT/V5+ = 9.0 [3.2]; mean [standard deviation] of Talairach coordinates for V1:  $x = -2.0$  [5.4],  $y = -81.4$  [4.4],  $z = 1.4$  [4.4], and for hMT/V5+:  $x = -41.6$  [1.9],  $y = -68.8.0$  [5.2],  $z = 7.2$  [3.3]). In order to test whether predictable motion in the context of apparent motion induces lower visual responses, we tested both in V1 and hMT/V5+ whether responses were lower for the most predictable motion-angle offset  $0^\circ$  as compared to the least predictable angle offset of  $90^\circ$ . Indeed, both



**Figure 3**

Left The grand-mean event-related BOLD responses for experiment 2 generated using deconvolution for the regions of interest representing the moving dots in V1 and hMT/V5+. Event-related responses are shown for motion that runs parallel to the apparent-motion direction as well as responses to motion deviating 30°, 60° and 90° anti-clockwise from the apparent-motion direction. Right Pearson correlation plots illustrating the positive correlation between the deviation of the random-dot motion direction from the apparent-motion direction and BOLD response amplitudes (average of data points 4-12) in V1 and hMT/V5+.

these areas exhibited a lower response when the random dots moved parallel to the apparent-motion direction (0°) as compared to responses to orthogonal motion (90°) (V1:  $p < .0005$ ; hMT/V5+:  $p < .0005$ , Bonferroni-corrected for the

number of ROIs). For both areas, we also tested whether there was a positive correlation between the angle offset and the visual response amplitudes. This turned out to be the case for both areas although the correlation in V1 did not reach significance (V1:  $p = .08$ ; hMT/V5+:  $p < .02$ , see figure 3 for more details).

### **Results psychophysical control experiment**

Previously we have shown that low-contrast stimuli that are predictable in the context of apparent motion are more readily detected (Schwiedrzik et al., 2007). In order to test whether this is also the case for the high contrast stimuli used in experiment 1, we performed a control experiment which contained both high- and low- contrast target stimuli. These stimuli were presented during both upward as well as downward apparent motion at two different positions along the apparent-motion path. The results of this experiment replicated our previous findings (Schwiedrzik et al., 2007). Thus, stimuli that are predictable in the context of apparent motion were detected more often than unpredictable stimuli (mean detection rate predictable = 38%, mean detection rate unpredictable = 32%,  $p < .03$  – repeated-measures ANOVA, two-sided test). Neither stimulus contrast, apparent-motion direction nor target position was found to interact with this effect (for details on the experimental procedure see the supplementary methods). From the current experiment, however, we cannot tell whether the difference in detection rates is due to a difference in  $d'$  or due to a response bias because our paradigm did not allow us to assess the correct rejections or the false-alarm rate. Our previous experiment (Schwiedrzik et al., 2007) however, indicated that the effect of predictability on detection rates was not due to a criterion shift. As the stimuli employed in the current psychophysical experiment are almost identical to those employed in our previous experiment, it is unlikely that the elevated detection rates reported here are due to a criterion shift.

The mean reaction time for predictable stimuli was 513 ms and for unpredictable stimuli 521 ms. This small difference in reaction time between these stimulus categories was, however, not significant ( $p > .05$ ). Reaction times were also not affected by stimulus contrast, apparent-motion direction or target position.

### **Discussion**

In this fMRI study we investigated whether predictable stimuli evoke smaller responses in V1 as implied by predictive-coding models (Rao & Ballard, 1999; Mumford, 1992). To this end, we measured BOLD responses in V1 to stimuli whose onset or motion direction could either be predicted or not predicted from their spatiotemporal context. Hence, we tested whether activation in a mapped

region in V1 was modulated by illusory motion induced by stimuli presented well outside the classical receptive field of this V1 region. Furthermore, we assessed whether stimulus predictability affected activation levels in the human visual motion area hMT/V5+.

The results of both experiments are in line with our hypothesis that stimulus predictability reduces activation levels in V1. The outcome of experiment 1 indicates that stimuli with a predictable onset give rise to lower V1 responses than identical stimuli presented with a less predictable onset. The second experiment shows that responses in V1 and hMT/V5+ are lowest when the direction of random-dot motion is predicted by the direction of apparent motion and that visual responses in these areas increase as the direction is made less predictable.

Our findings are in line with several other studies that have observed lower V1 responses for stimuli that fit their visual context. V1 has been shown to respond less to coherent than to incoherent motion (Bartels et al., 2008; Harrison et al., 2007; McKeefry et al., 1997) and less to grouped than to randomly arranged objects (Murray et al., 2002). Furthermore, face-selective areas in ventral visual cortex have been shown to respond less when a face stimulus is repeated in a continuous trajectory (Yi et al., 2008) and responses of neurons in the superior temporal sulcus of the monkey brain were shown to be suppressed and to occur at shorter latencies when stimulation consists of predictable sequences of natural images (Perrett et al., 2009). However, our study is the first to show that subtle changes in the spatiotemporal predictability of a stimulus affect stimulus processing in V1. Hence, in experiment 1 we show that V1 processes stimuli with less activation when their onset is predictable, even though luminance, size, position and duration of stimuli were kept constant. Experiment 1 also shows that this effect of predictability is constrained to the retinotopic representation of the test stimulus. Furthermore, we demonstrate in experiment 2 that responses in V1 decrease when the predictability of visual stimuli is parametrically increased.

Another important implication of experiment 1 is that lower responses in V1 can co-occur with higher detection rates. Although we did not measure behavioral responses inside the scanner during experiment 1, the results of our previous study (Schwiedrzik et al., 2007) taken together with the results of our psychophysical control experiment imply that the detectability of the high-contrast test stimuli used during this experiment should have been more detectable when they were predictable in the context of apparent motion. Thus, the present study implies that a predictable stimulus, that is more detectable, can induce a smaller BOLD response in V1 than an unpredictable and less detectable stimulus.

In experiment 1, we observed that the predictability of stimulus onset reduced V1 responses but that no similar effect was present in hMT/V5+. In experiment 2 however, both these areas were found to exhibit reduced responses when random-dot motion was more predictable in the context of apparent motion. One could conclude from these results that hMT/V5+ is affected by the predictability of motion direction but not by the predictability of stimulus onset. However, it is also possible that hMT/V5+ is sensitive to both these features but that we did not have a large enough signal-to-noise ratio to demonstrate this due to the low amplitude of this region's responses to static stimuli.

Given the results presented here, what can be said about the mechanisms that allow V1 to process predictable stimuli with less activation? According to the model of Rao and Ballard (1999), this would require feedback from higher-level visual areas specifying which stimulus input is likely to arrive in V1 given the current spatiotemporal context. Feedback from higher-level visual areas to V1 seems a likely explanation for the effects of stimulus predictability reported here as these areas have larger receptive fields than V1 allowing them to determine the trajectory of long-range apparent motion. (Angelucci & Bullier, 2003, Angelucci & Bressloff, 2006, Ichida et al., 2007). This fact, taken together with the observation that during long-range apparent motion hMT/V5+ sends feedback signals to V1 (Ahmed et al., 2008; Wibrall et al., 2009; Sterzer et al., 2006; Muckli et al., 2005), can be considered a strong indication that activation in hMT/V5+ drives the predictability effect in V1. However, several studies have suggested that local processing of feed-forward signals in V1 allows for more sophisticated neural computations than one would expect from classical receptive field models (Seriès et al., 2002; Seriès et al., 2003; Masland & Martin, 2007). Due to the low temporal resolution of fMRI we could not assess whether activation in hMT/V5+ precedes and drives the predictability effects in V1. Therefore, it still remains to be determined whether reduced responses in V1 to predictable stimuli results from feedback, local processing in V1 or, which is likely to be the case, an interaction between feedback and local processing in V1 (Erlhagen, 2003).

Yi and colleagues (2008) observed that ventral visual cortex responds less to continuously than to discontinuously moving objects and attributed this effect to subjects perceiving continuously moving objects more as a single entity, or a Gestalt. One could argue that the results presented here are due to a similar mechanism by claiming that predictable stimulus ensembles had a greater integrity as a Gestalt. Such an interpretation does, however, not stand in opposition to the predictive-coding model of Rao and Ballard (1999). They propose that predictions are based on statistical regularities of the natural world, which can be argued to be the basis of Gestalt principles (Brunswik & Kamiya, 1953; Elder & Goldberg, 2002). It is also worth mentioning in this context, that

the extra-classical receptive-field effects explained in the model of Rao and Ballard (1999) all relate to reduced neural responses to stimuli that form a Gestalt with their spatial surround based on collinearity.

Another explanation for a higher BOLD response in V1 to unpredictable stimuli could be that these types of stimuli induce greater pop-out. It could be that the unpredictable stimuli stood out more than the predictable stimuli due to their higher incompatibility with the surrounding apparent-motion stimuli. Such an attentional explanation would be in line with the finding of greater neural responses in V1 to stimuli that induce a stronger pop-out effect in macaques (Smith et al., 2007). However, as we have shown previously (Schwiedrzik et al., 2007) and replicated in the psychophysical control experiment: detection rates are lower for unpredictable flashes, which speaks against this attentional interpretation. Hence, if the unpredictable flash pops out more than the predictable flash, then it should also be detected more readily (Treisman, 1982). Furthermore, if the higher V1 response in experiment 1 was driven by attention, then, based on previous studies (Büchel et al., 1998; Beauchamp et al., 1997; Treue et al., 1996), one would expect that such motion-related attentional modulation would be even stronger in hMT/V5+, which is not compatible with our findings.

To summarize, in this study we show that the predictability of visual stimuli reduces neural responses in V1 and hMT/V5+. This finding provides strong empirical evidence for the idea that the visual cortex actively anticipates its visual input and that such anticipation allows predictable stimuli to be processed with less neural activation at the earliest cortical relay for visual processing. Furthermore, our results imply that predictable stimuli can be detected more readily than unpredictable stimuli, although unpredictable stimuli evoke greater V1 responses.

## **Supplementary Methods**

### **Subjects - Psychophysical Control Experiment**

Eleven healthy subjects (eight female) with normal or corrected to normal vision participated in the experiment. All subjects gave their informed consent after being introduced to the experimental procedure in accordance with the declaration of Helsinki.

### **Stimuli and task - Psychophysical Control Experiment**

Stimuli were generated using Presentation software (Version 10.3, Neurobehavioral Systems, Inc.) and presented at a 75 hertz refresh rate on a LCD monitor (Dell, 1908WFP, dimensions: 16 x 10 inches) which the subjects viewed

from 55 cm distance. We presented two bright squares (height: 3.3°, width: 3.3°, luminance: 222 cd/m<sup>2</sup>) on the right side of the screen with a horizontal eccentricity of 10 visual degrees which were located 9.5 visual degrees above and below the horizontal mid-line of the screen. The screen had a background luminance of 55.5 cd/m<sup>2</sup>. During a trial these squares were flashed consecutively on the screen for 67 ms with an inter-stimulus interval (ISI) of 67 ms which induced the percept of continuous vertical apparent motion with a full cycle frequency of 3.75 hertz. During the session there was always a fixation cross (0.3°, luminance: 222 cd/m<sup>2</sup>) present at the centre of the screen. During a session we presented 180 trials consisting of six apparent motion cycles with no inter-trial interval which induced a continuous apparent motion percept throughout the session. In total there were nine types of trials. During eight of those trial types a target stimulus (square, height: 3.3°, width: 3.3 °) was presented on the apparent motion path. These targets were presented during the third or fourth cycle in the trial in a randomized fashion and the target appeared either 6 visual degrees above the lower apparent motion inducer or 6 visual degrees below the upper apparent motion inducer. Furthermore, the target stimulus could appear in the ISI during which upward apparent motion was perceived or during the ISI during which downward apparent motion was perceived. The crucial manipulation was, however, the time at which the target stimulus was presented. This could be either early during the ISI (during the second frame which corresponds to approximately a delay of 20ms assuming that the stimulus was presented half way through the vertical refresh) or late during the ISI (during the fourth frame which corresponds to approximately a delay of 47ms). In the case of a presentation at the lower position the target was labelled as predictable when it was presented early during upward apparent motion or late during downward apparent motion. In the case of a presentation at the upper position the target was labelled as predictable when it was presented late during upward apparent motion or early during downward apparent motion. The other four stimulus configurations were labelled as unpredictable. All eight conditions were: predictable conditions: Upward Apparent Motion-Target Up-Late (UAM-TU-L), Upward Apparent Motion-Target Down-Early (UAM-TD-E), Downward Apparent Motion-Target Up-Early (DAM-TU-E), Downward Apparent Motion-Target Down-Late (DAM-TD-L); non-unpredictable conditions: Upward Apparent Motion-Target Up-Early (UAM-TU-E), Upward Apparent Motion-Target Down-Late (UAM-TD-L), Downward Apparent Motion-Target Up-Late (DAM-TU-L), Downward Apparent Motion-Target Down-Early (DAM-TD-E) (for an illustration see supplementary figure 4). The ninth trial type was a catch trial during which no target was presented and served to further decrease the predictability of the appearance of targets. Within a session we presented ten

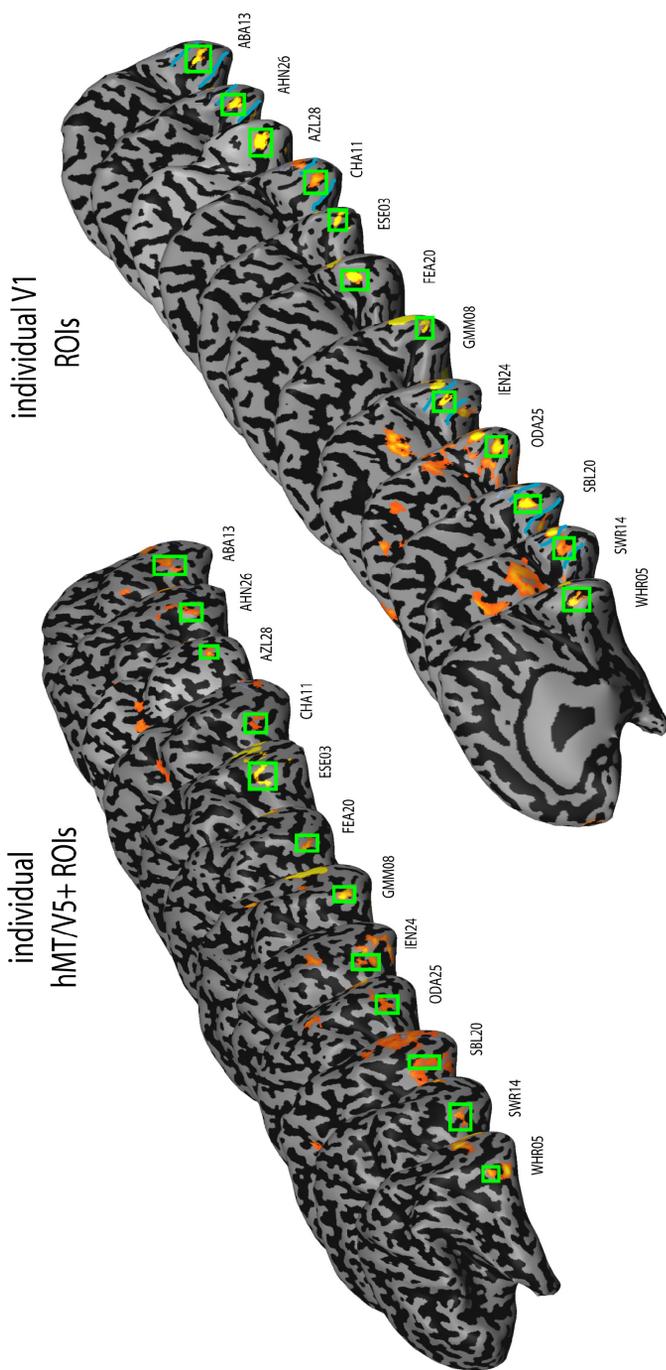
times each trial type using a randomized trial order. There were in total four sessions; two sessions during which the target stimulus was grey (luminance: 68.5 cd/m<sup>2</sup>.) and two sessions during which the target stimulus was white (luminance:222 cd/m<sup>2</sup>). The order of the sessions was counterbalanced across subjects with six subjects performing sessions in the order grey-white-grey-white and the other five subjects performing sessions in the order white-grey-white-grey. During these four sessions subjects were instructed to fixate at the fixation cross and to press the spacebar on the keyboard as fast as possible when they detected the target stimulus on the apparent motion path. Responses were registered as hits when they pressed the space bar in between 200 and 1500 ms after the target stimulus presentation.

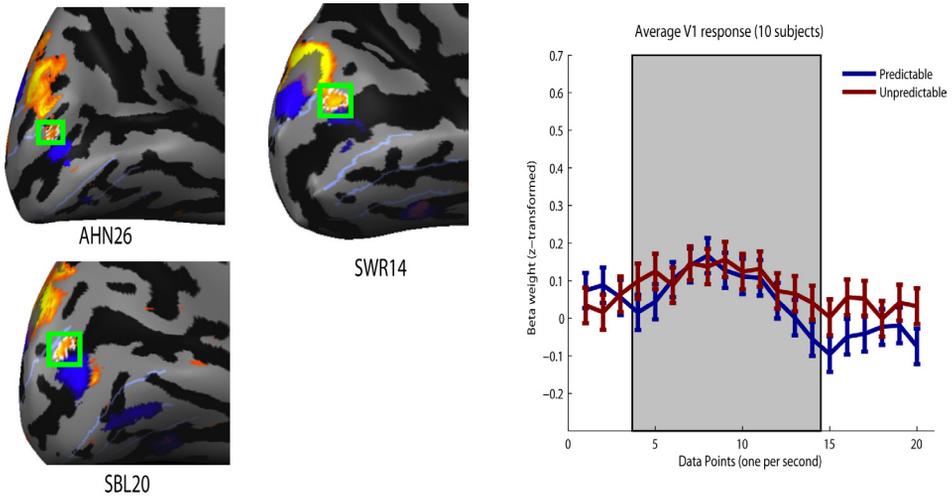
## Supplementary Figures

### Supplementary figure 1 (right page)

Lateral and medial view of cortex reconstructions of the left hemisphere for all twelve subjects measured in experiment 1 (medial view is mirrored for illustrative purposes) Right Regions of interest for V1 defined on the cortex segmentation for all twelve subjects. Regions of interest consist of the 500 mm<sup>3</sup> of cortex within the calcarine sulcus that responded most strongly when an inverting checkerboard was presented at the location of the test bar, but that showed no response when an inverting checkerboard was presented at the lower-bar location. The t-threshold was defined for each subject individually. For the ROIs in V1, the t-value was on average 5.15 with a standard deviation across subjects of 3.12. For the ROIs in hMT/V5+, the t-value was on average 3.39 with a standard deviation across subjects of 1.99. For half of the subjects, the V1-V2 border (blue lines) was defined using standard polar-angle mapping. Left Regions of interest for hMT/V5+ defined as 500 mm<sup>3</sup> of cortex within the posterior part of the middle temporal gyrus that responded to an inverting checkerboard presented at the location of the test bar.

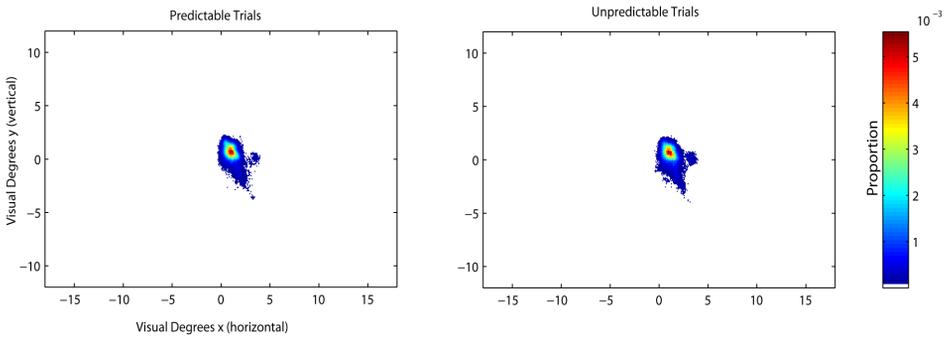
# Stimulus predictability reduces responses in primary visual cortex





**Supplementary figure 2**

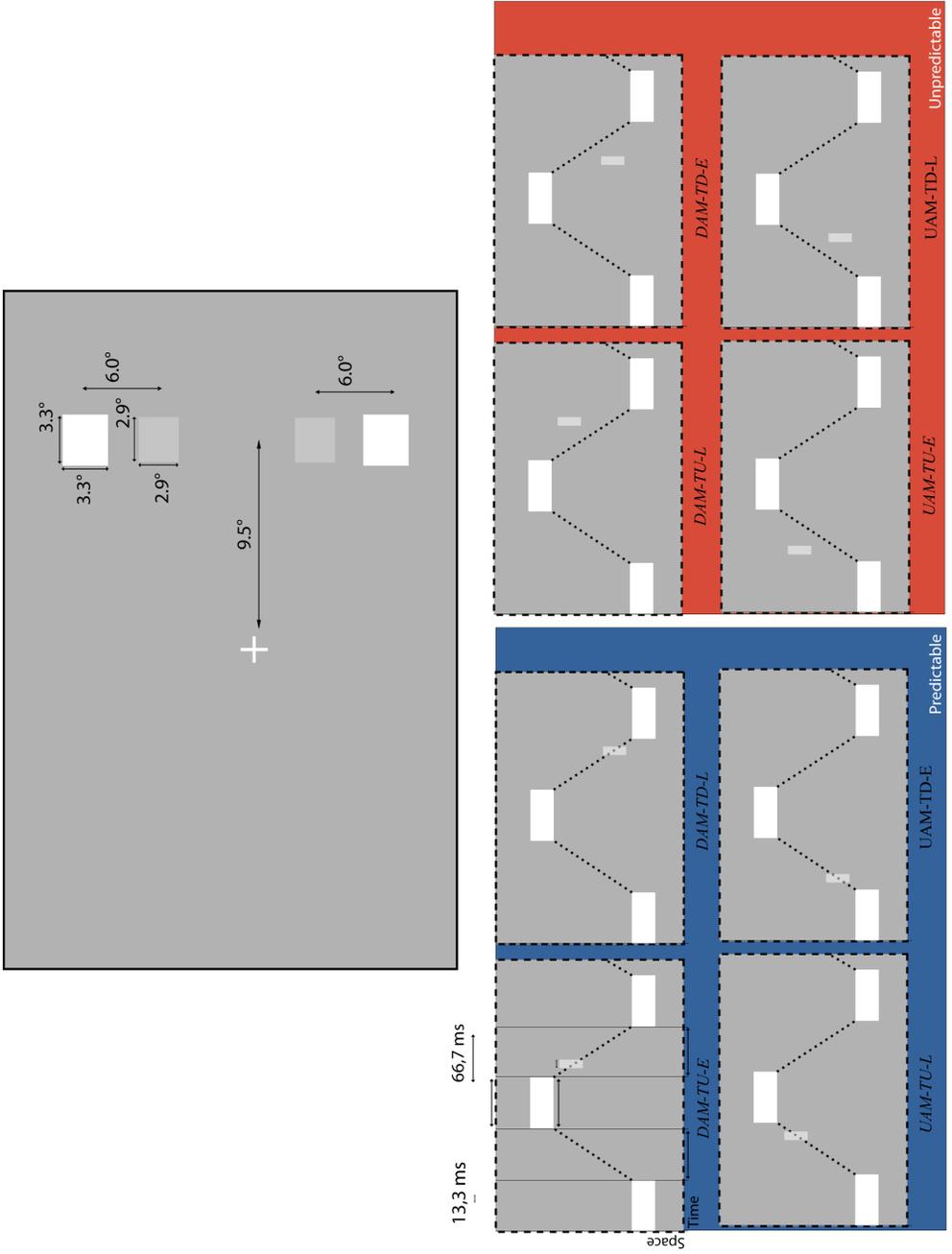
Left Medial view of occipital cortex for the three exemplary subjects AHN26, SBL20 and SWR14. Regions that respond to an inverting checkerboard presented during experiment 1 at the location of the test-bar stimulus location are colored blue and regions responding to an inverting checkerboard presented at the location of the lower apparent-motion inducer are colored yellow-orange. The yellow-orange areas inside the green boxes are the regions of interest for the lower apparent-motion-inducer representation in V1 for these three subjects. The  $t$ -thresholds were defined for each subject individually, which were on average 5.06 with a standard deviation across subjects of 2.69. Right Responses to the predictable and unpredictable test-bar stimuli within the region of interest for the lower apparent-motion-inducer representation in V1, pooled over 10 subjects.



**Supplementary figure 3**

Spatial distribution of fixation recorded during experiment 1 based on the eye-tracking data of eleven subjects during the predictable trials (left) and the unpredictable trials (right).

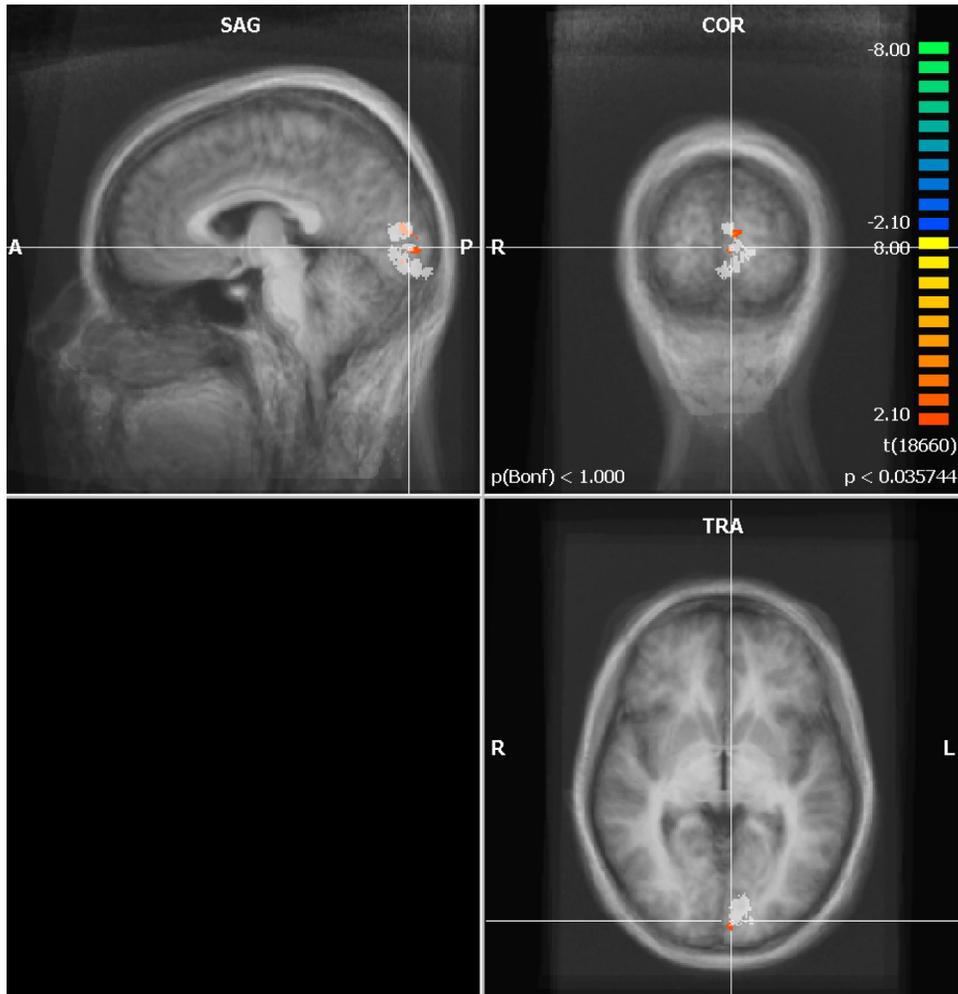
Stimulus predictability reduces responses in primary visual cortex



**Supplementary figure 4 (previous page)**

Stimuli presented in the psychophysical control experiment. a Schematic overview of the spatial layout of the stimuli b A schematic space-time plot that illustrates the time of presentation of the target stimulus relative to linear apparent motion. The dotted line represents the trajectory of linear apparent motion between the lower and the upper apparent-motion inducers. Abbreviations stand for: Upward Apparent Motion-Target Up-Late (UAM-TU-L), Upward Apparent Motion-Target Down-Early (UAM-TD-E), Downward Apparent Motion-Target Up-Early (DAM-TU-E), Downward Apparent Motion-Target Down-Late (DAM-TD-L), Upward Apparent Motion-Target Up-Early (UAM-TU-E), Upward Apparent Motion-Target Down-Late (UAM-TD-L), Downward Apparent Motion-Target Up-Late (DAM-TU-L), Downward Apparent Motion-Target Down-Early (DAM-TD-E).

## Stimulus predictability reduces responses in primary visual cortex



### Supplementary figure 5

A statistical map based on the group analysis of experiment 1 superimposed on the average anatomy across all twelve subjects aligned in Talairach space. The transparent white area running along the calcarine sulcus represents the brain volume in Talairach space that was selected at least once as V1 ROI across all subjects. This figure illustrates that the regions affected by stimulus predictability in experiment 1 ( $p < .05$ , uncorrected) are all located within or close by the representation of the test stimulus in V1.

### References

Ahmed B, Hanazawa A, Undeman C, Eriksson D, Valentiniene S, Roland PE (2008) Cortical dynamics subserving visual apparent motion. *Cereb. Cortex* 18:2796-2810

- Alink A, Singer W, Muckli L (2008) Capture of auditory motion by vision is represented by an activation shift from auditory to visual motion cortex. *J. Neurosci* 28:2690-2697
- Angelucci A, Bressloff PC (2006) Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons. *Prog. Brain Res* 154:93-120
- Angelucci A, Bullier J (2003) Reaching beyond the classical receptive field of V1 neurons: horizontal or feedback axons? *J. Physiol. Paris* 97:141-154
- Bar M (2007) The proactive brain: using analogies and associations to generate predictions. *Trends Cogn. Sci. (Regul. Ed.)* 11:280-289
- Bar M (2009) The proactive brain: memory for predictions. *Philos. Trans. R. Soc. Lond., B, Biol. Sci* 364:1235-1243
- Bartels A, Zeki S, Logothetis NK (2008) Natural vision reveals regional specialization to local motion and to contrast-invariant, global flow in the human brain. *Cereb. Cortex* 18:705-717
- Beauchamp MS, Cox RW, DeYoe EA (1997) Graded effects of spatial and featural attention on human area MT and associated motion processing areas. *J. Neurophysiol* 78:516-520
- Beck J, Elsner A, Silverstein C (1977) Position uncertainty and the perception of apparent movement. *Perception & Psychophysics* 21:33-38
- Brunswik E, Kamiya J (1953) Ecological cue-validity of 'proximity' and of other Gestalt factors. *The American journal of psychology* 66:20-32
- Büchel C, Josephs O, Rees G, Turner R, Frith CD, Friston KJ (1998) The functional anatomy of attention to visual motion. A functional MRI study. *Brain* 121 ( Pt 7):1281-1294
- Ekroll V, Faul F, Golz J (2008) Classification of apparent motion percepts based on temporal factors. *J Vis* 8:31.1-22
- Elder JH, Goldberg RM (2002) Ecological statistics of Gestalt laws for the perceptual organization of contours. *Journal of Vision* 2:324-353
- Enns JT, Lleras A (2008) What's next? New evidence for prediction in human vision. *Trends Cogn. Sci. (Regul. Ed.)* 12:327-333
- Erlhagen W (2003) Internal models for visual perception. *Biological cybernetics* 88:409-417
- Exner S (1875) \Über das Sehen von Bewegung und die theorie des zusammengesetzten Auges. *Sitzungsberichte/Akademie der Wissenschaften in Wien. Mathematisch-naturwissensch\äftliche Klassen. Abteilung III: Anatomie un Physiologie des Menschen und der Tiere, sowie theoretische Medizin* 72:156-190

- Friston K, Kiebel S (2009) Predictive coding under the free-energy principle. *Philos. Trans. R. Soc. Lond., B, Biol. Sci* 364:1211-1221
- Glover GH (1999) Deconvolution of impulse response in event-related BOLD fMRI. *Neuroimage* 9:416-429
- Grossberg S (2009) Cortical and subcortical predictive dynamics and learning during perception, cognition, emotion and action. *Philos. Trans. R. Soc. Lond., B, Biol. Sci* 364:1223-1234
- Harrison LM, Stephan KE, Rees G, Friston KJ (2007) Extra-classical receptive field effects measured in striate cortex with fMRI. *Neuroimage* 34:1199-1208
- Ichida JM, Schwabe L, Bressloff PC, Angelucci A (2007) Response facilitation from the "suppressive" receptive field surround of macaque V1 neurons. *J. Neurophysiol* 98:2168-2181
- Kolers PA (1963) A difference between real and apparent visual movement. *Nature* 197:271-272
- Masland RH, Martin PR (2007) The unsolved mystery of vision. *Current Biology* 17:577-582
- McKeefry DJ, Watson JD, Frackowiak RS, Fong K, Zeki S (1997) The activity in human areas V1/V2, V3, and V5 during the perception of coherent and incoherent motion. *Neuroimage* 5:1-12
- Muckli L, Kohler A, Kriegeskorte N, Singer W (2005) Primary visual cortex activity along the apparent-motion trace reflects illusory perception. *PLoS Biol* 3:e265
- Muckli L, Naumer MJ, Singer W (2009) Bilateral visual field maps in a patient with only one hemisphere. *Proc. Natl. Acad. Sci. U.S.A* 106:13034-13039
- Mumford D (1992) On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biol Cybern* 66:241-251
- Murray SO, Kersten D, Olshausen BA, Schrater P, Woods DL (2002) Shape perception reduces activity in human primary visual cortex. *Proc. Natl. Acad. Sci. U.S.A* 99:15164-15169
- Perrett DI, Xiao D, Barraclough NE, Keyzers C, Oram MW (2009) Seeing the future: Natural image sequences produce "anticipatory" neuronal activity and bias perceptual report. *Q J Exp Psychol (Colchester)* 62:2081-2104
- Rao RP, Ballard DH (1999) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci* 2:79-87
- Schwiedrzik CM, Alink A, Kohler A, Singer W, Muckli L (2007) A spatio-temporal interaction on the apparent motion trace. *Vision Res* 47:3424-3433
- Series P, Georges S, Lorceau J, Fregnac Y (2002) Orientation dependent modulation of apparent speed: a model based on the dynamics of feed-

- forward and horizontal connectivity in V1 cortex. *Vision Research* 42:2781–2798
- Seriès P, Lorenceau J, Frégnac Y (2003) The “silent” surround of V1 receptive fields: theory and experiments. *Journal of physiology-Paris* 97:453–474
- Smith MA, Kelly RC, Lee TS (2007) Dynamics of response to perceptual pop-out stimuli in macaque V1. *J. Neurophysiol* 98:3436–3449
- Sterzer P, Haynes J, Rees G (2006) Primary visual cortex activation on the path of apparent motion is mediated by feedback from hMT+/V5. *Neuroimage* 32:1308–1316
- Talairach J, Tournoux P Co-planar stereotaxic atlas of the human brain (1988) New York. NY: Thieme Medical Publishers
- Tootell R, Reppas J, Kwong K, Malach R, Born R, Brady T, Rosen B, Belliveau J (1995) Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *J. Neurosci.* 15:3215–3230
- Treisman A (1982) Perceptual grouping and attention in visual search for features and for objects. *J Exp Psychol Hum Percept Perform* 8:194–214
- Treue S, Maunsell JH (1996) Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature* 382:539–541
- Vanni S, Henriksson L, James AC (2005) Multifocal fMRI mapping of visual cortical areas. *Neuroimage* 27:95–105
- Weigelt S, Kourtzi Z, Kohler A, Singer W, Muckli L (2007) The cortical representation of objects rotating in depth. *J. Neurosci* 27:3864–3874
- Wertheimer M (1912) *Experimentelle Studien uber Sehen von Bewegung.*
- Wibral M, Bledowski C, Kohler A, Singer W, Muckli L (2009) The timing of feedback to early visual cortex in the perception of long-range apparent motion. *Cereb. Cortex* 19:1567–1582
- Yi DJ, Turk-Browne NB, Flombaum JI, Kim MS, Scholl BJ, Chun MM (2008) Spatiotemporal object continuity in human ventral visual cortex. *Proceedings of the National Academy of Sciences* 105:8840
- Zaitsev M, Hennig J, Speck O (2004) Point spread function mapping with parallel imaging techniques and high acceleration factors: fast, robust, and flexible method for echo-planar imaging distortion correction. *Magn Reson Med* 52:1156–1166
- Zeki S, Watson J, Lueck C, Friston K, Kennard C, Frackowiak R (1991) A direct demonstration of functional specialization in human visual cortex. *J. Neurosci.* 11:641–649

# Summary

## Summary

This dissertation describes and introduces four studies with the common theme motion perception. The main aim of the studies described in chapter 2, 3 and 4 is: finding out if the human brain processes visual and auditory motion signals in an independent or interactive manner. Chapter 5 of this dissertation relates solely to visual motion perception and describes a study that shows that stimuli that are more predictable induce lower responses in the primary visual cortex.

Three of the four experiments amassed in this dissertation make use of functional magnetic resonance imaging (fMRI). With this method one can monitor blood flow in neural tissue throughout the entire brain. Because increases in neural activity are accompanied by increases in local blood flow fMRI allows one to indirectly measure the level of activity in neural tissue. The spatial unit measured with this method is called a 'voxel' and typically covers 3x3x3mm of neural tissue. Despite the coarse spatial resolution of fMRI, it has been shown that with this method one can detect activation differences in neural populations subtending a volume much smaller than a typical voxel. This can be achieved by analyzing the information content of multi-voxel response patterns.

The study described in chapter 2 applies multi-voxel-pattern analysis to identify areas in the human brain that are sensitive to the direction of auditory motion. This analysis shows that auditory motion direction sensitivity affects activation patterns in the primary auditory cortex as well as in visual cortex directly adjacent to the visual motion area hMT/V5+. This implies that it is possible that auditory and visual motion signals interact with each other in hMT/V5+.

The fMRI experiment described in chapter 3 provides evidence for an interaction between visual and auditory motion signals in hMT/V5+ by showing that responses in this area are sensitive to the direction of auditory motion relative to visual motion. Furthermore, it is shown in this chapter that responses in hMT/V5+ and the auditory motion area, the planum temporale, are affected by an audiovisual motion illusion. In addition, both these areas appear to be activated by both auditory and visual motion presented in isolation.

Chapter 4 describes a psychophysical study demonstrating that the direction of moving sounds can alter the perceived directionality of a visual stimulus. This finding, taken together with previous studies, indicates that visual and auditory motion perception interact bidirectionally.

Chapter 5 describes two experiments indicating that responses in the primary visual cortex are reduced when the onset-time or motion-direction of a stimulus is predictable. This finding is in line with the idea that the human brain actively anticipates upcoming sensory events.

# Samenvatting

## Samenvatting

In dit proefschrift worden vier onderzoeksprojecten met als gezamenlijk thema “bewegingswaarneming” ingeleid en beschreven. De hoofddoelstelling van het onderzoek beschreven in hoofdstuk 2, 3 en 4, is: te weten komen of auditorische en visuele bewegingssignalen op een onafhankelijke of op een interactieve wijze in de hersenen verwerkt worden. In hoofdstuk 5 worden experimenten beschreven waarin aangetoond wordt dat wanneer stimuli in de context van visuele beweging voorspelbaar zijn ze een geringere respons teweegbrengen in de primaire visuele cortex.

Bij drie van de vier onderzoeksprojecten, bijeengebracht in dit proefschrift, wordt gebruik gemaakt van functionele kernspintografie (fMRI). Met deze methode kan bloeddorstrooming gemeten worden in neurale weefsel in het gehele brein. Omdat verhoogde neurale activiteit samengaat met een verhoogde lokale bloeddorstrooming maakt fMRI het mogelijk om indirect neurale activiteitsniveaus in de hersenen te meten. De ruimtelijke eenheid gemeten met deze methode, noemt men een ‘voxel’ en omvat in de regel  $3 \times 3 \times 3 \text{ mm}^3$  neurale weefsel. Ondanks de lage ruimtelijke resolutie van fMRI is het in vorige studies aangetoond dat men met deze methode activiteitsniveauverschillen kan meten van neuronpopulaties die een veel kleiner volume innemen dan een gebruikelijke voxel. Dit kan worden gerealiseerd door responspatronen op informatie-inhoud te analyseren.

In hoofdstuk 2 is zo'n multivoxel patroonanalyse toegepast om hersengebieden te identificeren die gevoelig zijn voor de richting van auditorische bewegingen. Deze analyse toont aan dat de bewegingsrichting van geluiden een invloed heeft op responspatronen in de primaire auditorische cortex en op responspatronen in visuele cortex aangrenzend aan hMT/V5+, een hersengebied dat visuele bewegingswaarneming ondersteunt. Deze laatste bevinding geeft aan dat het mogelijk is dat auditorische en visuele bewegingssignalen elkaar beïnvloeden in hMT/V5+.

In hoofdstuk 3 wordt voor zo'n interactie bewijs geleverd door te laten zien dat activiteit in hMT/V5+ gevoelig is voor de bewegingsrichting van geluiden ten opzichte van visuele beweging. In dit hoofdstuk wordt ook aangetoond dat een audio-visuele bewegingsillusie invloed uitoefent op activiteit in hMT/V5+ en het planum temporale, een hersengebied dat belangrijk is voor auditieve bewegingswaarneming. Tevens blijken deze hersengebieden zowel op visuele als op auditieve beweging te reageren.

Hoofdstuk 4 beschrijft psychofysisch onderzoek waarmee aangetoond wordt dat de richting van bewegende geluiden de waargenomen richting van bewegende visuele stimuli kan beïnvloeden. Deze bevinding, in samenhang met vorige studies, suggereert dat visuele en auditorische bewegingswaarneming elkaar wederzijds beïnvloeden.

In hoofdstuk 5 worden twee experimenten vermeld die aantonen dat responsen in de primaire visuele cortex kleiner zijn wanneer het moment van presentatie of de bewegingsrichting van stimuli voorspelbaar is. Dit is in overeenstemming met het idee dat de hersenen actief op sensorische gebeurtenissen anticiperen.



# Acknowledgements

## Acknowledgements

In this section I would like to thank all my family, friends and colleagues that have directly or indirectly contributed to this dissertation. If I, due to my flawed memory, did not adequately thank you for your indispensable influence on my work/social life/creativity/diet/mood/physical condition please indicate this on the next page and ask me for a signature.

I would like to thank:

Lars Muckli for taking me on as a diploma student and supporting me for so many years although I hardly did anything I was supposed to.

Axel Kohler for being the living example that science can be approached in an objective and thorough manner while remaining in a state of complete mental balance.

Professor Singer for being Professor Singer and always being there to support us.

Michael Wibral for making me aware of the pitfalls of fMRI, the hard way.

my ex-diploma students Caspar Schwiedrzik and Felix Euler without whom I would not have had much to write about in this dissertation.

all the members of the fMRI and EEG group for being such a relaxed bunch of nice colleagues.

Rainer Goebel, Bernadette Jansma and Mart Bles for getting me excited about odd things like ERPs and inflated cortex reconstructions.

all the members of the salad-club for providing me with leafy vegetables, appelstroop and good company and for allowing me to practice my daily mushroom-peeling-meditation.

Sara van Leeuwen, Tonio Heidegger, Florian Beißner, Christian Kell and Andreea Lazar for making life fun in Frankfurt.

my entire family for their constant support and in particular my mother for correcting my not-so-native-anymore Dutch in the "samenvatting".

Miriam Lee for helping with the design of this dissertation and for her unconditional support and affection.

I would like to thank \_\_\_\_\_ for

---

---

---

---

---

---

---

---

Yours Sincerely,  
Arjen Alink

---



# Curriculum Vitae

Arjen Alink was born on September 24 1983 in Nijmegen, the Netherlands. He completed his high school education at the Montessori College in Nijmegen after which started studying Psychology at the University of Maastricht (2000). During his studies he developed an interest for fundamental brain research and therefore he decided to participate in the study trajectory Cognitive Neuroscience. During the final year of his studies Arjen Alink performed his first scientific study, employing electroencephalography (EEG), together with Prof. Jansma and Dr. Bles at the University of Maastricht. After bringing this project to an end and finishing his exams he decided to go abroad to work on his master thesis as an intern at the renowned Max Planck Institute for Brain Research in Frankfurt, Germany. During this internship Arjen Alink learned how to investigate brain activation using functional magnetic resonance imaging (fMRI) and he was given the opportunity to investigate how the human brain integrates audiovisual motion signals. Arjen Alink finished his master thesis and obtained his master in Psychology in 2006 after which he started working as a PhD student within the same Max Planck department. During this period he continued investigating audiovisual integration of motion signals and performed research that led to the discovery that the human primary visual cortex processes predictable stimuli more efficiently than unpredictable stimuli. In the year 2009 Arjen Alink was awarded a post-doc position and was put in charge of the fMRI group consisting of two PhD students and several interns. During the last year he has gained expertise in applying multivoxel pattern analysis (MVPA) to identify areas in the human brain that encode the position and motion direction of sounds. Just before handing in the final version of his doctoral dissertation in 2010, the Dutch organisation for scientific research (NWO) has awarded Arjen Alink a research grant which allows him to perform post-doctoral research for a period of two years at the MRC Cognition and Brain Sciences Unit (CBU) in Cambridge (UK).

### **Publications in peer-reviewed journals**

- Alink, A., Schwiedrzik, C. M., Kohler, A., Singer, W., & Muckli, L. (2010). Stimulus predictability reduces responses in primary visual cortex. *Journal of Neuroscience*, 30(8), 2960
- Hein, G., Alink, A., Kleinschmidt, A., & Müller, N. G. (2009). The attentional blink modulates activity in the early visual cortex. *Journal of Cognitive Neuroscience*, 21(1), 197-206
- Alink, A., Singer, W., & Muckli, L. (2008). Capture of auditory motion by vision is represented by an activation shift from auditory to visual motion cortex. *The*

Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 28(11), 2690-2697

Schwiedrzik, C. M., Alink, A., Kohler, A., Singer, W., & Muckli, L. (2007). A spatio-temporal interaction on the apparent motion trace. *Vision Research*, 47(28), 3424-3433

Bles, M., Alink, A., & Jansma, B. M. (2007). Neural aspects of cohort-size reduction during visual gating. *Brain Research*, 1150, 143-154

Hein, G., Alink, A., Kleinschmidt, A., & Müller, N. G. (2007). Competing neural responses for auditory and visual decisions. *PLoS One*, 2(3), e320

Wibral, M., Muckli, L., Melnikovic, K., Scheller, B., Alink, A., Singer, W., & Munk, M. H. (2007). Time-dependent effects of hyperoxia on the BOLD fMRI signal in primate visual cortex and LGN. *NeuroImage*, 35(3), 1044-1063

### **Submitted manuscripts**

Alink, A., Euler, F., Kriegeskorte, N., Singer, W., Kohler, A. (2010). Auditory motion direction encoded in human primary auditory and high-level visual cortex. Submitted to the *Journal of Neuroscience*

### **Manuscripts in preparation for submission**

Alink, A., Euler, F., Galeano, E., Kriegeskorte, N., Singer, W., Kohler, A. (2010) Auditory motion capturing ambiguous visual motion. In preparation for submission to *Perception*

Melloni, L., van Leeuwen, S., Alink, A., Müller, N. (2010) Interaction between bottom-up saliency and top-down control in the human brain. In preparation for submission to the *Journal of Neuroscience*