SPATIAL ATTENTION AND EXPECTATION ACROSS AUDITION AND VISION

Arianna Zuanazzi

Primary supervisor: Prof. Uta Noppeney

A thesis submitted

to the University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

School of Psychology

College of Life and Environmental Sciences

University of Birmingham

September 2018

Abstract

To enable robust perception in the multisensory environment, the brain combines signals from different sensory modalities. Attention and expectation are two top-down mechanisms that facilitate perceptual processes. Previous research has shown that attentional resources are shared across sensory modalities. However, these studies often confounded attention and expectation, thus making it difficult to disentangle their respective effects.

We orthogonally manipulated spatial attention and expectation and evaluated their modality-specific and crossmodal behavioural and neural effects. We demonstrated that spatial attention and expectation are intimately linked by co-determining response probabilities. Furthermore, we showed that they rely on partly overlapping neural systems. From the multisensory perspective, we showed that behavioural modality-specific effects of response probability are larger than crossmodal effects and that auditory and visual expectation are differently built over time. At the neural level, we demonstrated that spatial attention engages neural resources interactively across sensory modalities, whereas spatial expectation is encoded in a modality-specific fashion. Finally, despite the pivotal role attributed to the posterior parietal cortex in arbitrating between integration and segregation of multisensory signals, we did not observe changes in the magnitude of audiovisual binding following application of transcranial magnetic stimulation to the right anterior intraparietal sulcus.

Keywords

Audition · ANOVA · Association cortices · Attention · Bayes factor · BCI · Behaviour · Decision-making · Crossmodal · Endogenous attention · Expectation · Eyetracking · fMRI · Integration · IPS · Modality-specific · Model comparison · MRI Multisensory · Neural correlates · Neuroimaging · Neuronavigation · Predictive coding · ROI · Perception · Perceptual decision making · Perceptual inference · Prediction · Response probability · RT · SDT · Segregation · Sensory cortices · Sensory reliability · Signal probability · Space · Task-relevance · TMS · Ventriloquist effect · Vision

Acknowledgements

Thanks to my supervisor prof. Uta Noppeney for our valuable scientific conversations, for her suggestions and support and for teaching me what determination really means.

Thanks to the examiners prof. Riikka Möttönen and prof. Chris Miall for kindly reading my thesis.

Thanks to the lab for, directly and indirectly, teaching me so many important aspects of science and life.

Thanks to my family and friends for being curious about my work and for encouraging me during these years.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ANOVA: analysis of variance

ACC: anterior cingulate cortex

AIC: Akaike information criterion

BCI: Bayesian causal inference

BIC: Bayesian information criterion

BOLD: blood oxygen level-dependent signal

CBF: cerebral blood flow

CC: calcarine cortex

CR: correct rejection

CSF: cerebrospinal fluid

cTBS: continuous TBS

D': d-prime

EEG: electroencephalography

EPI: echo-planar imaging

ERP: event-related potentials

FA: false alarm

FEF: frontal eye field

FWE: family-wise error

FIR: finite impulse response

fMRI: functional magnetic resonance imaging

FWHM: full width at half-maximum

GLM: general linear model

GM: grey matter

HRF: haemodynamic response function

IFG: inferior frontal gyrus

INS: insula

IPL: inferior parietal lobule

IPS: intraparietal sulcus

iTBS: intermittent TBS

LTD: long-term depression

LTP: long-term potentiation

MEP: motor evoked potentials

MFG: middle frontal gyrus

MNI: Montreal neurological institute

mPFC: medial prefrontal cortex

MR: magnetic resonance

MRI: magnetic resonance imaging

MT: motor threshold

NHST: null hypothesis significant testing

PPC: posterior parietal cortex

PT: phosphene threshold

RF: radiofrequency

RMSE: root mean square error

ROI: region of interest

RT: reaction times

rTMS: repetitive TMS

SE: spin echo

SDT: signal detection theory

SFS: superior frontal sulcus

SFG: superior frontal gyrus

SMG: superior marginal gyrus

SPL: superior parietal lobule

SPM: statistical parametric mapping

STG: superior temporal gyrus

T1: longitudinal (spin-lattice) relaxation time

T2: transversal (spin-spin) relaxation time

TBS: theta-burst stimulation

TE: echo time

TMS: transcranial magnetic stimulation

TPJ: temporoparietal junction

TR: repetition time

VE: ventriloquist effect

WM: white matter

List of Papers and Abstracts

The following four papers form chapters 3, 4, 5 and 6 of this thesis:

Published:

Zuanazzi A., Noppeney U. (2018). Additive and interactive effects of spatial attention and expectation on perceptual decisions. *Scientific Reports*, 8:6732.

In preparation/submission:

Zuanazzi A., Noppeney U. Modality-specific and crossmodal effects of response probability.

Zuanazzi A., Noppeney U. Distinct neural mechanisms of spatial attention and expectations guide perceptual inference in a multisensory world.

Zuanazzi A., Thielscher A., Noppeney U. No evidence that transcranial magnetic stimulation of right anterior intraparietal sulcus alters signal reliability or ventriloquist effect.

The following published conference abstracts are based on data included in this thesis:

Zuanazzi A., Noppeney U. (2018). To what extent do spatial attention and expectation rely on 'amodal' or modality-specific mechanisms? *CNS2018*, Boston, Massachusetts, USA.

Zuanazzi A., Noppeney U. (2017). Synergistic interplays between attention and expectation. A multisensory perspective. *TenYearsCIMeC*, Rovereto, Italy.

Zuanazzi A., Noppeney U. (2017). Do attention and expectation act interactively or additively? A multisensory perspective. *BNA*, Birmingham, UK.

Chapter 1: General Introduction

Previous research often confounded attention and expectation. Both attention and expectation facilitate perception but, while attention increases brain activity, expectation reduces it. Spatial attention engages multisensory processes in association cortices, i.e., PPC. However, the critical role of PPC in multisensory integration and segregation is debated. Moreover, multisensory properties of spatial expectation are yet not clear.

Chapter 2: Methodologies

1. Behavioural: GLM, model comparison, SDT
3. Eye-movements: eye-tracking
4. fMRI: preprocessing, GLM
5. TMS: experimental and stimulation protocols

Chapter 3: Study 1

Orthogonal manipulation of spatial attention and expectation produces seemingly contradictory interactive or additive behavioural effects. We explain these results by showing that attention and expectation are linked in co-determining response probability.

Chapter 4: Study 2

Crossmodal effects of response probability are similar in audition and vision. We attribute differences between modality-specific effects of response probability in audition and vision to the way probabilities are learnt over time, potentially influenced by sensory reliabilities.

Chapter 5: Study 3

Spatial attention and expectation engage partly overlapping neural systems. However, we show that attention activates frontoparietal regions interactively across audition and vision, whereas expectation activates sensory and frontoparietal regions selectively for the modality where it is directly manipulated.

Chapter 6: Study 4

Audiovisual binding depends on top-down attention, spatial proximity and sensory reliability, which we show is different in audition and vision. We find that neither audiovisual binding nor sensory reliability are altered by right anterior IPS-TMS.

Chapter 7: Final remarks

1. The operational definition of attention as ‘task-relevance’ and expectation as ‘signal probability’ sets up an artificial dichotomy.
2. It remains an open question whether it is possible to investigate attention (and expectation) selectively at the perceptual level by fully factoring out task and decision-making processes.
3. Sensory-specific features of the signal and different functions of sensory modalities in perception and action are pivotal when investigating modality-specific and multisensory processes, especially in relation to attention and expectation.

Thesis at a glance

Chapter 1:

General introduction

1.1 ATTENTION AND EXPECTATION: TWO SIDES OF THE SAME COIN?

In everyday life, attention and expectation often interact. For instance, when looking for an object, prior knowledge of the location of the object makes us direct our attention toward such location. Another example is surprising events. Our attention is usually attracted by events we do not expect to happen. While pioneering studies of spatial attention capitalised on this interaction and intentionally manipulated attention via stimulus probability, in more recent years a collection of studies strived to characterise attention and expectation in isolation and aimed to unravel their potential interactions when not confounded. The key ambition of these studies was to find out whether attention and expectation are two sides of the same coin or whether it is possible to outline a dichotomy in which attention and expectation are fully dissociated.

1.1.1 Definition of attention and expectation

The human brain is constantly exposed to an overwhelming amount of information and, due to its limited processing capacities, has to separate what is relevant from what is noise. Attention is a process which facilitates perception and action by prioritising some information and ignoring other (Carrasco, 2011). According to the *biased competition* hypothesis, stimuli which are relevant for behaviour compete for processing capacity and neural representation. This competition is biased towards attended stimuli. Furthermore, processing of these stimuli is facilitated whereas processing of irrelevant stimuli is suppressed (Desimone & Duncan, 1995; Beck & Kastner, 2005).

To date, attention has been investigated from multiple perspectives. Attention was studied in relation to a specific sensory modality (i.e. *modality specific attention*, Odegaard et al., 2016), to an object or its feature (i.e. *object- and feature-based attention*, Soto & Blanco, 2004) or to a region of space (i.e. *spatial attention*, Macaluso & Doricchi, 2013). The way attention is directed may be voluntary (i.e. *endogenous*/*sustained*) or induced by an abrupt change in the environment (i.e. *exogenous/transient*). Moreover, attention can be oriented with (i.e. *overt*) or without (i.e. *covert*) a change of eye fixation and head movement (Van Der Stigchel, 2007). In light of these different features of attention, it is overall clear that this mechanism entails a process of selection of a region/feature/object which, in the laboratory environment, is achieved by informing participants via cues or instructions where to direct their attention to and what to select. For this reason, *task-relevance* is often used as synonym of attention (Nobre & van Ede, 2017).

Similarly to attention, expectation (or prediction) facilitates perception and action. It does so by constraining forthcoming sensory signals and their features on the basis of prior information (Summerfield & Egner, 2009). This process is formally quantified in the *predictive coding* framework which postulates that the brain strives to minimize surprise, or rather *prediction error*. Prediction error is defined as the discrepancy between the brain self-generated predictions (i.e., *priors*) against the actual sensory input. Thus, perception is reduced to the problem of estimating priors given the data (i.e., posterior distribution in terms of Bayesian inference) and of minimizing prediction error (Friston, 2010). The predictive coding model assumes that information is passed across different levels of the cortical layers and cortical hierarchy. Specifically, at each level the prior is updated until prediction error is minimised. In particular, prediction error is computed at each level as the difference

between prediction from higher levels and sensory input from lower levels, and it is then propagated up the hierarchy (Friston, 2005). At the cellular level, it has been postulated that there exist *representational* units which encode the prior expectation about sensory information, and *error* units which encode the prediction error (Friston, 2010).

1.1.2 Effects of attention and expectation

1.1.2.1 Behavioural level: similar effects of attention and expectation

Facilitation at the behavioural level is observed for stimuli which are attended. A vast literature explored the effects of attention on behaviour, generally employing discrimination judgment or detection tasks. Facilitation is observed in terms of reaction times (Coull & Nobre, 1998; Geng & Behrmann, 2002, 2005), sensitivity (Wyart et al., 2012) and accuracy (van Ede et al., 2012). Attention also increases reliability of stimuli in the attended sensory modality (Vercillo & Gori, 2015). A similar pattern is evident for expectation: stimuli that are predicted are associated to improved performance at the behavioural level (Doherty et al., 2005; Rohenkohl et al., 2014).

1.1.2.2 Neural level: different effects of attention and expectation

At the neural level, attention is associated to increased brain activity in sensory areas. In particular, functional magnetic resonance imaging (fMRI) studies have shown that attention increases neural activity in primary sensory areas in the presence or absence of stimulation (Grady et al., 1997; Tzourio et al., 1997; Watanabe et al. 1998; Kastner et al., 1998, 1999; Somers et al., 1999; Kastner & Ungerleider, 2000; Liu et al., 2007; Silver et al., 2007; Serences et al., 2009; Carrasco, 2011). The effect of attention is not limited to sensory cortices but can also be observed in higher-order regions such as parietal and frontal cortices. A dorsal frontoparietal network (i.e.,

dorsal attention network) including intraparietal sulcus (IPS), superior parietal lobule (SPL) and frontal eye field (FEF) is activated when attention is (endogenously) oriented to a specific object or location (Kastner et al., 1999; Corbetta et al., 2000; Hopfinger et al., 2000; Shomstein & Yantis, 2006; Corbetta & Shulman, 2002, 2011). A (right-lateralised) ventral frontoparietal network (i.e., *ventral attention network*) including superior marginal gyrus (SMG), superior temporal gyrus (STG), inferior and middle frontal gyrus (IFG, MFG) and insula is activated for task-relevant (exogenous) reorienting of attention (i.e. stimulus driven reorienting, Kincade et al., 2005; Shulman et al., 2010; Corbetta & Shulman, 2011) or contextual updating (Geng et al., 2013). Reorienting to irrelevant events is prevented by suppressing the ventral network (Corbetta & Shulman, 2002). Conversely, when behaviourally relevant stimuli entail reorienting of attention, the two networks interact to shift attention (Corbetta et al., 2008). From an electrophysiological perspective, modulation of attention is broadly associated to a decreased power of alpha (8-12 Hz) oscillations for task-irrelevant stimuli (Worden et al., 2000; Thut et al., 2006; Händel et al., 2011; Weisz et al., 2011; Mazaheri et al., 2014). It has been hypothesised that alpha could reflect functional inhibition which would serve to gate sensory processing within the cortex (Jensen & Mazaheri, 2010; Jensen et al, 2014).

Differently from attention, expectation is shown to reduce neural response in sensory cortices when stimuli are expected or repeated (Murray et al., 2002; Grill-Spector et al., 2006; den Ouden et al., 2009; Alink et al., 2010; Todorovic et al., 2011; Kok et al., 2012a). A correspondent increase of activation is observed in frontal areas (Summerfield et al., 2006), which would support predictive coding models that hypothesise that prior expectation is sent top-down. In terms of electrophysiological findings, predicting an event contributes to delta-theta (2-8 Hz) phase resetting at the

prestimulus level and desynchronization of alpha oscillations at the stimulus onset (Arnal & Giraud, 2012).

Conversely, unexpected relative to expected stimuli elicit higher activation in primary sensory areas. This is interpreted in the predictive coding framework as evidence for larger prediction error in correspondence to unexpected events (Yoshiura et al., 1999; den Ouden et al., 2009, Lee & Noppeney, 2014). When a deviant event (i.e., a stimulus with different frequency/duration/location) interrupts a sequence of identical events, it is possible to observe a negative event-related potential (ERP) peaking at 150-250 ms from the deviant stimulus. This ERP is called *mismatch-negativity* and, together with an increase of gamma-band activity (> 30Hz), is thought to represent prediction error (Friston 2005; Näätänen et al., 2007; Arnal & Giraud, 2012).

1.1.3 Two sides of the same coin?

1.1.3.1 The Posner task and the confound between spatial attention and spatial expectation

Behavioural and neural aspects of spatial attention have been widely investigated in *hemineglect* patients due to the distinctive spatial connotation of this syndrome. Spatial neglect (or hemineglect) is a disabling neurological condition associated to unilateral right hemisphere lesion after a stroke (the right hemisphere is lesioned in over 90% of stroke survivors with spatial neglect, Corbetta et al., 2005) which is characterized by difficulties to process (i.e., attend, respond to) stimuli in the contralesional hemifield (generally the left hemifield), in the case when vision and motor control are not impaired. As a consequence, these stimuli often get unreported. Neuropsychological research on spatial attention has often employed the Posner task (or Posner cuing paradigm, Posner, 1980). In this task, participants are presented with

a cue which indicates the probability for a subsequent stimulus to be presented at a certain location. Participants are instructed to shift their attention to the most likely location. In the pioneering work of Corbetta et al. (2005), neglect patients were presented with a central arrow indicating which hemifield to attend. In 75% of the trials, the target was valid (i.e., presented in the hemifield indicated by the arrow) and in 25% of the trials the target was presented in the other hemifield (i.e. invalid). At the behavioural level, facilitation effects of validity (or attention) were observed. At the neural level, fMRI results showed that the activity in the parietal and dorsolateral prefrontal cortex (i.e., dorsal attention network) and in the tempo-parietal junction (TPJ) (i.e., part of the ventral attention network) in relation to attention orienting was weaker compared to that of healthy participants who performed the same task. The authors concluded that deficits in orienting and reorienting of attention were caused by an abnormal interaction between ventral and dorsal attention networks, underlying the importance of the interplay between these two networks for attentional processes (Corbetta & Shulman, 2011). Therefore, this unbalanced interaction would be the cause of the right-lateralised spatial bias that can be observed in neglect patients and also of their critical difficulties in disengaging attention to reorient it elsewhere (Corbetta et al., 2000; Corbetta & Shulman, 2011). Critically, because in the Posner task spatial attention is manipulated via signal probability, in this study attention and expectation were confounded. In fact, the experimental design of Corbetta et al. (2005) can be reduced to two main conditions: either stimuli are attended and expected (e.g., 75% of trials are presented in the valid hemifield, which is indicated by the attention cue) or they are unattended and unexpected (e.g., 25% of trials are presented in the invalid hemifield). Consequently, behavioural and neural effects of spatial attention and spatial expectation cannot be disentangled, making it difficult to separate

differential contributions of these two mechanisms in attention orienting (Summerfield & Egner, 2009).

To date, a conspicuous number of behavioural and neuroimaging studies employed the Posner cuing paradigm to investigate spatial attention across different healthy and patient populations, with different techniques and across different sensory modalities (for a brief review, Hayward & Ristic, 2013). Hence, facilitation effects of spatial attention and the neural correlates of orienting and reorienting reported in previous studies employing the Posner task (Nobre et at., 2000; Kincade et al. 2005; Bressler et al., 2008; Smith et al., 2010; Langner et al., 2011) might reflect expectation manipulation or the interaction between attention and expectation. Only recently has research delved further into the confound that the Posner task introduces.

1.1.3.2 Beyond the Posner task: spatial attention not confounded with spatial expectation

To disentangle the effects of spatial attention and expectation, the attention cue should not carry any information about the probability for a stimulus to be presented in either hemifield. Shulman et al. (2009) separately manipulated reorienting of attention and the probability of reorienting. The attention cue instructed participants about whether to maintain attention in a specific hemifield or to shift it to the other hemifield. Thus, the target stimulus always appeared in the cued hemifield. Additionally, the probability to maintain or shift attention was manipulated. Their results showed that both maintained attention and attentional shifts activated the dorsal attention stream. The precuneus showed higher activations for maintained attention, whereas IPS for shifts of attention in an increasing fashion for unexpected vs expected cue. FEF showed a mixing of both. Conversely, the right TPJ showed higher activations in correspondence to stimulus-driven reorienting independently of

expectation and thus functioned as a 'switch' to disengage attention. Separately, it also showed higher activation for breaches of expectation. These two effects were independent. Differently, right inferior frontal gyrus (IFG) was activated for unexpected reorienting only. Overall, these results suggest that, when attention and expectation are not confounded, different brain regions exhibit activations in correspondence to one or the other. Nevertheless, attention and expectation also share neural systems. In fact, certain brain regions (i.e., IPS, TPJ) showed a pattern of activation which depended on the contribution of the two.

The study of Kok et al. (2012b) similarly aimed to investigate the different contributions of spatial attention and expectation in the visual cortex. They manipulated spatial attention via an arrow indicating which hemifield to direct attention to and expectation via a cue that, before each block, indicated the hemifield where the stimulus was more likely to appear. Participants had to perform a discrimination task only for targets in the attended hemifield which could be presented with high or low probability in different blocks. The results of this orthogonal manipulation showed that unexpected relative to expected stimuli elicited higher activations in the primary visual cortex when they were unattended but the opposite pattern was found when stimuli were attended. Kok et al. concluded that attention reversed the effect of expectation. This interactive pattern speaks against the hypothesis that attention and expectation have additive effects and rather supports the hypothesis suggested by the predictive coding framework that attention and expectation function synergistically (Friston, 2009). Accordingly, attention would increase the precision of perceptual inference. This hypothesis postulates that attention works as a filter of relevant information which gets thereby higher precision, where precision is the inverse variance. Prediction errors are weighted by the precision of the

sensory information and by the prior information. Thus, perception is optimised by a smaller mismatch between prior expectation and sensory signal but also by a high precision of prediction errors. Hence, the sensory silencing induced by expectation would reverse in the presence of attention (Friston, 2009).

1.1.4 First aim of this thesis

The first aim of this thesis is to contribute to the understanding of the (independent or interactive) role that spatial attention and expectation play in perceptual decisions. In chapter 3, 4 and 5 we manipulate attention and expectation in an orthogonal fashion using a novel multisensory paradigm. Our multisensory approach allows us to evaluate the effect of spatial attention also at the behavioural level. In fact, if attention is manipulated as task-relevance, i.e., participants are asked to perform a task only for targets presented in the attended hemifield, no response is available for the unattended hemifield. This makes it impossible to investigate the main effect of attention and its interaction with expectation at the behavioural level. Using this multisensory approach, in chapter 3 we evaluate whether the behavioural effects of spatial attention and expectation are additive or interactive. In chapter 4 we extend the results of chapter 3 by swapping the roles of sensory modalities in the experimental design. Chapter 5 follows up the behavioural results of chapter 3 to further explore the (additive or interactive) effects of spatial attention and expectation at the neural level.

1.2 MULTISENSORY PROCESSES

Attention and expectation are often investigated in one single sensory modality. Nevertheless, to build a coherent percept of the surrounding environment the brain has to make sense of signals in multiple sensory modalities. Among others, previous studies focused on the following research questions: (1) where and how in the brain multisensory information is integrated and how this reflects on behaviour and (2) how multisensory processes are affected by attention and expectation.

1.2.1 Multisensory integration and segregation

1.2.1.1 Neural correlates of multisensory processes

A given brain region is defined as *multisensory* in virtue of the fact that unimodal areas converge to it, it shows activations for stimuli in different sensory modalities or if, when lesioned, a deficit involving multiple sensory modalities can be observed (Ghazanfar & Schroeder, 2006; Schroeder et al., 2009). Brain areas that are traditionally considered to be the locus of convergence of unisensory information (i.e., *association* or *heteromodal* areas) are the frontal cortex, the superior temporal sulcus (STS) and IPS (Calvert et al., 2001; Calvert & Thesen, 2004; Ghazanfar & Schroeder, 2006). Importantly, also areas which were initially considered as unisensory (i.e., specialised in the processing of signals in one sensory modality) have been shown to have multisensory properties (Falchier et al., 2002; Leitão et al., 2013). For instance, activity of the visual cortex can be modulated by auditory information and vice-versa, in the form of activation or deactivation (Laurienti et al., 2002; Johnson & Zatorre, 2005; Werner & Noppeney, 2011). Furthermore, unisensory areas reveal response enhancement following stimulation in multiple sensory modalities (Calvert et al.,

1999; Werner & Noppeney, 2011). Corroborating these results, animal and human studies have demonstrated that auditory and visual cortices are interconnected (Falchier et al., 2002; Eckert, 2008).

1.2.1.2 Multisensory integration in the posterior parietal cortex

In the multisensory field, much attention was devoted to the role of the posterior parietal cortex (PPC) in multisensory processes. Neuroimaging studies provided evidence for overlapping parietal cortex activations (especially in IPS) when auditory, visual or tactile signals are presented in isolation (Bremmer et al., 2001; Klemen & Chambers, 2012), interpreting this result as prove that PPC integrates signals from different sensory modalities (Calvert & Thesen, 2004). Multisensory integration provides an evolutionary advantage as the combination of different information deriving from different senses disambiguates external stimuli (Stein, 1998). In particular, when signals are integrated under the assumption that they are produced by a common source, an improvement in behavioural performance is observed. However, integration is not always the optimal strategy and, when signals are produced by different sources, to integrate them would be behaviourally detrimental. Previous fMRI studies have demonstrated that the brain can indeed perform integration and segregation of sensory signals in function of common-source judgment (Rohe & Noppeney, 2015a, Rohe & Noppeney, 2016). They demonstrated that integration and segregation are performed at different stages of the cortical hierarchy and that PPC is centrally involved in multisensory integration and segregation. In particular, primary sensory areas would process sensory stimuli under the assumption that they come from different sources, whereas the posterior IPS (IPS1-2) under the assumption that they come from a common source (i.e., *forced fusion*). It is only at the level of the anterior IPS (IPS3-4) that common and independent sources assumptions would be combined,

weighted by their posterior probability as predicted by the Bayesian Causal Inference (BCI) account (Körding et al, 2007).

To complement neuroimaging results which can only offer correlational information about the involvement of a brain region in a given task, transcranial magnetic stimulation (TMS) is often employed to draw conclusions on the causal brain region-task causal relationship. Previous neurostimulation studies investigated the causal role of association cortices, especially PPC, and of temporoparietal cortex, in multisensory processes. While some studies demonstrated that PPC is indeed causally involved in multisensory processes (Ro et al., 2004; Chambers et al., 2007; Pasalar et al., 2010; Kamke et al., 2012), others concluded that only temporoparietal regions are causally involved in such processes, whereas PPC activations found in fMRI studies would only constitute an epiphenomenon (Bolognini et al., 2009; Bertini et al., 2010). In light of these contrasting results, the causal and critical involvement of PPC in multisensory processes remains to date unclear.

1.2.1.3 Multisensory effects on behaviour

The study of behavioural effects of presentation of stimuli in multiple sensory modalities dates back to at least the beginning of the 20th century (Todd, 1912). Facilitation at the behavioural level is observed when targets in different sensory modalities are presented simultaneously, at the same location or are semantically congruent (Radeau, 1994; Vroomen & de Gelder, 2000; Frassinetti et al., 2002; Teder-Sälejärvi et al., 2005). For instance, target stimuli in one sensory modality presented together with a simultaneous stimulus in another sensory modality are detected faster compared to the target stimuli presented alone (i.e., *redundant target effect*: Gielen et al., 1983). To explain this phenomenon, the *race model approach* has suggested that each signal elicits a process in parallel to the others and the process which arrives first

to the response generator determines the response and its speed (Raab, 1962). Alternatively, the two processes could be combined and such integration could be responsible for the observed response facilitation (i.e., *coactivation model*, Miller, 1982).

Multisensory facilitation is less effective when stimuli are discrepant and perceptual illusions may instead be observed. The *McGurk effect* is the illusion of perceiving a certain syllable (e.g., DA) as the result of the combination of two other syllables presented in the visual and auditory modalities (e.g., GA and BA) (McGurk & MacDonald, 1976). Another type of illusion observed for spatially discrepant multisensory signals is the *ventriloquist effect*. When, for instance, a visual and an auditory stimuli are presented close in time but at different spatial locations and the observer is asked to localise only stimuli in one sensory modality, a displacement in space towards the other sensory modality can be observed (Bertelson, et al., 1998). Studies showed that the magnitude of this effect depends on stimulus disparity and on the relative reliability of the sensory stimuli, with visual stimuli generally exerting a stronger biasing effect on auditory stimuli (Charbonneau at al., 2013; Rohe & Noppeney, 2015a).

In summary, behavioural facilitation or perceptual illusions are observed for spatially congruent or incongruent signals respectively. Previous studies aimed to further characterise such multisensory processes by investigating whether they are affected by attention and expectation mechanisms.

1.2.2 The interplay between multisensory processes, attention and expectation

1.2.2.1 Attention and multisensory integration

In a multisensory environment, attention often selects relevant information across several sensory signals. The question arises as to whether mechanisms of

attention are modality-specific (i.e., there are different attentional resources for different sensory modalities) or if orienting of attention is supramodal (i.e., attentional resources are shared across modalities). The pioneering work of Spence & Driver (1996) played a critical role in the investigation of behavioural mechanisms of multisensory attention. Using a Posner task, they presented participants with auditory and visual stimuli in the left or the right hemifield and asked them to judge their elevation. For one sensory modality only (i.e., primary modality), they instructed participants to shift their endogenous spatial attention to the hemifield where the probability for a target in the primary modality to appear was higher. They observed that participants were faster for targets presented in the attended hemifield in the primary modality but also for targets in the other sensory modality (i.e., secondary modality), although spatial attention was not manipulated in this latter modality. This effect could be observed for audition as primary modality and vision as secondary modality but also vice-versa. These results indicate that the effects of spatial attention are not modality-specific but synergistic across modalities. The effect in the secondary modality was nevertheless smaller than that in the primary modality, indicating that the attention system is not fully supramodal. Therefore, the authors concluded that auditory and visual attention resources are 'separable-but-linked'. Similar behavioural results were later found for other combinations of sensory modalities (Spence et al., 2000; Lloyd et al., 2003).

From the neural perspective, following up results investigating the neural correlates of unisensory spatial attention, previous studies showed that the frontoparietal attentional networks are recruited regardless of sensory modality. Nevertheless, for the auditory modality the distinction between dorsal and ventral attention system is less clear than in the visual modality (Santangelo & Macaluso,

2012). To date, a conspicuous number of studies (many of them again employing the Posner task) showed that similar attentional networks are involved in spatial orienting of attention in different sensory modalities (e.g., Macaluso et al., 2000a, 2003; Wu et al., 2007). Moreover, previous research demonstrated that spatial attention modulates also unisensory areas regardless of stimulus modality (Macaluso et al., 2000b, 2002). All these studies bring data in support of the hypothesis that the attentional system is to some degree supramodal and that spatial proximity plays an important role in how attentional resources are shared across sensory modalities.

However, whether multisensory processes depend on (endogenous or exogenous) attention has been matter of debate. In fact, multisensory integration has been shown to take place in some cases preattentively (Bertelson, et al., 2000a, 2000b, Vroomen et al., 2001, Odegaard et al., 2015) and in other cases in function of attention (van Ee et al., 2009; Santangelo & Macaluso, 2012; Michail & Keil, 2018). In support of the hypothesis that attention and multisensory processes are interconnected, fMRI studies found enhancements of sensory activation for bimodal versus unimodal stimulation when the location was attended than unattended (Busse et al. 2005, Ciaramitaro et al., 2007) and larger bottom-up attentional capture for multisensory stimuli (Van der Burg et al., 2008). Furthermore, multisensory integration processes were shown to be influenced by modality-specific attention (Oruc et al., 2008, Vercillo & Gori, 2015) and parietal cortices were shown to integrate signals weighted by task-relevance (i.e., modality-specific attention) (Rohe & Noppeney, 2016). Overall, all these findings support the hypothesis that attention and multisensory processes interact in multifaceted ways (Talsma et al., 2010).

1.2.2.2 Expectation and multisensory integration

Only few studies to date investigated the effects of expectation on multisensory integration, thus it remains unclear whether expectation is a modality-specific or supramodal process. Expectation manipulated as audio-visual congruency or spatial/temporal prediction has been generally shown to interact with multisensory processes. For instance the McGurk illusion is more likely to be observed if participants expect the auditory and the visual signals to be incongruent, with inferior frontal sulcus arbitrating between integration and segregation based on prior expectations (Gau & Noppeney, 2016); auditory processing is influenced by visual signals predicting the onset of an auditory event (Stekelenburg & Vroomen, 2007) or the location of such event (Stekelenburg & Vroomen, 2012), as evidenced by N1 suppression as a measure of lower computational demand in a multisensory context; moreover, temporal predictability increases in multisensory contexts (Ball et al, 2018) and, if targets are presented at the expected time, responses are speeded and ERP enhanced regardless of stimulus modality (Lange & Röder, 2006). However, it should be noted that spatial/temporal expectation may induce attentional orienting which in fact could be responsible of such facilitation. Importantly, modality-specific effects of expectation have also been observed. Temporal expectation increases the perceptual weight of signal in the visual but not the auditory modality (Menceloglu et al., 2017). Additionally, if targets are expected in a specific modality, performance is reduced for targets in another modality (Spence et al., 2001).

In conclusion, different results were obtained for different types of expectation manipulation. Moreover, previous research suggested that manipulation of expectation often engages attentional mechanisms, making attention and expectation processes difficult to disentangle under the hypothesis that they are in fact separable.

1.2.3 Second and third aim of this thesis

The second aim of this thesis is to investigate whether spatial attention and expectation similarly engage multisensory brain systems and whether they each have similar effects on behaviour regardless of stimulus modality. In chapter 4 we qualitatively evaluate modality-specific and crossmodal effects of the behavioural interplay of spatial attention and expectation (i.e., response probability) in audition and vision and quantitatively measure differences attributed to sensory specificity. In chapter 5 we explore the multisensory effects of spatial attention and expectation at the neural level by comparing their effects for auditory and visual signals. fMRI studies only provide correlational information about the involvement of association areas in multisensory integration and segregation. The third aim of this thesis is to specifically investigate the causal role of the right anterior intraparietal sulcus in multisensory processes in function of modality specific attention. Thus, in chapter 6 we employ TMS to alter the excitability of IPS and evaluate behavioural effects of such manipulation.

REFERENCES

Alink, A., Schwiedrzik, C. M., Kohler, A., Singer, W., & Muckli, L. (2010). Stimulus predictability reduces responses in primary visual cortex. *The Journal of Neuroscience*, *30*(8), 2960–6. http://doi.org/10.1523/JNEUROSCI.3730-10.2010

Arnal, L. H., & Giraud, A.-L. L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, *16*(7), 390–398. http://doi.org/10.1016/j.tics.2012.05.003

Ball, F., Michels, L. E., Thiele, C., & Noesselt, T. (2018). The role of multisensory interplay in enabling temporal expectations. *Cognition*, *170*(September 2017), 130–146. http://doi.org/10.1016/j.cognition.2017.09.015

Beck, D. M., & Kastner, S. (2005). Stimulus context modulates competition in human extrastriate cortex. *Nature Neuroscience*, *8*(8), 1110–1116. http://doi.org/10.1038/nn1501

Bertelson, P., & Ascherleben, G. (1998). Automatic visual bias of perceived auditory location PAUL. *Psychonomic Bulletin & Review*, *5*(3), 482–489.

Bertelson, P., Pavani, F., Ladavas, E., Vroomen, J., & De Gelder, B. (2000). Ventriloquism in patients with unilateral visual neglect. *Neuropsychologia*, *38*(12), 1634–1642. http://doi.org/10.1016/S0028-3932(00)00067-1

Bertelson, P., Vroomen, J., De Gelder, B., & Driver, J. (2000). The ventriloquist effect does not depend on the direction of deliberate visual attention. *Perception & Psychophysics*, *62*(2), 321–332.

Bertini, C., Leo, F., Avenanti, A., & Làdavas, E. (2010). Independent mechanisms for ventriloquism and multisensory integration as revealed by theta-burst stimulation. *European Journal of Neuroscience*, *31*(10), 1791–1799. http://doi.org/10.1111/j.1460-9568.2010.07200.x

Bolognini, N., Miniussi, C., Savazzi, S., Bricolo, E., & Maravita, A. (2009). TMS modulation of visual and auditory processing in the posterior parietal cortex. *Experimental Brain Research*, *195*(4), 509–517. http://doi.org/10.1007/s00221-009-1820-7

Bremmer, F., Schlack, A., Shah, N. J., Zafiris, O., Kubischik, M., Hoffmann, K.-P., … Fink, G. R. (2001). Polymodal motion processing in posterior parietal and premotor cortex. *Neuron*, *29*, 287–296. http://doi.org/10.1016/S0896-6273(01)00198-2

Bressler, S. L., Tang, W., Sylvester, C. M., Shulman, G. L., & Corbetta, M. (2008). Top-down control of human visual cortex by frontal and parietal cortex in anticipatory visual spatial attention. *J.Neurosci.*, *28*(1529–2401 (Electronic)), 10056–10061. http://doi.org/10.1523/JNEUROSCI.1776-08.2008

Busse, L., Roberts, K. C., Crist, R. E., Weissman, D. H., & Woldorff, M. G. (2005). The spread of attention across modalities and space in a multisensory object. *Proceedings of the National Academy of Sciences*, *102*(51), 18751–18756. http://doi.org/10.1073/pnas.0507704102

Calvert, G. A., Brammer, M. J., Bullmore, E. T., Campbell, R., Iversen, S. D., & David, A. S. (1999). Response amplification in sensory-specific cortices during crossmodal binding. *NeuroReport*, *10*(12), 2619–2623. http://doi.org/10.1097/00001756-199908200-00033

Calvert, G. A., Hansen, P. C., Iversen, S. D., & Brammer, M. J. (2001). Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. *NeuroImage*, *14*(2), 427–438. http://doi.org/10.1006/nimg.2001.0812

Calvert, G. A., & Thesen, T. (2004). Multisensory integration: methodological approaches and emerging principles in the human brain. *Journal of Physiology-Paris*, *98*(1–3), 191–205. http://doi.org/10.1016/j.jphysparis.2004.03.018

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, *51*(13), 1484–1525. http://doi.org/10.1016/j.visres.2011.04.012

Chambers, C. D., Payne, J. M., & Mattingley, J. B. (2007). Parietal disruption impairs reflexive spatial attention within and between sensory modalities. *Neuropsychologia*, *45*(8), 1715–1724. http://doi.org/10.1016/j.neuropsychologia.2007.01.001

Charbonneau, G., Veronneau, M., Boudrias-Fournier, C., Lepore, F., & Collignon, O. (2013). The ventriloquist in periphery: Impact of eccentricity-related reliability on audio-visual localization. *Journal of Vision*, *13*(12), 20–20. http://doi.org/10.1167/13.12.20

Ciaramitaro, V. M., Buracas, G. T., & Boynton, G. M. (2007). Spatial and cross-modal attention alter responses to unattended sensory information in early visual and auditory human cortex. *Journal of Neurophysiology*, *98*(4), 2399–2413. http://doi.org/10.1152/jn.00580.2007

Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., & Shulman, G. L. (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature Neurosci.*, *3*(3), 292–296. http://doi.org/10.1038/73009

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron*, *58*(3), 306–324. http://doi.org/10.1016/j.neuron.2008.04.017

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews. Neuroscience*, *3*(3), 201–215. http://doi.org/10.1038/nrn755

Corbetta, M., & Shulman, G. L. (2011). *Spatial neglect and attention networks. Annual review of neuroscience* (Vol. 34). http://doi.org/10.1146/annurev-neuro-061010-113731

Corbetta M, Kincade MJ, Lewis C, Snyder AZ, A., S., Corbetta, M., … Sapir, A. (2005). Neural basis and recovery of spatial attention deficits in spatial neglect. *Nat Neurosci.*, *8*(11), 1603–1610. http://doi.org/10.1038/nn1574

Coull, J. T., & Nobre, A. C. (1998). Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *Journal of Neuroscience*, *18*(18), 7426–7435. http://doi.org/0270-6474/98/187426-10$05.00/0

Den Ouden, H. E. M., Friston, K. J., Daw, N. D., McIntosh, A. R., & Stephan, K. E. (2009). A dual role for prediction error in associative learning. *Cerebral Cortex*, *19*(5), 1175–1185. http://doi.org/10.1093/cercor/bhn161

Desimone, R., & Duncan, J. S. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*(1), 193–222. http://doi.org/10.1146/annurev.ne.18.030195.001205

Doherty, J. R., Rao, A., Mesulam, M. M., & Nobre, A. C. (2005). Synergistic effect of combined temporal and spatial expectations on visual attention. *The Journal of Neuroscience*, *25*(36), 8259–8266. http://doi.org/10.1523/JNEUROSCI.1821-05.2005

Eckert, M. A., Kamdar, N. V, Chang, C. E., Beckmann, C. F., Greicius, M. D., & Menon, V. (2008). A cross-modal system linking primary auditory and visual cortices: evidence from intrinsic fMRI connectivity analysis. *Human Brain Mapping*, *29*(7), 848–857. http://doi.org/10.1002/hbm.20560.A

Falchier, A., Clavagnier, S., Barone, P., & Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *22*(13), 5749–59. http://doi.org/20026562

Frassinetti, F., Bolognini, N., & Làdavas, E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Experimental Brain Research*, *147*(3), 332–343. http://doi.org/10.1007/s00221-002-1262-y

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360*(1456), 815–36. http://doi.org/10.1098/rstb.2005.1622

Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends in Cognitive Sciences*, *13*(7), 293–301. http://doi.org/10.1016/j.tics.2009.04.005

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews. Neuroscience*, *11*(2), 127–138. http://doi.org/10.1038/nrn2787

Gau, R., & Noppeney, U. (2016). How prior expectations shape multisensory perception. *NeuroImage*, *124*, 876–886. http://doi.org/10.1016/j.neuroimage.2015.09.045

Geng, J. J., & Behrmann, M. (2002). Probability cuing of target location facilitates visual search implicitly in normal participants and patients with hemispatial neglect. *Psychological Science*, *13*(6), 520–525. http://doi.org/10.1111/1467-9280.00491

Geng, J. J., & Behrmann, M. (2005). Spatial probability as an attentional cue in visual search. *Perception & Psychophysics*, *67*(7), 1252–1268. http://doi.org/10.3758/BF03193557

Geng, J. J., & Vossel, S. (2013). Re-evaluating the role of TPJ in attentional control: Contextual updating? *Neuroscience and Biobehavioral Reviews*, *37*(10), 2608–2620. http://doi.org/10.1016/j.neubiorev.2013.08.010

Ghazanfar, A., & Schroeder, C. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, *10*(6), 278–285. http://doi.org/10.1016/j.tics.2006.04.008

Gielen, S. C. A. M., Schmidt, R. A., & Van Den Heuvel, P. J. M. (1983). On the nature of intersensory facilitation of reaction time. *Perception &Psychophysics*, *34*(2), 161–168. http://doi.org/10.1016/0001-6918(84)90040-4

Grady, C. L., Van Meter, J. W., Maisog, J. M., Pietrini, P., Krasuski, J., & Rauschecker, J. P. (1997). Attention-related modulation of activity in primary and secondary auditory cortex. *Neuroreport*, *8*(11), 2511–2516.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*(1), 14–23. http://doi.org/10.1016/j.tics.2005.11.006

Händel, B. F., Haarmeier, T., & Jensen, O. (2011). Alpha oscillations correlate with the successful inhibition of unattended stimuli. *Journal of Cognitive Neuroscience*, *23*(9), 2494–2502. http://doi.org/10/fv25rd

Hayward, D. A., & Ristic, J. (2013). Measuring attention using the Posner cuing paradigm: the role of across and within trial target probabilities. *Frontiers in Human Neuroscience*, *7*, 1–11. http://doi.org/10.3389/fnhum.2013.00205

Hopfinger, J. B., Buonocore, M. H., & Mangun, G. R. (2000). The neural mechanisms of top-down attentional control. *Nature Neuroscience*, *3*(3), 284–291. http://doi.org/10.1038/72999

Jensen, O., Gips, B., Bergmann, T. O., & Bonnefond, M. (2014). Temporal coding organized by coupled alpha and gamma oscillations prioritize visual processing. *Trends in Neurosciences*, *37*(7), 357–369. http://doi.org/10.1016/j.tins.2014.04.001

Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Frontiers in Human Neuroscience*, *4*(November), 1–8. http://doi.org/10.3389/fnhum.2010.00186

Johnson, J. A., & Zatorre, R. J. (2005). Attention to simultaneous unrelated auditory and visual events: Behavioral and neural correlates. *Cerebral Cortex*, *15*(10), 1609–1620. http://doi.org/10.1093/cercor/bhi039

Kamke, M. R., Vieth, H. E., Cottrell, D., & Mattingley, J. B. (2012). Parietal disruption alters audiovisual binding in the sound-induced flash illusion. *NeuroImage*, *62*(3), 1334–1341. http://doi.org/10.1016/j.neuroimage.2012.05.063

Kastner, S., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1998). Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. *Science*, *282*(5386), 108–111.

Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron*, *22*, 751–761.

Kastner, S., & Ungerleider, L. G. (2000). Mechanisms of visual attention in the Human Cortex. *Annual Review of Neuroscience*, *23*, 315–341.

Kincade, J. M., Abrams, R. A., Astafiev, S. V., Shulman, G. L., & Corbetta, M. (2005). An event-related functional magnetic resonance imaging study of voluntary and stimulus-driven orienting of attention. *Journal of Neuroscience*, *25*(18), 4593–4604. http://doi.org/10.1523/JNEUROSCI.0236-05.2005

Klemen, J., & Chambers, C. D. (2012). Current perspectives and methods in studying neural mechanisms of multisensory interactions. *Neuroscience and Biobehavioral Reviews*, *36*(1), 111–133. http://doi.org/10.1016/j.neubiorev.2011.04.015

Kok, P., Jehee, J. F. M. F. M., & de Lange, F. P. (2012). Less is more: expectation sharpens representations in the primary visual cortex. *Neuron*, *75*(2), 265–270. http://doi.org/10.1016/j.neuron.2012.04.034

Kok, P., Rahnev, D., Jehee, J. F. M., Lau, H. C., & de Lange, F. P. (2012). Attention reverses the effect of prediction in silencing sensory signals. *Cerebral Cortex*, *22*(9), 2197–2206. http://doi.org/10.1093/cercor/bhr310

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal inference in multisensory perception. *PLoS ONE*, *2*(9). http://doi.org/10.1371/journal.pone.0000943

Lange, K., & Röder, B. (2006). Orienting attention to points in time improves stimulus processing both within and across modalities. *Journal of Cognitive Neuroscience*, *18*(5), 715–729. http://doi.org/10.1162/jocn.2006.18.5.715

Langner, R., Kellermann, T., Boers, F., Sturm, W., Willmes, K., & Eickhoff, S. B. (2011). Modality-specific perceptual expectations selectively modulate baseline activity in auditory, somatosensory, and visual cortices. *Cerebral Cortex*, *21*(12), 2850–2862. http://doi.org/10.1093/cercor/bhr083

Laurienti, P. J., Burdette, J. H., Wallace, M. T., Yen, Y.-F., Field, A. S., & Stein, B. E. (2002). Deactivation of sensory-specific cortex by cross-modal stimuli. *Journal of Cognitive Neuroscience*, *14*(3), 420–429. http://doi.org/10.1162/089892902317361930

Lee, H., & Noppeney, U. (2014). Temporal prediction errors in visual and auditory cortices. *Current Biology*, *24*(8), R309–R310. http://doi.org/10.1016/j.cub.2014.02.007

Leitão, J., Thielscher, A., Werner, S., Pohmann, R., & Noppeney, U. (2013). Effects of parietal TMS on visual and auditory processing at the primary cortical level-a

concurrent TMS-fMRI study. *Cerebral Cortex*, *23*(4), 873–884. http://doi.org/10.1093/cercor/bhs078

Liu, T., Larsson, J., & Carrasco, M. (2007). Feature-based attention modulates orientation-selective responses in human visual cortex. *Neuron*, *55*(2), 313–323. http://doi.org/10.1016/j.neuron.2007.06.030.Feature-Based

Lloyd, D. M., Merat, N., Francis, M., & Spence, C. (2003). Crossmodal links between audition and touch in covert endogenous spatial attention. *Perception & Psychophysics*, *65*(6), 901–924. http://doi.org/10.1016/S1366-5545(02)00012-1

Macaluso, E., & Doricchi, F. (2013). Attention and predictions: control of spatial attention beyond the endogenous-exogenous dichotomy. *Frontiers in Human Neuroscience*, *7*(October), 685. http://doi.org/10.3389/fnhum.2013.00685

Macaluso, E., Frith, C. D., & Driver, J. (2002). Crossmodal spatial influences of touch on extrastriate visual areas take current gaze direction into account. *Neuron*, *34*(4), 647–658. http://doi.org/10.1016/S0896-6273(02)00678-5

Macaluso, E., Frith, C., & Driver, J. (2000). Modulation of human visual cortex by crossmodal spatial attention. *Science*, *289*(2000), 1206–1208. http://doi.org/10.1126/science.289.5482.1206

Macaluso, E., Frith, C., & Driver, J. (2000). Selective spatial attention in vision and touch: unimodal and multimodal mechanisms revealed by PET. *Journal of Neurophysiology*, *83*(5), 3062–3075.

Mazaheri, A., van Schouwenburg, M. R., Dimitrijevic, A., Denys, D., Cools, R., & Jensen, O. (2014). Region-specific modulations in oscillatory alpha activity serve to facilitate processing in the visual and auditory modalities. *NeuroImage*, *87*, 356–362. http://doi.org/10.1016/j.neuroimage.2013.10.052

Menceloglu, M., Grabowecky, M., & Suzuki, S. (2017). Temporal expectation weights visual signals over auditory signals. *Psychonomic Bulletin and Review*, *24*(2), 416–422. http://doi.org/10.3758/s13423-016-1109-3

Michail, G., & Keil, J. (2018). High attentional load enhances the susceptibility to non-speech audiovisual illusions. *Scientific Reports*, 1–11. http://doi.org/10.1038/s41598-018-30007-6

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, *14*(2), 247–279. http://doi.org/10.1016/0010-0285(82)90010-X

Murray, S. O., Kersten, D., Olshausen, B. A., Schrater, P., & Woods, D. L. (2002). Shape perception reduces activity in human primary visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(23), 15164–9. http://doi.org/10.1073/pnas.192579399

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, *118*(12), 2544–2590. http://doi.org/10.1016/j.clinph.2007.04.026

Nobre, A. C., Gitelman, D. R., Dias, E. C., & Mesulam, M. M. (2000). Covert visual spatial orienting and saccades: overlapping neural systems. *NeuroImage*, *11*(3), 210–216. http://doi.org/10.1006/nimg.2000.0539

Nobre, A. C., & van Ede, F. (2017). Anticipated moments: temporal structure in attention. *Nature Reviews Neuroscience*, *19*(1), 34–48. http://doi.org/10.1038/nrn.2017.141

Odegaard, B., Wozny, D. R., & Shams, L. (2015). Biases in visual, auditory, and audiovisual perception of space. *PLoS Computational Biology*, *11*(12), 1–23. http://doi.org/10.1371/journal.pcbi.1004649

Odegaard, B., Wozny, D. R., & Shams, L. (2016). The effects of selective and divided attention on sensory precision and integration. *Neuroscience Letters*, *614*, 24–28. http://doi.org/10.1016/j.neulet.2015.12.039

Oruc, I., Sinnett, S., Bischof, W. F., Soto-Faraco, S., Lock, K., & Kingstone, A. (2008). The effect of attention on the illusory capture of motion in bimodal stimuli. *Brain Research*, *1242*, 200–208. http://doi.org/10.1016/j.brainres.2008.04.014

Pasalar, S., Ro, T., & Beauchamp, M. S. (2010). TMS of posterior parietal cortex disrupts visual tactile multisensory integration. *European Journal of Neuroscience*, *31*(10), 1783–1790. http://doi.org/10.1111/j.1460-9568.2010.07193.x

Posner, M. I. (1980). Orienting of attention. *Q J Exp Psychol*, *32*(1), 3–25. http://doi.org/10.1080/00335558008248231

Raab, D. H. (1962). Statistical facilitation of simple reaction-times. *Transactions of the New York Academy of Sciences*, *24*(5), 574–590.

Ro, T., Wallace, R., Hagedorn, J., Farné, A., & Pienkos, E. (2004). Visual enhancing of tactile perception in the posterior parietal cortex. *Journal of Cognitive Neuroscience*, *16*(1), 24–30. http://doi.org/10.1162/089892904322755520

Rohe, T., & Noppeney, U. (2015a). Cortical hierarchies perform Bayesian Causal Inference in multisensory perception. *PLOS Biology*, *13*(2), e1002073. http://doi.org/10.1371/journal.pbio.1002073

Rohe, T., & Noppeney, U. (2015b). Sensory reliability shapes Bayesian Causal Inference in perception via two mechanisms. *Journal of Vision*, *15*, 1–38. http://doi.org/10.1167/15.5.22.doi

Rohe, T., & Noppeney, U. (2016). Distinct computational principles govern multisensory integration in primary sensory and association cortices. *Current Biology*, *26*(4), 509–514. http://doi.org/10.1016/j.cub.2015.12.056

Rohenkohl, G., Gould, I. C., Pessoa, J., & Nobre, A. C. (2014). Combining spatial and temporal expectations to improve visual perception. *Journal of Vision*, *14*(4), 1–13. http://doi.org/10.1167/14.4.8

Santangelo, V., & Macaluso, E. (2012). Spatial attention and audiovisual processing. *The New Handbook of Multisensory Processing*, 359–370.

Schroeder, C. E., John, S., Kaiming, F., Tammy, M., O'Connell, M., & Hackett, T. (2009). Anatomical mechanisms and functional implications of multisensory convergence in early cortical processing Charles. *International Journal of Psychophysiology*, *258*(1–2), 235–246. http://doi.org/10.1016/S0167-8760

Serences, J. T., Saproo, S., Scolari, M., Ho, T., & Muftuler, L. T. (2009). Estimating the influence of attention on population codes in human visual cortex using voxel-based tuning functions. *NeuroImage*, *44*(1), 223–231. http://doi.org/10.1016/j.neuroimage.2008.07.043

Sheliga, B. M., Riggio, L., & Rizzolatti, G. (1994). Orienting of attention and eye movements. *Experimental Brain Research*, *98*, 507–522. http://doi.org/10.1007/BF00233988

Shomstein, S., & Yantis, S. (2006). Parietal cortex mediates voluntary control of spatial and nonspatial auditory attention. *Journal of Neuroscience*, *26*(2), 435–439. http://doi.org/10.1523/JNEUROSCI.4408-05.2006

Shulman, G. L., Astafiev, S. V, Franke, D., Pope, D. L. W., Abraham, Z., Mcavoy, M. P., & Corbetta, M. (2009). Interaction of stimulus-driven reorienting and expectation in ventral and dorsal fronto-parietal and basal ganglia-cortical networks. *Journal of Neuroscience*, *29*(14), 4392–4407. http://doi.org/10.1523/JNEUROSCI.5609-08.2009

Shulman, G. L., Pope, D. L. W., Astafiev, S. V., McAvoy, M. P., Snyder, A. Z., & Corbetta, M. (2010). Right hemisphere dominance during spatial selective attention and target detection occurs outside the dorsal frontoparietal network. *Journal of Neuroscience*, *30*(10), 3640–3651. http://doi.org/10.1523/JNEUROSCI.4085-09.2010

Silver, M. A., Ress, D., & Heeger, D. J. (2007). Neural correlates of sustained spatial attention in human early visual cortex. *Journal of Neurophysiology*, *97*, 229–237. http://doi.org/10.1152/jn.00677.2006.

Smith, D. V, Davis, B., Niu, K., Healy, E. W., Bonilha, L., Fridriksson, J., … Rorden, C. (2010). Spatial attention evokes similar activation patterns for visual and auditory stimuli. *Journal of Cognitive Neuroscience*, *22*(2), 347–61. http://doi.org/10.1162/jocn.2009.21241

Somers, D. C., Dale, A. M., Seiffert, A. E., & Tootell, R. B. (1999). Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(4), 1663–8. http://doi.org/10.1073/pnas.96.4.1663

Soto, D., & Blanco, M. J. (2004). Spatial attention and object-based attention: A comparison within a single task. *Vision Research*, *44*(1), 69–81. http://doi.org/10.1016/j.visres.2003.08.013

Spence, C., & Driver, J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(4), 1005–1030. http://doi.org/10.1037/0096-1523.22.4.1005

Spence, C., Nicholls, M. E., & Driver, J. (2001). The cost of expecting events in the wrong sensory modality. *Perception & Psychophysics*, *63*(2), 330–336. http://doi.org/10.3758/BF03194473

Spence, C., Pavani, F., & Driver, J. (2000). Crossmodal links between vision and touch in covert endogenous spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(4), 1298–1319. http://doi.org/10.3758/BF03194823

Stein, B. E. (1998). Neural mechanisms for synthesizing sensory information and producing adaptive behaviors. *Experimental Brain Research*, *123*(1–2), 124–135. http://doi.org/10.1007/s002210050553

Stekelenburg, J. J., & Vroomen, J. (2007). Neural correlates of multisensory integration. *Journal of Cognitive Neuroscience*, *19*(12), 1964–1973.

Stekelenburg, J. J., & Vroomen, J. (2012). Electrophysiological correlates of predictive coding of auditory location in the perception of natural audiovisual events. *Frontiers in Integrative Neuroscience*, *6*(May), 1–7. http://doi.org/10.3389/fnint.2012.00026

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*, *13*(9), 403–409. http://doi.org/10.1016/j.tics.2009.06.003

Summerfield, C., Egner, T., Greene, M., Koechlin, E., Mangels, J., & Hirsch, J. (2006). Predictive codes for forthcoming perception in the frontal cortex. *Science*, *314*(5803), 1311–1314.

Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences*, *14*(9), 400–410. http://doi.org/10.1016/j.tics.2010.06.008

Teder-Sälejärvi, W. A., Di Russo, F., McDonald, J. J., & Hillyard, S. A. (2005). Effects of spatial congruity on audio-visual multimodal integration. *Journal of Cognitive Neuroscience*, *17*(9), 1396–1409. http://doi.org/10.1162/0898929054985383

Thut, G., Nietzel, A., Brandt, S. A., & Pascual-Leone, A. (2006). Alpha-Band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *Journal of Neuroscience*, *26*(37), 9494–9502. http://doi.org/10.1523/JNEUROSCI.0875-06.2006

Todd, J. W. (1912). *Reaction to multiple stimuli*. New York: Science Press.

Todorovic, A., van Ede, F., Maris, E., & de Lange, F. P. (2011). Prior expectation mediates neural adaptation to repeated sounds in the auditory cortex: an MEG study. *Journal of Neuroscience*, *31*(25), 9118–9123. http://doi.org/10.1523/JNEUROSCI.1425-11.2011

Tzourio, N., El Massioui, F., Crivello, F., Joliot, M., Renault, B., & Mazoyer, B. (1997). Functional anatomy of human auditory attention studied with PET. *NeuroImage*, *5*(1), 63–77. http://doi.org/10.1006/nimg.1996.0252

Van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W., & Theeuwes, J. (2008). Pip and pop: nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(5), 1053–1065. http://doi.org/10.1037/0096-1523.34.5.1053

van Ede, F., de Lange, F. P., & Maris, E. (2012). Attentional cues affect accuracy and reaction time via different cognitive and neural processes. *Journal of Neuroscience*, *32*(30), 10408–10412. http://doi.org/10.1523/JNEUROSCI.1337-12.2012

van Ee, R., van Boxtel, J. J. A., Parker, A. L., & Alais, D. (2009). Multisensory congruency as a mechanism for attentional control over perceptual selection. *Journal of Neuroscience*, *29*(37), 11641–11649. http://doi.org/10.1523/JNEUROSCI.0873-09.2009

Vercillo, T., & Gori, M. (2015). Attention to sound improves auditory reliability in audio-tactile spatial optimal integration. *Frontiers in Integrative Neuroscience*, *9*, 1–8. http://doi.org/10.3389/fnint.2015.00034

Vroomen, J., Bertelson, P., & de Gelder, B. (2001). The ventriloquist effect does not depend on the direction of automatic visual attention. *Perception & Psychophysics*, *63*(4), 651–659. http://doi.org/10.3758/BF03194427

Vroomen, J., & De Gelder, B. (2000). Sound enhances visual perception: Cross-modal effects of auditory organization on vision. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(5), 1583–1590. http://doi.org/10.1037/0096-1523.26.5.1583

Watanabe, T., Sasaki, Y., Miyauchi, S., Putz, B., Fujimaki, N., Nielsen, M., … Miyakawa, S. (1998). Attention-regulated activity in human primary visual cortex. *Journal of Neurophysiology*, *79*, 2218–2221.

Weisz, N., Hartmann, T., Müller, N., Lorenz, I., & Obleser, J. (2011). Alpha rhythms in audition: Cognitive and clinical perspectives. *Frontiers in Psychology*, *2*(APR), 1–15. http://doi.org/10.3389/fpsyg.2011.00073

Werner, S., & Noppeney, U. (2011). The contributions of transient and sustained response codes to audiovisual integration. *Cerebral Cortex*, *21*(4), 920–931. http://doi.org/10.1093/cercor/bhq161

Worden, M. S., Foxe, J. J., Wang, N., & Simpson, G. V. (2000). Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. *The Journal of Neuroscience*, *20*(6), RC63. http://doi.org/Rc63

Wu, C. T., Weissman, D. H., Roberts, K. C., & Woldorff, M. G. (2007). The neural circuitry underlying the executive control of auditory spatial attention. *Brain Research*, *1134*(1), 187–198. http://doi.org/10.1016/j.brainres.2006.11.088

Wyart, V., Nobre, A. C., & Summerfield, C. (2012). Dissociable prior influences of signal probability and relevance on visual contrast sensitivity. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(9), 3593–3598. http://doi.org/10.1073/pnas.1120118109

Yoshiura, T., Zhong, J., Shibata, D. K., Kwok, W. E., Shrier, D. A., & Numaguchi, Y. (1999). Functional MRI study of auditory and visual oddball tasks. *Neuroreport*, *10*(8), 1683–1688. http://doi.org/10.1097/00001756-199906030-00011

Chapter 2:

Methodologies

2.1 ANALYSIS OF BEHAVIOURAL DATA

This chapter aims to give a brief overview of the methods used in this thesis to analyse behavioural data.

2.1.1 General linear model and model comparison

A general linear model (GLM) is broadly described as:

(1) $Y_i = \beta_0 + \beta_1 X_{i1}$

(Rutherford, 2013), where Y represents the dependent variable (e.g., response times), X represents the independent variable or predictor (e.g., direction of attention) and β represents the contribution of the independent variable to the dependent variable's prediction. Multiple predictors can enter the model (e.g., direction of attention, direction of expectation etc.). Importantly, the GLM assumes linearity with respect to the parameters β. This means that no parameter is divided or multiplied with another parameter and no parameter is above the first power (Rutherford, 2013). In a GLM analysis, the aim is to estimate β for each of the predictors which, in terms of linear regression, represents the slope of the regression line. Compared to linear regression, the GLM is more flexible as it introduces the possibility to have multiple dependent variables which are linearly combined. In the GLM an error term ε is included in the model to account for unexplained differences between the data and the model (Rutherford, 2013):

(2) $Y_i = \beta_0 + \beta_1 X_{i1} + \varepsilon_i$

where ε is the error term ($\varepsilon \sim N(0, \sigma^2 I)$, where I is the identity matrix).

A more flexible (but still linear) generalization employed when the error distribution is not normal or when linearity with respect to the predictors is violated is the Generalised Linear Model. In this case a link function relates the mean of each observation to its linear predictor (McCullagh and Nedler, 1989).

A linear model which includes both fixed and random predictors is called Linear Mixed Model. Fixed predictors are factors of interest for which all levels are included whereas random predictors represent factors which are randomly sampled from a larger population, thus all level of the factors are not included in the model and inferences at the population level has to be made. A Linear Mixed Model is defined as:

(3) $Y_i = \beta_1 X_{i1} + b_i Z_i + \varepsilon_i$

where X_i is the fixed predictor, Z_i the random predictor, β is the fixed effect parameter, $b_i \sim N(0, D)$ is the random effect parameter and $\varepsilon_i \sim N(0, \sigma^2 I)$ is the error term. D is the variance-covariance matrix of the multivariate normal distribution defined by the variance of each random effect and the covariances between each two random effects (West et al., 2014; Devore & Farnum, 2015).

GLM incorporates regression, t-test and analysis of variance (ANOVA). The t-test is used to compare different means obtained in different experimental conditions. For instance, in an independent samples t-test, the *t* value is calculated as follows (Devore & Farnum, 2015):

$$(4)\ t = \frac{(\bar{x}_1 - \bar{x}_2)}{S_{\bar{x}_1 - \bar{x}_2}}$$

where $\bar{x}_1$ and $\bar{x}_2$ are the sample means and $S_{\bar{x}_1 - \bar{x}_2}$ is the pooled standard error of the mean. This corresponds to testing whether the slope β of the regression line defined by the two sample means is different from 0. When the two sample means are not independent (e.g., they are from the same individual), a paired-sample t-test is appropriate. In this case the numerator in equation (4) is the sample mean of the differences and the denominator is the sample standard deviation of the differences (Devore & Farnum, 2015).

In the ANOVA, the *F* value is calculated as the ratio of the between-samples variation (i.e., explained variance) and the within-samples variation (i.e., unexplained variance), where the variation is expressed in terms of sum of squares, as follows:

$$(5)\ F = \frac{\frac{SS_R}{k-1}}{\frac{SS_E}{n-k}}$$

where SS_R represents the between-samples variation, or rather the variability between each sample and the mean of all samples, SS_E represents the within-samples variation, or rather the variability between the mean of all samples and the linear regression model. k is the number of groups and n is the sample size (Devore & Farnum, 2015). When the samples are not independent, a repeated measures ANOVA is appropriate. In this case the error term in the denominator (SS_E) of equation (5) is calculated by subtracting the subject variability from the within-samples variation. This reduces SS_E and increases the power of the test. To compare single main effects, a multiple comparison correction is applied to avoid inflation of type I error (i.e., to reject the null

hypothesis when it is true). For example, Bonferroni tests can be employed, by dividing the significance value by the number of comparisons performed (Frane, 2015).

Frequentist inference uses *F* and *t* values for hypothesis testing. A null (e.g., two group means are equal) and an alternative hypothesis (e.g., two group means are different) are formulated. The *F* value comes from an F distribution under the null hypothesis and the *t* value from a t distribution under the null hypothesis. The *p*-value is calculated as the probability of obtaining an *F* or *t* value equal to or more extreme than the one obtained, under the assumption that the null hypothesis is true (Schervish, 1996). Traditionally, a *p*-value equal to or smaller than 0.05 (i.e. a probability < 0.05) is used to reject the null hypothesis.

To measure the magnitude of the effect, the effect size is calculated. The effect size is defined as the degree to which the effect exists in the population (Cohen, 1988). For paired-samples t-tests, the Cohen's d (d_{av}) is calculated as the ratio of the difference of two means and the average standard deviation (Cohen, 1988). For one sample t-test, d_{av} is calculated as the ratio of the sample mean minus the test value (e.g., 0), and the sample standard deviation. For ANOVAs, the partial eta squared (η_p^2) is calculated as the ratio of the between-samples variation and the sum of the between-samples variation and the within-samples variation (Levine & Hullett, 2002; for a discussion on the interpretation of the effect size values, Lakens, 2013). In the null hypothesis significant testing (NHST) approach, the main aim is to disprove the null hypothesis but when the null hypothesis cannot be rejected, it is not possible to prove whether the data indeed favour the null hypothesis. To overcome this problem, the frequentist approach can be complemented with a Bayesian framework. This latter

approach is based on the Bayes theorem which, for parameter estimation, is defined as:

(6) $p(\theta \mid D, M_k) = \frac{p(D \mid \theta, M_k)p(\theta \mid M_k)}{p(D \mid M_k)}$ or posterior $= \frac{\text{likelihood x prior}}{\text{model evidence}}$

where θ are the parameters, M is the model and D is the data, (Lee & Wagenmakers, 2014). $p(\theta \mid D, M_k)$ is the *posterior*, i.e. the probability of a set of parameters θ of a model M_k given the observed data D. $p(D \mid \theta, M_k)$ is the *likelihood*: the probability of the data D given the set of parameters θ of a model M_k. $p(\theta \mid M_k)$ is the *prior*: the probability of the set of parameters θ of a model M_k. $p(D \mid M_k)$ is the *marginal likelihood* (or *model evidence*): the probability of observing the data D as modelled by M_k. The marginal likelihood is calculated as:

(7) $p(D \mid M_1) = \sum_{i=1}^{k} p(D \mid \theta_i, M_1)\, p(\theta_i \mid M_1)$

(Lee & Wagenmakers, 2014). To select one model over another (e.g., the null hypothesis vs the alternative hypothesis), the probability of the first model given the data is compared to the probability of the second model given the data, i.e., $p(M_1 \mid D)$ vs $p(M_2 \mid D)$. Following the Bayes theorem:

(8) $p(M_k \mid D) = \frac{p(D \mid M_k)\, p(M_k)}{p(D)}$

Assuming that $p(M_1)$ is equal to $p(M_0)$ then:

(9) $\frac{p(M_1 | D)}{p(M_0 | D)} = \frac{p(D | M_1)}{p(D | M_0)}$

where, in this example, M_1 is the alternative hypothesis and M_0 the null hypothesis and $p(D | M_1)$ and $p(D | M_0)$ are the marginal likelihood of the alternative hypothesis and that of the null hypothesis. The ratio of two marginal likelihoods of two different models is the Bayes Factor (Quintana & Williams, 2018). The Bayes factor is conventionally used to select the alternative or null model as follows: a Bayes factor larger than 3 is considered as moderate to very strong evidence in favour of the hypothesis under consideration (i.e., the one specified in the numerator) (Quintana & Williams, 2018).

To compare models with different factors (or even nested models), maximum likelihood estimation may be employed (i.e., parameters which maximise the likelihood function, or rather generate the given data with the highest probability, are estimated) and the information criterion is used to select one model over the others. The Akaike information criterion (AIC) and Bayesian information criterion (BIC) are formally defined as:

(10) AIC = $-2*\ln(\hat{L}) + 2*$number of free parameters

(11) BIC = $-2*\ln(\hat{L}) + \log($number of observations$)*$number of free parameters

The BIC approach is based on the Bayes theorem mentioned in (6) and:

(12) $\hat{L} = p(D | \hat{\theta}, M)$

where $\hat{\theta}$ are the parameters maximising the likelihood function (i.e., $\hat{L}$).

The model with the smallest AIC or BIC, meaning the model with the maximum likelihood (or log likelihood, ln) penalised by the number of free parameters, is selected as the best model and so are its parameters (Chakrabarti & Ghosh, 2011).

2.1.2 Signal detection theory

Signal detection theory (SDT) provides measures to describe and quantify observers' performance. This approach is used to analyse data of a task where a decision between two types of trials has to be performed. Each type of trial comes from one of two different distributions. One distribution is called *noise distribution*; the other is called *signal distribution* (Fig. 2.1). In e.g. a detection task, when a signal trial is responded to (i.e., *yes* response), a *hit* is produced; when it is not responded to (i.e., *no* response), a *miss* is produced. When a noise trial is responded to (i.e., *yes* response), a *false alarm* (FA) is produced; when it is not responded to (i.e., *no* response), a *correct rejection* (CR) is produced (see Table 2.1).

Table 2.1: Types of responses in SDT.

	Yes response	*No* response
Signal trial	Hit	Miss
Noise trial	False alarm	Correct rejection

It follows that:

a. Hit rate = Number of hits / number of signal trials

b. False alarm rate = Number of false alarms / number of noise trials

c. Miss rate = 1 - hit rate

d. Correct rejection rate = 1 – FA rate

The relationship between noise and signal distributions is the observer's *sensitivity*, or the *d-prime* (*d'*), and it is calculated as follows (Wickens, 2002):

$$(13)\ \mathrm{d'} = \mathrm{z(hit\ rate)} - \mathrm{z(false\ alarm\ rate)}$$

with z(·) being the inverse of the cumulative Gaussian distribution.

When d' is equal to zero, it means that the two distributions are identical (Fig. 2.1). Conversely, large d' values imply that the observer in more capable of discriminating signals from noise. Any propensity of the observer to respond yes or no (i.e., response bias, Fig. 2.1) is called *criterion* or λ_{center} and it is calculated relative to half the distance between signal and noise distributions' means, as follows (Wickens, 2002):

$$(14)\ \lambda_{\mathrm{center}} = -\frac{\mathrm{z(false\ alarm\ rate)} + \mathrm{z(hit\ rate)}}{2}$$

A $\lambda_{\mathrm{center}}$ equal to 0 means that the observer has no preference for signal or noise trials.

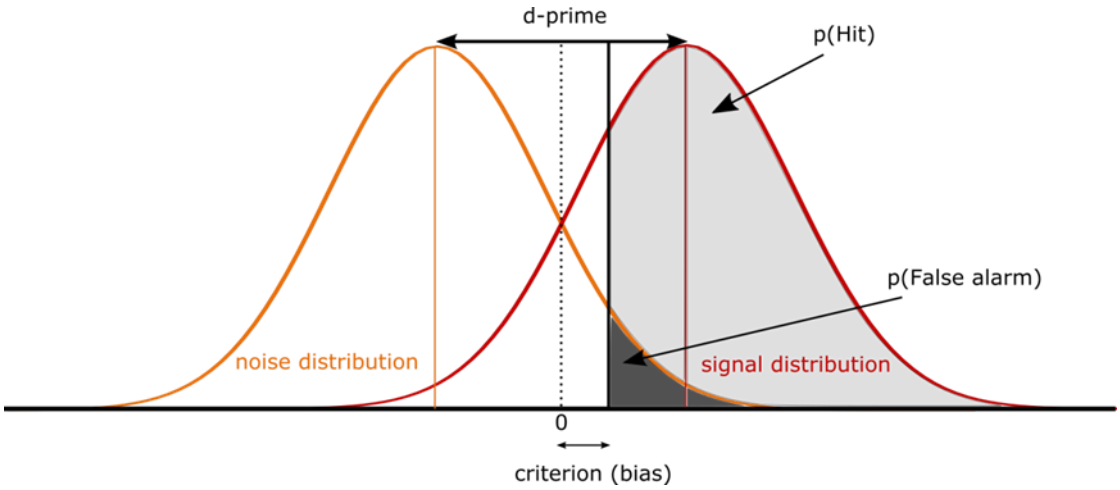


Figure 2.1 Illustration of noise distribution and signal distribution.

d' represents the distance between the two distributions' means and the criterion defines the bias of the observer towards one of the two distributions, relative to half of the two means' difference.

2.2 EYE-MOVEMENT ANALYSIS

2.2.1 Attention and eye-movements

Attention can be oriented with (*overt*) or without (*covert*) a change of eye fixation and head movement. To study the effects of attention independently from other systems, previous studies required participants to direct attention in a covert fashion. There is evidence, in fact, showing that attention and eye-movements are independent (Posner & Petersen, 1990). The idea that saccade-preparation is not involved in spatial attention has nevertheless been challenged by the *premotor theory of attention* (Rizzolatti et al., 1987) which argues that attention is in fact the consequence of eye-movement preparation (Sheliga et al., 1994) and thus, even in the absence of saccades (i.e., eye movements towards a certain location), the oculomotor system is involved (Van der Stigchel & Theeuwes, 2007, Hamker, 2005). Conventionally, if the aim is to investigate covert attention, participants are explicitly instructed to maintain fixation at a fixation point which is generally located directly in front of their eyes, 0° of visual angle along the azimuth and the vertical line (e.g., Spence & Driver, 1996).

2.2.2 Eye-tracking

In order to monitor participants' fixation, an eye-tracking device may be employed. Depending on the device used, fixation and events other than fixation (e.g. saccades and blinks) are determined using conceptually similar algorithms.

Salvucci & Goldberg (2000) proposed a series of algorithms to identify fixation events based on spatial and temporal information. When using a slow system (e.g., Tobii EyeX, Tobii, Sweden, ~60 Hz sampling rate) a dispersion-based approach is

preferable to a velocity-based approach as a fine-grained temporal profile of the eye-movements cannot be obtained. Dispersion-based algorithms define fixation as a cluster of samples which are close to one another within a radius threshold (i.e., area of interest) centred at a predefined x and y coordinates. Thus, fixation maintenance is based on the spread distance of consecutive samples (e.g. three consecutive samples, as suggested by ASL, Eyenal system) within a particular dispersion threshold. Blignaut (2009) suggested that a threshold radius between 0.7° and 1.3° ensures replicable results in terms of reliability of fixation definition. The algorithm consists in (1) collecting all the samples for each trial and (2) calculating the dispersion from the centre of fixation. If within one trial more than three consecutive samples are located outside the threshold radius, then fixation is not maintained (i.e., a saccade is produced). Importantly, blinks also incorporate saccades due to eyelid-closure and eyelid-opening events. The difference between a saccade and a blink is that a blink is followed by an absence of data samples (i.e., eyes are closed) for a duration of 50-500 ms (Caffier et al., 2003). Thus, to disentangle blinks from saccades the following approach can be used: a saccade followed by a period of no data between 50 and 500 ms is classified as blink, whereas a 'true' saccade is not followed by periods of no data.

When using a fast system (e.g. EyeLink, SR Research, 2007, 2000 Hz sampling rate), motion, velocity and acceleration based algorithms are employed to give a reliable classification of fixation, saccade and blink events. For instance, EyeLink classifies saccades based on the instantaneous computation of motion, velocity and acceleration to evaluate whether each data sample falls within the thresholds. Examples of thresholds are: saccade velocity of 30°/sec, saccade acceleration of 8000°/sec^2 and saccade motion of 0.15° (SR Research, 2007). Blinks are defined as samples where the pupil is missing. A dispersion approach may be employed on events

classified as ‘saccades’ by EyeLink when the aim of the analysis is to purely ensure that a good fixation is maintained.

2.3 FUNCTIONAL MAGNETIC RESONANCE IMAGING

Functional magnetic resonance imaging (fMRI) is a neuroimaging method first introduced in 1990 which is used to measure changes in the brain metabolism over time. Changes in the metabolic activity can represent changes in cognitive states due to e.g. task performance. For this reason, fMRI is widely used in cognitive neuroscience to indirectly measure brain activity (Glover et al. 2011).

2.3.1 Physics of MRI and fMRI

Magnetic resonance (MR) scanners use large electric current flowing into a loop to produce a strong static magnetic field (B_0). The hydrogen protons contained in the water of brain tissue spin about their axes and act like randomly oriented magnets. When a magnetic field is applied to the brain, the spin of hydrogen protons aligns parallel or in the opposite direction to B_0, which defines the longitudinal plane. The difference between the magnetization in the same and opposite direction to B_0 is called *net magnetization* and is the source of the MR signal. A radiofrequency (RF) coil is used to transmit energy to the protons with pulses delivered at the *Larmor* frequency. Larmor frequency is equal to a constant (42.6 MHz/T for hydrogen protons) times the main magnetic field strength. This oscillating/rotating magnetic field (B_1) interacts with the spinning of the protons, causing their precession. This energy makes the net magnetization rotate away from the longitudinal direction into the transverse plane. After the pulse, the net magnetization moves back into the original direction (Longitudinal or T1 relaxation) (Fig. 2.2A) and the magnetic signal is acquired by receiver coils. Tissues of different types contain protons with different T1 rate. T1 rate is defined as the time it takes for the longitudinal magnetization to move back to 63%

of its original value. After the RF pulse, protons start to dephase due to e.g. spin-spin interactions, magnetic field inhomogeneities, magnetic susceptibility or chemical shift effects (T2* dephasing) or exclusively to spin-spin interactions (T2 dephasing) (Fig. 2.2B). As for T1, T2 relaxation rate is different for different tissues and it is defined as the time it takes for the transverse magnetization to move back to 37% of its original value. T1 relaxation and T2 dephasing occur at the same time. After the 90° RF pulse, a 180° RF is applied. This makes the spins start rephrasing and this creates an echo (spin echo). Echo time (TE) is the time between two RF pulses' peaks and repetition time (TR) is the time of a whole pulse sequence (Pooley, 2005; Currie et al., 2012).

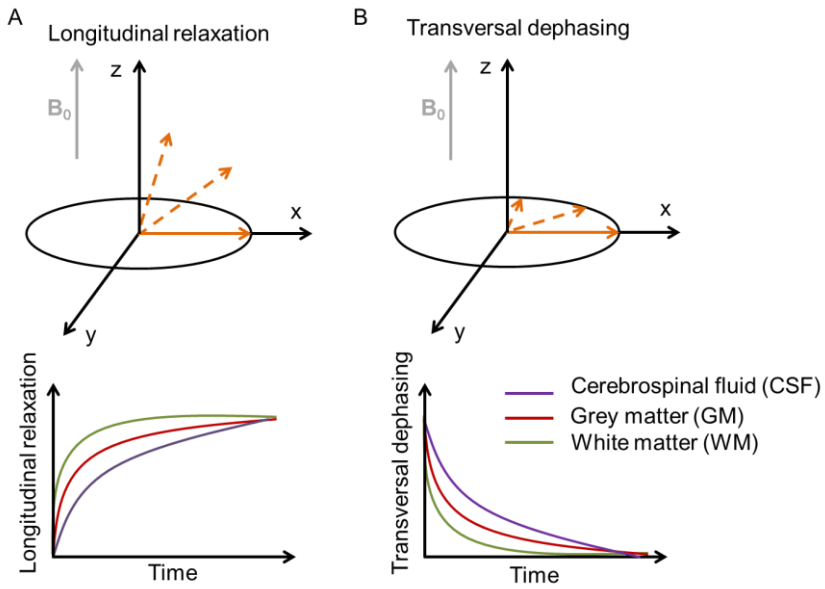


Figure 2.2: Relaxation and dephasing.

A. Longitudinal relaxation and ***B.*** Transversal dephasing. The time it takes for the longitudinal magnetization to move back to 63% of its original value is T1. The time it takes for the transverse magnetization to move back to 37% of its original value is T2. Cerebrospinal fluid, grey matter and white matter have different T1 and T2 rate, with CSF having the longest T1 and T2 time, white matter having the shortest T1 and

T2 time and grey matter having an intermediate T1 and T2 time. Adapted from Pooley (2005).

Depending on which effect is maximised (T1 or T2), different contrast images are produced: T1 images are used to discriminate anatomical differences in the brain, whereas T2 images are used to indirectly measure brain activity. Echo-planar imaging (EPI) is a fast MRI technique which permits to acquire the signal after a single RF pulse (Currie et al., 2012).

Changes (i.e. increase and decrease) of brain activity are related to changes in oxygenation concentration in haemoglobin (i.e., *neurovascular coupling*). Oxygenated haemoglobin has weak diamagnetic properties and reduces the magnetic field applied to the brain tissue, whereas deoxygenated haemoglobin has paramagnetic properties and enhances the magnetic field. The signal which is produced by these magnetic field changes is referred to as Blood Oxygen Level Dependent (BOLD) signal. When brain activity increases (e.g. during a task), oxygenated haemoglobin drops whereas deoxygenated haemoglobin increases (initial dip) followed by an increase of cerebral blood flow and oxygenated haemoglobin after two to six seconds (peak), which is the time for the blood to move from arteries to capillaries to veins. An undershoot follows the peak (Logothetis 2003, Fig. 2.3). More specifically, the BOLD signal arises from changes of transversal relaxation of the protons which experience dephasing due to changes in homogeneity, thus it is T2 (or T2*) which defines the BOLD signal. A slower phase decoupling in correspondence to a higher oxygenated/deoxygenated haemoglobin ratio corresponds to a higher T2* and thus to a higher BOLD signal.

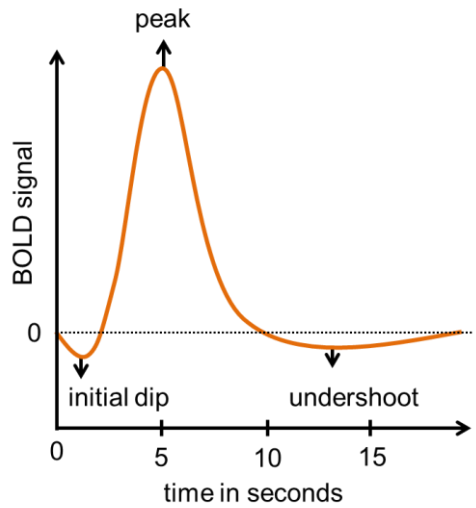


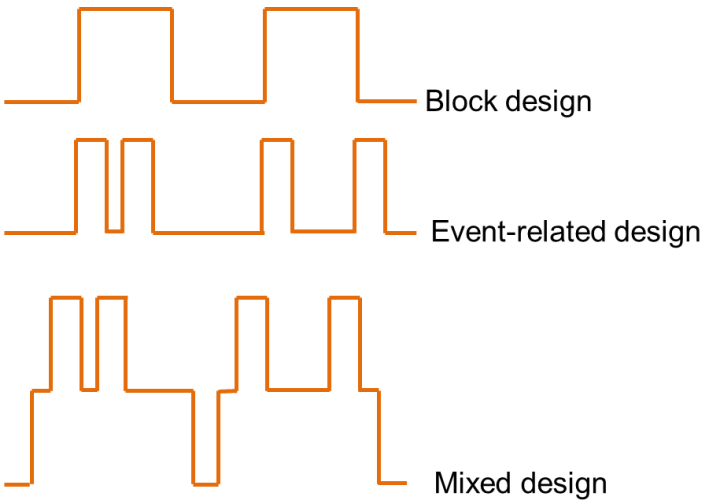
Figure 2.3: BOLD signal in function of time.
An initial dip in correspondence to an increase of deoxygenated blood precedes a peak of oxygenated blood which is followed by an undershoot.

2.3.2 fMRI design

To increase signal-to-noise ratio, experiment designs allow for numerous stimuli in each experimental condition. Block designs involve the presentation of trials in the same condition over blocks. This type of experimental design is characterised by a high statistical power, suitability for detecting regions of interest (ROI) and it is used in certain tasks where the randomization of stimuli is not possible. Limitations of this type of design are e.g. susceptibility to habituation effects and impossibility to distinguish different effects across the same block. Event-related designs involve intermixed presentation of trials of different conditions and at different inter-trials intervals. Despite having less power compared to the block design, this design allows for a better characterization of the hemodynamic response and for a post-hoc characterization of trials. Mixed designs consist of a combination of block and event-related designs where, within each block, stimuli from different conditions are

presented, thus allowing for a characterization of trial-specific activity but also sustained response (Fig. 2.4) (Petersen & Dubis, 2012).

Figure 2.4: Different experimental designs for fMRI experiments.



fMRI designs often use baseline periods (e.g., periods of rest of fixation or null trials) against which it is possible to contrast task activations. During this period participants do not perform an active task, nevertheless zero activity is never really achieved (Stark & Squire, 2001). If a brain region is significantly more active than baseline during a certain task, it is conventionally concluded that that region is involved in the processing of that specific task.

Parametric designs assume that a variation of a cognitive process corresponds to a variation in brain activation in a linear or non-linear fashion and include factors with multiple parametric levels. Conversely, factorial designs include categorical factors and are used to study the interaction between such factors (Friston et al., 2006).

2.3.3 fMRI data analysis

2.3.3.1 Preprocessing

Before fMRI data are analysed, volumes (i.e. images) are preprocessed to account for possible artefacts created by the MRI acquisition or by participants (e.g., breathing) and to prepare the data for following stages. Preprocessing is applied to both anatomical (e.g., T1 image) and EPI images. Different software packages are available to perform preprocessing and data analysis of fMRI data. Moreover, preprocessing pipelines closely depend on the experimental design, data acquisition and analysis (Lindquist, 2008, Poldrack et al., 2011; Ashburner et al., 2016). Finally, some steps can be applied in different orders and others are optional. The steps that are described in this paragraph (Fig. 2.5, Friston, 2006; Poldrack et al., 2011) refer to that used in this thesis (software: SPM, Friston et al., 1995).

Head motion can create a mismatch between the locations of single images in a time series. To correct for this mismatch, *realignment* is applied. Realignment is performed by aligning each image (with rigid body transformations, i.e. transformations which do not deform the shape of the head but only move the image via translation and rotation in the x, y and z directions) to a reference image which is usually the first in the time series or the middle one. Motion which is correlated to the task is nevertheless difficult to be accounted for because, by removing it, task-related activation can also be removed.

Unwarping refers to spatial realignment within the same volume and accounts for distortion in the volume due to magnetic field's inhomogeneities (i.e., *movement-by-distortion interactions*). This correction is applied interactively with realignment to reach a maximum likelihood solution.

To correct for differences arising due to different acquisition times (volumes are in fact acquired one at a time in an ascending or descending order), *slice-time correction* is applied. This correction is necessary to align the volumes in time as the model applied during data analysis assumes that all volumes were acquired at the same time. Slice-time correction is performed after selecting a reference slice and by applying temporal interpolation to all the other slices.

To allow for anatomical localization of brain activation and to improve normalization accuracy, the anatomical image is aligned to a predefined (e.g., mean) EPI image (i.e., *coregistration*) via rigid body (or *affine*) transformations.

Finally, a low-pass filter is applied to all images (i.e., *smoothing*). Images are smoothed to increase signal-to-noise ratio and to reduce variability across voxels (i.e., three dimensional units in the brain) across different brains. Smoothing is performed by convolving the volume to a Gaussian kernel of the magnitude of the full width at half the kernel's maximum height (FWHM). If a group analysis is performed, smoothing is usually applied after normalization. If a single-subject analysis is performed, smoothing is usually applied with a smaller kernel.

To be able to generalise fMRI results, group-analysis are used. To account for between-subjects brain variability, images of different subjects are aligned: first, the brain tissue of each participant has to be segmented into grey matter, white matter and CSF. This preprocessing step is called *segmentation* and it is based on the calculation of differences in the intensity of different tissues. Second, a volume-based registration to a template image is performed (i.e., *normalization*). Conventionally, the Montreal Neurological Institute template is employed (MNI: MNI305 and MNI152 were created as an average of a large group of individuals' anatomical scans). Image intensity and

prior spatial knowledge derived from the segmentation step are used to normalise anatomical and EPI images to the MNI template.

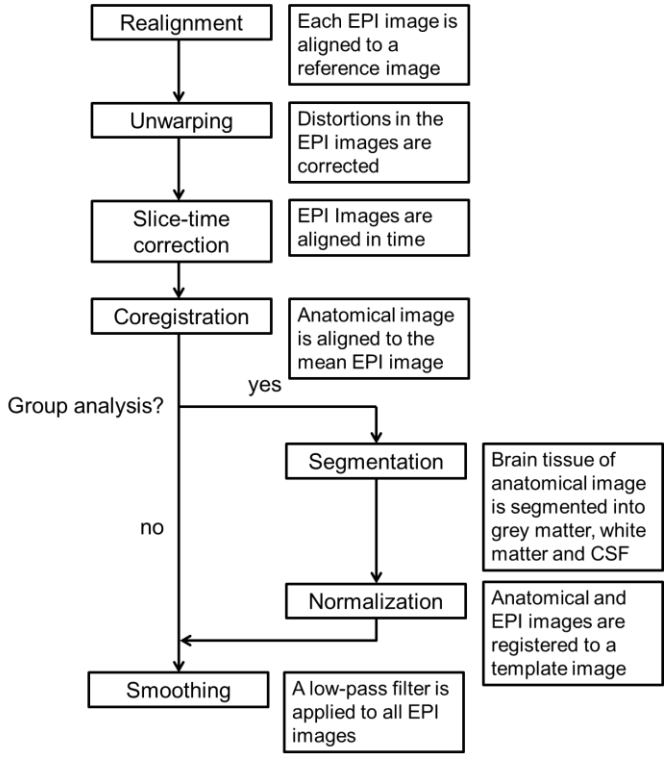


Figure 2.5: fMRI preprocessing pipeline used in this thesis.

2.3.3.2 First and second level data analysis

After preprocessing is performed, regions activated by a certain task are localised. This practically means to verify whether the BOLD signal changed in correspondence to the task. A GLM is used to model the BOLD signal which functions as dependent variable. Thus, the aim of the GLM analysis is to identify the predictors that can account for the BOLD signal in a voxel-to-voxel fashion, taking into consideration the fact that physiological response is sluggish.

The BOLD signal is assumed to have a canonical shape which develops over time and repeats time-locked to the stimulus. The expected BOLD response is the

result of the convolution between the stimulus function (i.e., *stimulus onsets*) and the hemodynamic response function (HRF). A stimulus onset time series is defined as the stimulus presentation over time and depends on the experimental paradigm. The HRF is characterised by an initial dip, a peak and an undershoot (this function is also known as double-gamma function, Friston et al., 1998). A more flexible version of the HRF model uses the canonical response function plus its derivatives (Friston et al., 1998). While the time derivative is used to model the temporal shift of the HRF, the dispersion derivative models the width of the function. An important property of the BOLD signal is that, if the neural activity is scaled by a factor of *n*, the BOLD signal is scaled by the same factor (*linearity*). Moreover, if the neural activity is shifted in time by *t* seconds, the BOLD signal is shifted in time by t seconds (*time invariance*). Thus, the GLM is defined as follows (as previously described in paragraph 2.1.1):

(15) $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \varepsilon$

In this case, Y is the BOLD signal and $X_{1:n}$ are the regressors of the experimental design after convolution with the HRF and its derivatives. Motion regressors calculated in the realignment step of the preprocessing can be included as nuisance regressors when they are not highly correlated with the task. To account for temporal autocorrelation of fMRI time series, *prewhitening* is applied and temporal autocorrelation not related to the task is removed.

The aim of the GLM analysis is to find the contribution of each regressor to the BOLD activity, calculated as follows:

(16) $\hat{\beta} = (X^T X)^{-1} X^T Y$

Once β is estimated, a statistical test can be carried out by specifying a design matrix and a contrast *c*. For instance, if a one-sample t-tests is to be performed, the null hypothesis can be specified as $c\beta = 0$ where $c = [1]$.

The result of the test is a statistical map (e.g., t-statistics maps) corrected for multiple comparisons (e.g., corrected for the number of voxels in the brain or the number of clusters), which represents how well each of the regressors explains the BOLD signal.

This analysis is run for each participant (i.e., *first level analysis*) and then contrast images are passed to the second level where an across participants GLM is performed (mixed-effects analysis, Holmes & Friston, 1998). Finally, to evaluate which contrasts are significant at the group level, further statistical tests are performed (e.g., ANOVA, t-test).

Such analyses can be run for the whole brain or within a specific region of interest (ROI). ROI analyses are employed, for instance, when a specific hypothesis about a given activation is available (e.g., anatomical information) (Poldrack et al., 2011). Importantly, ROI should not be defined based on functional results which are not independent to the results of interest (i.e., *double-dipping*, Kriegeskorte et al., 2009).

2.4 TRANSCRANIAL MAGNETIC STIMULATION

TMS is a non-invasive neuroimaging technique that allows for stimulation of the brain cortex in a relatively painless way. Differently from other neuroimaging techniques which only provide indirect (i.e., correlational) information about the link between a cognitive task and a brain system or area (e.g., fMRI), this technique permits to draw conclusions on the causal relationships between a given task and the stimulated brain region (Pascual-Leone et al., 1999). The physical principles of TMS were discovered by Michael Faraday in 1881. TMS was officially first introduced in 1984 by Anthony Barker and in 1994 Alvaro Pascual-Leone first introduced repetitive TMS (rTMS). Since then, TMS is used both for research and as a therapeutic tool (Horvath et al., 2011).

2.4.1 Physics and physiology of TMS

The TMS coil consists of a coil of wire through which passes a current which rapidly changes in time. This produces a magnetic field (1.5-2 tesla strength) orthogonally to the direction of the coil. This magnetic field passes through the head tissue without impedance and, for the principle of inductance, it induces another weaker current (150 Volt/meter, due to the low conductivity of biological tissue) in the brain up to 3 cm depth, parallel to the coil (Wassermann, 1998; Rossi et al, 2009). The strength of the magnetic field rapidly weakens with distance from the coil, making it difficult to stimulate deep brain areas (Siebner et al., 2009). TMS more likely induces initiation of action potential in the axons of neurons rather that in the cells, because axons are shown to be efficiently activated by pulses with short duration (Amassian et al., 1998; Siebner et al., 2009). The exact way TMS interferes with a certain brain

region is not completely understood. Studies suggest that this could take place (1) by silencing neuronal activity (via GABA increase) and (2) by introducing random neural noise (Siebner et al., 2009; Sandrini et al., 2011). Importantly, the state of the cortex during stimulation is crucial to determine the type of cortical response (i.e., *state-dependency effects*, Cattaneo et al., 2005).

Coil orientation has a direct influence on the effects of TMS, with maximum effects when the coil is positioned with its axis oriented perpendicular to the gyrus (Gomez-Tames et al., 2018). In fact, currents are oriented parallel to the coil's plane and preferentially flow through brain tissue oriented horizontally to the brain surface (Pascual-Leone et al., 1999). Effects of TMS also depend on the orientation of the stimulated tissue and are maximal where axons bend out of the electrical field. In such areas lower stimulation thresholds are reported (Maccabee et al., 1993; Pascual-Leone et al., 1999). To date, it remains unclear which exact set of axons is activated when TMS is applied, especially given the variability of the cortical anatomy (Siebner et al., 2009). Moreover, TMS effects (especially rTMS) can spread following the path of least resistance and affect distant brain regions through cortico-cortical connections (Ilmoniemi et al., 1997; Pascual-Leone et al., 1999). For this reason, TMS effects are not always limited to focal areas.

2.4.2 TMS protocols

Various TMS protocols are available and, depending on the type of coil, the frequency of the TMS pulse, the time when TMS is applied and the intensity of stimulation (Fig. 2.6), different physiological and behavioural effects are observed.

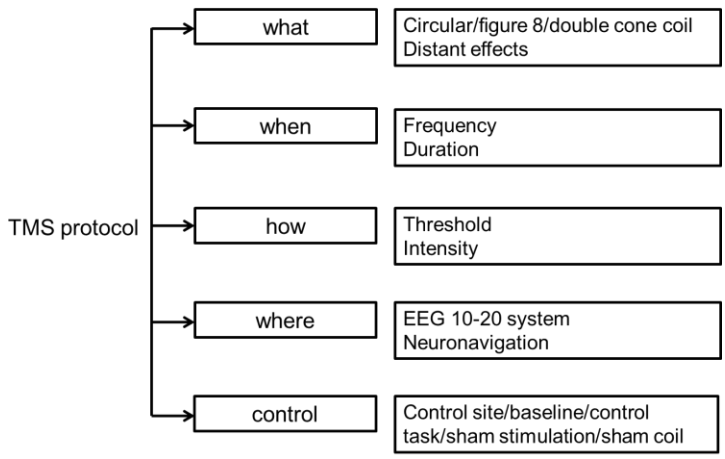


Figure 2.6: TMS parameters.

Different TMS protocols can be used depending on the aim of the experiment. Aspects which deserve consideration when designing a TMS experiment are: (1) which regions to stimulate (e.g. deeper or more focal regions) and which other areas are influenced by stimulation of a specific spot (i.e., distant effects); (2) when to stimulate: with which frequency and for how long (safety guidelines need to be considered); (3) the amount of intensity based on previous hypothesis and/or on participants' motor or phosphene threshold; (4) where to stimulate: neuronavigation systems allow for a better precision than EEG 10-20 systems; (5) one or more control conditions allow for more sound conclusions (adapted from Robertson et al., 2003).

2.4.2.1 Coils, waveforms and direction of stimulation

Circular coils (ring-shaped) produce a less focal stimulation but, because of their larger diameter, they allow for deeper stimulations. The double cone coil consists of a combination of two circular wings which form an angle of 95 degrees. This coil is used

to stimulate deep brain regions such as the cerebellum. The figure 8 coil allows for a more focal but shallower stimulation right beneath the central segment where the two wings meet (Pascual-Leone et al., 1999; Rossi et al., 2009, Fig. 2.7A).

Monophasic or biphasic pulses are available for stimulation. Biphasic stimuli are more effective than monophasic stimuli, for the initial current with the same amplitude (Sommer et al., 2006). In light of the fact that more efficient stimulation is achieved during the second phase of biphasic stimuli and when the current flows in the posterior-to-anterior direction for both waveforms, the coil has to be positioned differently for monophasic or biphasic stimulation (i.e., with the handle pointing in opposite directions, Kammer et al., 2001).

2.4.2.2 Frequency and time of application

TMS can be applied in single pulses (*single-pulse* TMS), in pairs of pulses (*paired-pulse* TMS), in a repetitive fashion (*repetitive* TMS, rTMS) or in patterns of short rTMS burst (e.g., *theta burst* TMS, TBS). At least 7 seconds are thought to be sufficient to separate pulses in the single-pulse TMS to prevent pulses to interact to each other (Robertson et al., 2003). In paired-pulse TMS, two pulses are applied to the same or to different regions. This protocol is usually used to measure intracortical or cortico-cortical effects (Rossi et al., 2009). The frequency of rTMS is conventionally defined *high* when > 1 Hz and *low* when ≤ 1 Hz. TMS can be applied when the task is performed (*online* TMS) or before or after the task is performed (*offline* TMS) (Fig. 2.7B). For theta-burst stimulation (TBS), burst of 50 Hz applied at 5Hz are used in a continuous or intermittent (i.e. 2 seconds every 10 seconds) fashion (cTBS and iTBS).

Online rTMS has been shown to induce a larger effect with higher frequencies. Offline low frequency rTMS and cTBS are associated to a decrease of cortical excitability, whereas offline high frequency rTMS and iTBS are associated to cortical

excitability increase (Robertson et al., 2003; Rossi et al., 2009, Suppa et al., 2015). However, behavioural outcomes for these stimulation protocols are less clear when brain regions other than the motor cortex are stimulated. For instance, following offline low-frequency rTMS, both worsening and improvement of performance were reported (Hilgetag et al., 2001; Pell et al., 2011).

rTMS and TBS have been shown to have longer-lasting after effects. Previous research demonstrated that the length of the aftereffect can last up to 50-100% of the stimulation duration (Robertson et al., 2003 & Hoogendam et al., 2010), whereas 20 seconds of iTBS or cTBS and 40 seconds of cTBS shows an after-effect of 20 and 60 minutes respectively (Suppa et al., 2015). Importantly, differences in terms of length of after-effects are observed depending on the stimulated area and more results are still needed to make conclusive inferences for brain areas other than the motor cortex (Robertson et al., 2003, Suppa et al., 2015). The fact that rTMS shows longer-lasting effects suggests that this type of stimulation affects synaptic plasticity through processes like long-term potentiation (LTP) and long-term depression (LTD). Synaptic plasticity refers to a form or brain reorganization at the synaptic level that follows and outlasts the experimental training or stimulation (Robertson et al., 2003). LTP refers to a potentiation of the synaptic strength and it is broadly associated to increased cortical excitability, whereas LTD refers to a weakening of synaptic strength and it is associated to decreased cortical excitability (Hoogeman et al., 2010).

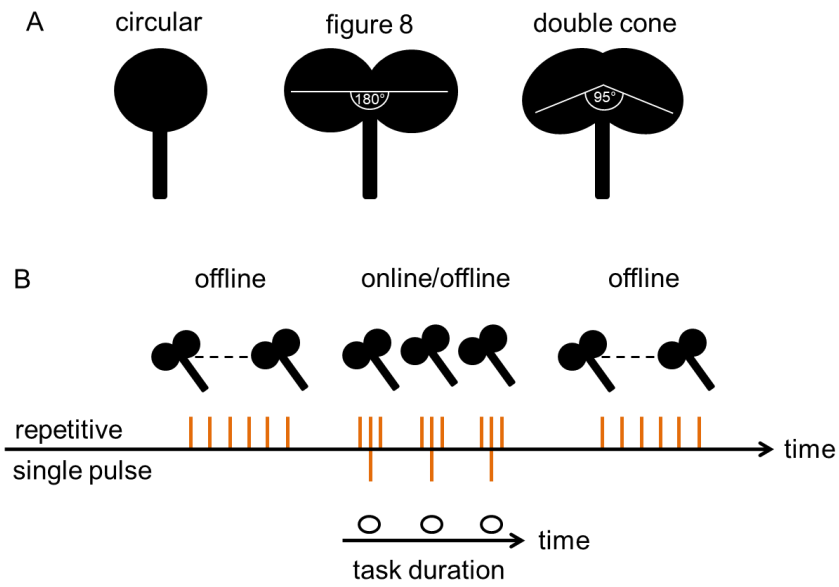


Figure 2.7: Coils and protocols.

A. Different types of available coils (in black). ***B.*** Example of stimulation times (offline or online) and frequencies (repetitive or single pulse) and their application time in function of task duration and stimulus presentation. Pulses are represented by orange lines and stimuli by white discs.

2.4.2.3 Intensity, threshold and safety guidelines

Intensity of stimulation can either be fixed for all participants (when a solid hypothesis is available and depending on the other parameters involved) or adjusted for each participant. In fact, cortical excitability changes depending on the brain area and is variable across participants. Individual motor or phosphene threshold (MT or PT) are used to determine the individual motor or visual cortex excitability (Rossini et al., 1994; Sandrini et al, 2011). MT is measured as the lowest stimulation intensity which can produce 5 out of 10 visible first dorsal interosseous muscle twitches (motor evoked potentials, MEP) or 5 out of 10 MEP of no less than 50µV when the limb is at rest (*resting* MT) and 200µV when the muscle is 10-20% contracted (*active* MT) (López-Alonso et al., 2014, Vernet et al., 2014). PT is measured as the lowest

stimulation intensity which can produce 5 out of 10 reported phosphenes (Stewart et al., 2001). Despite the wide use of such measures, it is not yet clear whether MT and PT correlate (Stewart et al., 2001; Boroojerdi et al., 2002; Antal et al., 2004; Deblieck et al., 2008). Moreover, the effect of MT and PT intensity on brain areas other than motor and visual cortices is controversially discussed (although some guidelines are available, e.g., Oliveri et al., 2000; Oliver et al., 2009).

Once a threshold has been established, stimulation can be performed at threshold, above or below threshold. In order to minimize the risk of seizure and other adverse effects, guidelines were published to guide experimenters and help them build a safe protocol which takes into consideration intensity, frequency and duration of stimulation in healthy individuals and patients (Wasserman, 1998; Rossi et al, 2009). Moreover, safety questionnaires (Rossi et al, 2009) are used by experimenters to evaluate participant inclusion depending on the type of study (i.e., research, diagnostic, therapeutic).

2.4.2.4 Neuronavigation

After a brain region of interest has been established, a suitable stimulation site has to be identified on the participant's scalp. Due to subjects' brain and scalp variability, there is not a 1:1 correspondence between brain target region and scalp entry. The international 10-20 EEG system can be used to position the coil, as it is possible to correlate the electrode placement with the underlying cortical area (Herwig et al., 2003; Okamoto et al., 2004). Alternatively, for a more accurate positioning, neuroimaging-based methods are used, such as *image-guided frameless stereotaxic neuronavigation*. With this method, each participant's MRI image is coregistered to their head in space via facial/cranial landmarks (e.g., nasion, inion, ears, eyes) (Sparing

et al., 2010). This method permits to accurately navigating the participant's brain image from their scalp.

2.4.2.5 Control conditions

Different control conditions are used to better evaluate the effects of TMS. For instance: (1) Stimulation can be delivered over a control site which is thought not to be involved in the task (i.e., control site); (2) The task can be performed without stimulation (i.e., baseline); (3) Stimulation over the same site can be used together with another task which does not involve the cognitive processes which are thought to be involved in the task of interest (i.e., control task). (4) The coil can be tilted 45° or 90° from the scalp in order to resemble the auditory clicking sound of the real stimulation with only little or no voltage introduced in the brain (i.e., sham stimulation) (Lisanby et al., 2001); (5) A sham coil can be used, which consists of a regular coil provided with a thick shield and which resemble the auditory clicking sound of the real stimulation (i.e., sham coil). Depending on the task and the protocol, one or more control conditions should be employed to achieve robust conclusions (Robertson et al., 2003).

REFERENCES

Antal, A., Nitsche, M. A., Kincses, T. Z., Lampe, C., & Paulus, W. (2004). No correlation between oving phosphene and motor thresholds: a transcranial magnetic stimulation study. *Neuroreport*, *15*(2), 297–302. http://doi.org/10.1097/01.wnr.0000099600.41403.ec

Ashburner, J., Chen, C., Moran, R., Henson, R., & Phillips, C. (2016). SPM12 manual, 1–508.

Blignaut, P. (2009). Fixation identification: the optimum threshold for a dispersion algorithm. *Attention, Perception & Psychophysics*, *71*(4), 881–895.

Boroojerdi, B., Meister, I. G., Foltys, H., Sparing, R., Cohen, L. G., & Töpper, R. (2002). Visual and motor cortex excitability: a transcranial magnetic stimulation study. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *113*, 1501–1504. http://doi.org/10.1016/S1388-2457(02)00198-0

Caffier, P. P., Erdmann, U., & Ullsperger, P. (2003). Experimental evaluation of eye-blink parameters as a drowsiness measure. *European Journal of Applied Physiology*, *89*(3–4), 319–325. http://doi.org/10.1007/s00421-003-0807-5

Cattaneo, L., Voss, M., Brochier, T., Prabhu, G., Wolpert, D. M., & Lemon, R. N. (2005). A cortico-cortical mechanism mediating object-driven grasp in humans. *Proceedings of the National Academy of Sciences*, *102*(3), 898–903. http://doi.org/10.1073/pnas.0409182102

Chakrabarti, A., & Ghosh, J. K. (2011). AIC, BIC and recent advances in model selection. *Philosophy of Statistics*, *7*, 583–605. http://doi.org/10.1016/B978-0-444-51862-0.50018-6

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. *Routledge Academic*. http://doi.org/10.2307/2529115

Currie, S., Hoggard, N., Craven, I. J., Hadjivassiliou, M., & Wilkinson, I. D. (2012). Understanding MRI: Basic MR physics for physicians. *Postgraduate Medical Journal*, *89*(1050), 209–223. http://doi.org/10.1136/postgradmedj-2012-131342

Deblieck, C., Thompson, B., Iacoboni, M., & Wu, A. D. (2008). Correlation between motor and phosphene thresholds: A transcranial magnetic stimulation study. *Human Brain Mapping*, *29*(6), 662–670. http://doi.org/10.1002/hbm.20427

Devore, J. L., & Farnum, N. R. (1999). *Applied statistics for engineers and scientists.* Pacific Grove, Calif: Duxbury Press.

Frane, A. V. (2015). Are per-family type I error rates relevant in social and behavioral science? *Journal of Modern Applied Statistical Methods*, *14*(1), 12–23. http://doi.org/10.22237/jmasm/1430453040

Friston, K. J. (2006). *Statistical Parametric Mapping: the analysis of functional brain images*. *Functional neuroimaging: Technical*. Academic Press.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., & Frackowiak, R. S. J. (1995). Statistical parametric maps in functional imaging : A general linear approach. *Human Brain Mapping*, *2*, 189–210.

Gomez-Tames, J., Hamasaka, A., Laakso, I., Hirata, A., & Ugawa, Y. (2018). Atlas of optimal coil orientation and position for TMS: A computational study. *Brain Stimulation*, *11*(4), 839–848. http://doi.org/10.1016/j.brs.2018.04.011

Hamker, F. H. (2005). The reentry hypothesis: The putative interaction of the frontal eye field, ventrolateral prefrontal cortex, and areas V4, IT for attention and eye movement. *Cerebral Cortex*, *15*(4), 431–447. http://doi.org/10.1093/cercor/bhh146

Herwig, U., Satrapi, P., & Schönfeldt-Lecuona, C. (2003). Using the international 10-20 EEG system for positioning of transcranial magnetic stimulation. *Brain Topography*, *16*(2), 95–99. http://doi.org/10.1023/B:BRAT.0000006333.93597.9d

Hilgetag, C. C., Théoret, H., & Pascual-leone, A. (2001). Enhanced visual spatial attention ipsilateral to rTMS-induced “virtual lesions” of human parietal cortex. *Nature Neuroscience*, *4*(9), 953–957. http://doi.org/10.1038/nn0901-953

Holmes, A. P., & Friston, K. J. (1998). Generalisability, random effects and population inference. *NeuroImage*, *7*, S754.

Hoogendam, J. M., Ramakers, G. M. J., & Di Lazzaro, V. (2010). Physiology of repetitive transcranial magnetic stimulation of the human brain. *Brain Stimulation*, *3*(2), 95–118. http://doi.org/10.1016/j.brs.2009.10.005

Horvath, J. C., Perez, J. M., Forrow, L., Fregni, F., & Pascual-Leone, A. (2011). Transcranial magnetic stimulation: a historical evaluation and future prognosis of therapeutically relevant ethical concerns. *Journal of Medical Ethics*, *37*(3), 137–143. http://doi.org/10.1136/jme.2010.039966

Ilmoniemi, R. J., Virtanen, J., Ruohonen, J., Karhu, J., Aronen, H. J., Näätänen, R., & Katila, T. (1997). Neuronal responses to magnetic stimulation reveal cortical reactivity and connectivity. *NeuroReport*, *8*(16), 3537–3540. http://doi.org/10.1097/00001756-199711100-00024

Kammer, T., Beck, S., Thielscher, A., Laubis-Hermann, U., & Topka, H. (2001). Motor threshold in humans: a transcranial magnetic stimulation study comparing different pulse waveforms, current directions and stimulator types. *Clinical Neurophysiology*, *112*, 250–8. http://doi.org/10.1016/S1388-2457(00)00513-7

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S. F., & Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nature Neuroscience*, *12*(5), 535–540. http://doi.org/10.1167/8.6.88

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*(NOV), 1–12. http://doi.org/10.3389/fpsyg.2013.00863

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian Modeling for Cognitive Science: A Practical Course*. Cambridge University Press.

Levine, T. R., & Hullett, C. R. (2002). Eta squared, partial eta squared, and misreporting of effect size in communication research. *Human Communication Research*, *28*(4), 612–625. http://doi.org/10.1093/hcr/28.4.612

Lindquist, M. A. (2008). The statistical analysis of fMRI data. *Statistical Science*, *23*(4), 439–464. http://doi.org/10.1214/09-STS282

Lisanby, S. H., Gutman, D., Luber, B., Schroeder, C., & Sackeim, H. A. (2001). Sham TMS: Intracerebral measurement of the induced electrical field and the induction of motor-evoked potentials. *Biological Psychiatry*, *49*(5), 460–463. http://doi.org/10.1016/S0006-3223(00)01110-0

Logothetis, N. K. (2003). The underpinnings of the BOLD functional magnetic resonance imaging signal. *The Journal of Neuroscience*, *23*(10), 3963–3971. http://doi.org/23/10/3963 [pii]

López-Alonso, V., Cheeran, B., Río-Rodríguez, D., & Fernández-Del-Olmo, M. (2014). Inter-individual variability in response to non-invasive brain stimulation paradigms. *Brain Stimulation*, *7*(3), 372–380. http://doi.org/10.1016/j.brs.2014.02.004

Maccabee, P. J., Amassian, V. E., Eberle, L. P., & Cracco, R. Q. (1993). Magnetic coil stimulation of straight and bent amphibian and mammalian peripheral nerve in vitro: locus of excitation. *The Journal of Physiology*, *460*(1), 201–219. http://doi.org/10.1113/jphysiol.1993.sp019467

McCullagh, P., & Nelder, J. A. (1989). *Generalized Linear Models*. London: Chapman & Hall. http://doi.org/10.1007/978-1-4899-3242-6

Meindertsma, T., Kloosterman, N. A., Engel, A. K., Wagenmakers, E. J., & Donner, T. H. (2018). Surprise about sensory event timing drives cortical transients in the Beta frequency band. *The Journal of Neuroscience*, 0307-18. http://doi.org/10.1523/JNEUROSCI.0307-18.2018

Okamoto, M., Dan, H., Sakamoto, K., Takeo, K., Shimizu, K., Kohno, S., … Dan, I. (2004). Three-dimensional probabilistic anatomical cranio-cerebral correlation via the international 10–20 system oriented for transcranial functional brain mapping. *NeuroImage*, *21*(1), 99–111. http://doi.org/10.1016/j.neuroimage.2003.08.026

Oliver, R., Bjoertomt, O., Driver, J., Greenwood, R., & Rothwell, J. (2009). Novel "hunting" method using transcranial magnetic stimulation over parietal cortex disrupts visuospatial sensitivity in relation to motor thresholds. *Neuropsychologia*, *47*(14), 3152–61. http://doi.org/10.1016/j.neuropsychologia.2009.07.017

Oliveri, M., Caltagirone, C., Filippi, M. M., Traversa, R., Cicinelli, P., Pasqualetti, P., & Rossini, P. M. (2000). Paired transcranial magnetic stimulation protocols reveal a pattern of inhibition and facilitation in the human parietal cortex. *Journal of Physiology*, *529*(2), 461–468. http://doi.org/10.1111/j.1469-7793.2000.00461.x

Pascual-Leone, A., Bartres-Fazf, D., & Keenan, J. P. (1999). Transcranial magnetic stimulation: studying the brain-behaviour relationship by induction of "virtual lesions." *Philosophical Transactions of the Royal Society B: Biological Sciences*, *354*(1387), 1229–1238. http://doi.org/10.1098/rstb.1999.0476

Pell, G. S., Roth, Y., & Zangen, A. (2011). Modulation of cortical excitability induced by repetitive transcranial magnetic stimulation: Influence of timing and geometrical parameters and underlying mechanisms. *Progress in Neurobiology*, *93*(1), 59–98. http://doi.org/10.1016/j.pneurobio.2010.10.003

Petersen, S. E., & Dubis, J. W. (2012). The mixed block/event-related design. *NeuroImage*, *15*(62), 1177–1184. http://doi.org/10.1016/j.asieco.2008.09.006.EAST

Poldrack, R. A., Mumford, J. A., & Nichols, T. (2011). *Handbook of functional MRI data analysis*. *Cambridge University Press*.

Pooley, R. A. (2005). AAPM/RSNA physics tutorial for residents: fundamental physics of MR imaging. *Radiographics : A Review Publication of the Radiological Society of North America, Inc*, *25*(4), 1087–1099. http://doi.org/10.1148/rg.254055027

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, *13*, 25–42.

Quintana, D. S., & Williams, D. R. (2018). Bayesian alternatives for common null-hypothesis significance tests in psychiatry: A non-technical guide using JASP. *BMC Psychiatry*, *18*(178), 1–8. http://doi.org/10.17605/OSF.IO/WUN5V

Rizzolatti, G., Gentilucci, M., Matelli, M. (1985). Selective spatial attention: One center, one circuit or many circuits. In *Attention and Performance XI*, ed. M. Posner,O. S. M. Marin, pp. 251-65.

Robertson, E. M., Théoret, H., & Pascual-Leone, A. (2003). Studies in cognition: the problems solved and created by transcranial magnetic stimulation. *Journal of Cognitive Neuroscience*, *15*(7), 948–960. http://doi.org/10.1162/089892903770007344

Rossi, S., Hallett, M., Rossini, P. M., & Pascual-Leone, A. (2009). Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clinical Neurophysiology*, *120*(12), 2008–2039. http://doi.org/10.1016/j.clinph.2009.08.016

Rossini, P. M., Barker, A. T., Berardelli, A., Caramia, M. D., Caruso, G., Cracco, R. Q., … Tomberg, C. (1994). Non-invasive electrical and magnetic stimulation of the brain, spinal cord and roots: basic principles and procedures for routine clinical application. Report of an IFCN committee. *Electroencephalography & Clinical Neurophysiology*, *91*, 79–92. http://doi.org/10.1016/0013-4694(94)90029-9

Rutherford, A. (2001). *ANOVA and ANCOVA: A GLM Approach*. *SAGE* (SAGE). London. http://doi.org/10.1002/9781118491683

Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. *Proceedings of the Eye Tracking Research and Applications Symposium*, 71–78. http://doi.org/10.1145/355017.355028

Sandrini, M., Umiltà, C., & Rusconi, E. (2011). The use of transcranial magnetic stimulation in cognitive neuroscience: A new synthesis of methodological issues. *Neuroscience and Biobehavioral Reviews*, *35*(3), 516–536. http://doi.org/10.1016/j.neubiorev.2010.06.005

Schervish, M. (1996). P values: What they are and what they are not. *American Statistician*, *50*(3), 203–206.

Siebner, H. R., Hartwigsen, G., Kassuba, T., & Rothwell, J. (2009). How does transcranial magnetic stimulation modify neuronal activity in the brain ? - Implications for studies of cognition. *Cortex*, *45*(9), 1035–1042. http://doi.org/10.1016/j.cortex.2009.02.007.How

Sommer, M., Alfaro, A., Rummel, M., Speck, S., Lang, N., Tings, T., & Paulus, W. (2006). Half sine, monophasic and biphasic transcranial magnetic stimulation of the human motor cortex. *Clinical Neurophysiology*, *117*(4), 838–844. http://doi.org/10.1016/j.clinph.2005.10.029

Sparing, R., Hesse, M. D., & Fink, G. R. (2010). Neuronavigation for transcranial magnetic stimulation (TMS): Where we are and where we are going. *Cortex*, *46*(1), 118–120. http://doi.org/10.1016/j.cortex.2009.02.018

Spence, C., & Driver, J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(4), 1005–1030. http://doi.org/10.1037/0096-1523.22.4.1005

Stark, C. E. L., & Squire, L. R. (2001). When zero is not zero: The problem of ambiguous baseline conditions in fMRI. *Proceedings of the National Academy of Sciences*, *98*(22), 12760–12766. http://doi.org/10.1073/pnas.221462998

Stewart, L. M., Walsh, V., & Rothwell, J. C. (2001). Motor and phosphene thresholds: A transcranial magnetic stimulation correlation study. *Neuropsychologia*, *39*(4), 415–419. http://doi.org/10.1016/S0028-3932(00)00130-5

Suppa, A., Huang, Y. Z., Funke, K., Ridding, M. C., Cheeran, B., Di Lazzaro, V., … Rothwell, J. C. (2015). Ten years of theta burst stimulation in humans: established knowledge, unknowns and prospects. *Brain Stimulation*, *9*(3), 323–335. http://doi.org/10.1016/j.brs.2016.01.006

Van Der Stigchel, S., & Theeuwes, J. (2007). The relationship between covert and overt attention in endogenous cuing. *Perception and Psychophysics*, *69*(5), 719–731. http://doi.org/10.3758/BF03193774

Vernet, M., Bashir, S., Yoo, W.-K., Oberman, L., Mizrahi, I., Ifert-Miller, F., … Pascual-Leone, A. (2014). Reproducibility of the effects of theta burst stimulation on motor cortical plasticity in healthy participants. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *125*(2), 320–6. http://doi.org/10.1016/j.clinph.2013.07.004

Wassermann, E. M. (1998). Risk and safety of repetitive transcranial magnetic stimulation: report and suggested guidelines from the international workshop on the safety of repetitive Transcranial Magnetic Stimulation, June 5 – 7, 1996. *Electroencephalography and Clinical Neurophysiology*, *108*, 1–16.

West, B. T., Welch, K. B., & Gałecki, A. T. (2006). *Linear Mixed Models : A practical guide using statistical software*. *CRC Press* (CRC Press). http://doi.org/doi:10.1201/b17198-3

Westwood, S. J., Olson, A., Miall, R. C., Nappo, R., & Romani, C. (2017). Limits to tDCS effects in language: Failures to modulate word production in healthy participants with frontal or temporal tDCS. *Cortex*, *86*, 64–82. http://doi.org/10.1016/j.cortex.2016.10.016

Wickens, T. (2002). *Elementary signal detection theory*. New York: Oxford University Press.

Chapter 3:

Additive and interactive effects of spatial attention and expectation on perceptual decisions

Arianna Zuanazzi[1], Uta Noppeney[1]

[1]Computational Cognitive Neuroimaging Laboratory, Computational Neuroscience and Cognitive Robotics Centre, University of Birmingham, B15 2TT Birmingham, UK

Published paper:
Scientific Reports (2018), 8:6732, doi:10.1038/s41598-018-24703-6

Author contributions:
Conceived and designed the experiments: **Arianna Zuanazzi**, Uta Noppeney.
Performed the experiments: **Arianna Zuanazzi**.
Contributed to data analysis: **Arianna Zuanazzi**, Uta Noppeney.
Contributed to the writing of the manuscript: **Arianna Zuanazzi**, Uta Noppeney.

ABSTRACT

Spatial attention and expectation are two critical top-down mechanisms controlling perceptual inference. Based on previous research it remains unclear whether their influence on perceptual decisions is additive or interactive.

We developed a novel multisensory approach that orthogonally manipulated spatial attention (i.e. task-relevance) and expectation (i.e. signal probability) selectively in audition and evaluated their effects on observers' responses in vision. Critically, while experiment 1 manipulated expectation directly via the probability of task-relevant auditory targets across hemifields, experiment 2 manipulated it indirectly via task-irrelevant auditory non-targets.

Surprisingly, our results demonstrate that spatial attention and signal probability influence perceptual decisions either additively or interactively. These seemingly contradictory results can be explained parsimoniously by a model that combines spatial attention, general and spatially selective response probabilities as predictors with no direct influence of signal probability. Our model provides a novel perspective on how spatial attention and expectations facilitate effective interactions with the environment.

Keywords:

spatial attention · spatial expectation · response probability · multisensory · perceptual decision making

3.1 INTRODUCTION

Generating a coherent representation of the world from the sensory signals with which we are bombarded is a fundamental challenge facing us in everyday life. Crucially, perceptual inference is not purely driven by bottom-up signals but also guided by top-down selective attention and prior expectations (or predictions). Selective attention shapes perception by prioritizing processing of signals that are relevant for our current goals. Conversely, prior expectations encode the probabilistic structure of the environment[1]. Based on past events and experiences we generate expectations or predictions for the future. While attention and expectation are both thought to facilitate perception at the behavioural level[2-10], they have traditionally been associated with distinct neural effects. Attention amplifies neural activity and signal to noise ratio[11-14] while expectation often leads to a reduction in stimulus-induced neural activity[15-20].

According to the notion of predictive coding, perceptual inference emerges within the cortical hierarchy via iterative adjustment of top-down predictions against bottom-up sensory evidence. While backwards connections impose predictions from higher to subordinate level, forwards connections furnish the prediction error, i.e. the discrepancy between prediction and sensory evidence, from lower to higher hierarchical levels[21-25]. Attention may influence perceptual inference by enhancing the precision (i.e. inverse of variance) of the prediction and/or prediction error signal leading to an increase in sensory gain for attended signals[24-26]. As a consequence, attention and expectation jointly shape perceptual inference and decisions.

Surprisingly, research to date has mostly conflated attention and expectation[8,27,28]. Most prominently, the so-called Posner cuing paradigm[29]

manipulates observers' endogenous spatial attention using a cue that probabilistically predicts the location of the subsequent signal thereby confounding spatial attention and spatial expectation (i.e. signal probability)[30]. Only recently have studies attempted to dissociate spatial attention and expectation by orthogonally manipulating task-relevance (i.e. response requirement) and spatial signal probability. Using fMRI, a previous study by Kok et al.[31] showed that spatial attention and expectation influence neural responses in an interactive fashion. More specifically, attention reversed the activation increase for unexpected relative to expected signals that were observed for unattended signals. The interactive effects between attention and spatial signal probability were interpreted as in line with the notion of precision weighted prediction errors as embodied in predictive coding models[24,31]. Yet, a critical limitation of those neuroimaging experiments is that synergistic effects between attention and expectation could be evaluated only at the neural but not the behavioural level[31,32], because spatial attention was manipulated as response requirement over space[32]. As a result, observers did not respond to the unattended, i.e. task-irrelevant, signals and the effects of spatial expectation on response times could only be evaluated for signals in the attended hemifield[31]. This raises the critical question whether the interactions between spatial attention and spatial signal probability are behaviourally relevant for effective interactions with the environment. How does the brain optimize detection of signals across the spatial field depending on attention and expectations formed based on signal probability? In the current study we have developed a novel multisensory approach to determine whether spatial attention and signal probability influence behaviour additively and/or interactively in a target detection task.

3.2 RESULTS

In a series of two experiments, participants were presented on each trial either with an auditory burst of white noise or a visual flash in their left or right hemifield. We orthogonally manipulated spatial attention (i.e. response requirement for signals presented in a particular hemifield) and expectation (i.e. signal probability in a particular hemifield) selectively in audition and evaluated their effects on target detection in audition and vision, where signal probability and response requirements were held constant. This multisensory generalization approach[27,28] provides us with the novel opportunity to evaluate the putative additive or interactive effects of spatial attention and signal probability at the behavioural level rather than only implicitly at the neural level as in previous unisensory research[31].

Critically, experiment 1 manipulated auditory spatial expectation directly via the probability of auditory targets across the two hemifields, which led to differences in the general response probability across conditions (i.e. run type A (blue) vs run type B (green), see Figure 3.1a, 3.1b and 3.1d, Supplementary Table 8.1 and [31] for related design). By contrast, experiment 2 manipulated auditory expectation via task-irrelevant non-targets that never required a response and was thereby able to hold the general response probability constant across all conditions (see Figure 3.2a, 3.2b and 3.2d, Supplementary Table 8.1 and [32] for related design).

First, we asked for each experiment independently whether spatial attention and signal probability shapes target detection responses in an additive or interactive fashion. This question can be addressed only for visual signals where responses were collected over both attended and unattended hemifields. To assess whether the effects of expectation generalise from audition to vision in the attended hemifield, we also report the results for auditory targets in the attended hemifield.

3.2.1 Experiment 1

For the auditory modality, the two-sided paired-sample t-tests on hit rates and response times in the attended hemifield showed non-significantly higher hit rates ($t(14) = 2.06$, $p = 0.058$, Cohen's d_{av} [95% CI] = 0.56 [0.02, 1.12][*]) and significantly faster responses when the hemifield was expected than unexpected ($t(14) = -3.23$, p=.006, Cohen's d_{av} [95% CI] = -0.31 [-0.52, -0.09] (see Supplementary Figure 8.1, left bar plot of Figure 3.1e and Table 3.1).

For the visual modality, the 2 (attended vs. unattended) x 2 (expected vs. unexpected) rmANOVAs revealed a significant main effect of attention on hit rates ($F(1, 14) = 5.56$, $p = 0.033$, η_p^2 [90% CI] = 0.29 [0.01, 0.51]) and response times ($F(1, 14) = 42.81$, $p < 0.001$, η_p^2 [90% CI] = 0.75 [0.48, 0.84]). Critically, we also observed a significant crossover interaction between attention and expectation for hit rates ($F(1, 14) = 4.89$, $p = 0.044$, η_p^2 [90% CI] = 0.26 [0.004, 0.49]) and response times ($F(1, 14) = 8.90$, $p = 0.010$, η_p^2 [90% CI] = 0.39 [0.06, 0.59]). Participants had non-significantly higher hit rates ($t(14) = 2.01$, $p = 0.064$, Cohen's d_{av} [95% CI] = 0.44 [-0.02, 0.90][†]) and responded significantly faster ($t(14) = -2.16$, $p = 0.049$, Cohen's d_{av} [95% CI] = -0.2 [-0.39, -0.001]) to visual targets in the attended hemifield when this hemifield was expected than unexpected (see Supplementary Figure 8.1, right bar plot of Figure 3.1e and Table 3.1). This simple main effect mimics the response profile observed for auditory targets, suggesting that the effect of expectation generalises from audition to vision. By contrast, participants had non-significantly lower hit rates ($t(14) = -1.49$, $p = 0.159$, Cohen's d_{av} [95% CI] = -0.23 [-0.53, 0.09][‡]) and responded significantly more

---

[*] We repeated this simple contrasts using a non-parametric 2-tailed Wilcoxon Signed Ranks: $p = 0.018$
[†] We repeated this simple contrasts using a non-parametric 2-tailed Wilcoxon Signed Ranks: $p = 0.049$

[‡] We repeated this simple contrasts using a non-parametric 2-tailed Wilcoxon Signed Ranks: $p = 0.141$

slowly ($t(14) = 3.21$, $p = 0.006$, Cohen's d_{av} [95% CI] = 0.18 [0.05, 0.30]) to visual targets in the unattended hemifield, when this hemifield was expected than unexpected (see Supplementary Figure 8.1, right bar plot of Figure 3.1e and Table 3.1).

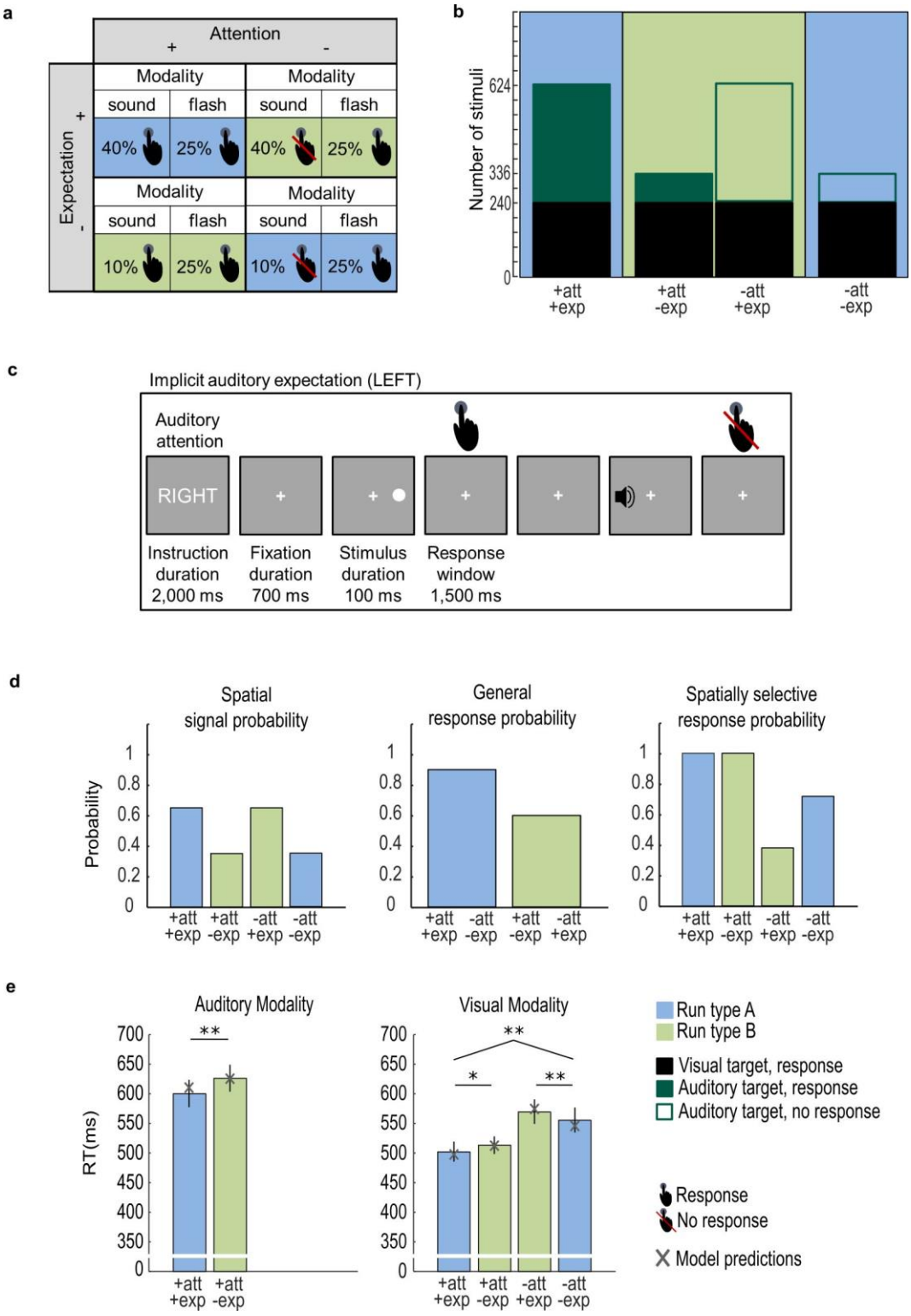


Figure 3.1: Experiment 1: Design, example trial, probabilities, observed and predicted response times.

a. Auditory spatial attention and expectation (i.e. signal probability) were manipulated in a 2 (attended vs. unattended) x 2 (expected vs. unexpected) x 2 (auditory vs. visual) factorial design. Colours indicate run type (blue: attention and expectation are

congruent; green: attention and expectation are incongruent). Presence vs. absence of response requirement is indicated by the hand symbol.

b. Number of auditory (dark green) and visual (black) trials in the 2 (attended vs. unattended) x 2 (expected vs. unexpected) conditions. The bar plots are filled (i.e. response requirement) or not filled (i.e. no response requirement). The fraction of filled area for each bar represents the spatially selective response probability. The fraction of filled area pooled over the two bars of one particular run type (e.g. blue) represents the general response probability in a run; it is greater for the 'blue runs' where attention and expectation are congruent.

c. At the beginning of each run, a cue informed participants whether to attend and respond to auditory signals selectively in their left or right hemifield throughout the entire run. On each trial participants were presented with an auditory or visual stimulus (100 ms duration) either in their left or right hemifield. They were instructed to respond to auditory stimuli only in the 'attended' hemifield and to all visual stimuli irrespective of hemifield as fast and accurately as possible with the same finger. The response window was limited to 1500 ms. Participants were not explicitly informed that auditory signals were more likely in one of the two hemifields. Instead, spatial expectation (i.e. spatial signal probability) was implicitly learnt over runs.

d. The bar plots show i. spatial signal probability: the probability that a signal (pooled over visual and auditory modalities) was presented in a particular hemifield, ii. general response probability: the probability that a signal required a response in a particular (green or blue) run type regardless of hemifield, iii. spatially selective response probability: the probability that a signal required a response conditioned on that it was presented in a particular hemifield.

e. The bar plots show across subjects' mean (±SEM) response times for each of the six conditions with response requirements. The brackets and stars indicate significance of main effects and interactions. * $p < 0.05$; ** $p < 0.01$. The grey crosses show the condition-specific response times predicted by the 'winning' $\text{Model}_{\text{Log}}$ 2 that includes attention, general response probability and spatially selective response probability as predictors.

3.2.2 Experiment 2

For the auditory modality, the two-sided paired-sample t-tests showed slower responses ($t(20) = 3.43$, $p = 0.003$, Cohen's d_{av} [95% CI] = 0.21 [0.07, 0.35]) in the attended hemifield, when this hemifield was expected than unexpected (see left bar plot of Figure 3.2e and Table 3.1).

For the visual modality, the 2 (attended vs. unattended) x 2 (expected vs. unexpected) rmANOVA on response times revealed a significant main effect of attention ($F(1, 20) = 62.58$, $p < 0.001$, η_p^2 [90% CI] = 0.76 [0.55, 0.83]). As in experiment 1, participants responded faster to visual stimuli in their attended than unattended hemifield. Yet, in contrast to experiment 1, we did not observe a significant interaction between attention and expectation in experiment 2. Instead, the rmANOVA revealed a significant main effect of expectation for both bias ($F(1, 20) = 5.17$, $p = 0.034$, η_p^2 [90% CI] = 0.20 [0.009, 0.42]) and response times ($F(1, 20) = 5.23$, $p = 0.033$, η_p^2 [90% CI] = 0.21 [0.009, 0.42]). Irrespective of whether the hemifield was attended or unattended, participants had a greater bias and responded more slowly to visual stimuli in the expected than unexpected spatial hemifield, again mimicking the effect of expectation in the auditory modality (see Supplementary Figure 8.1, right bar plot of Figure 3.2e and Table 3.1).

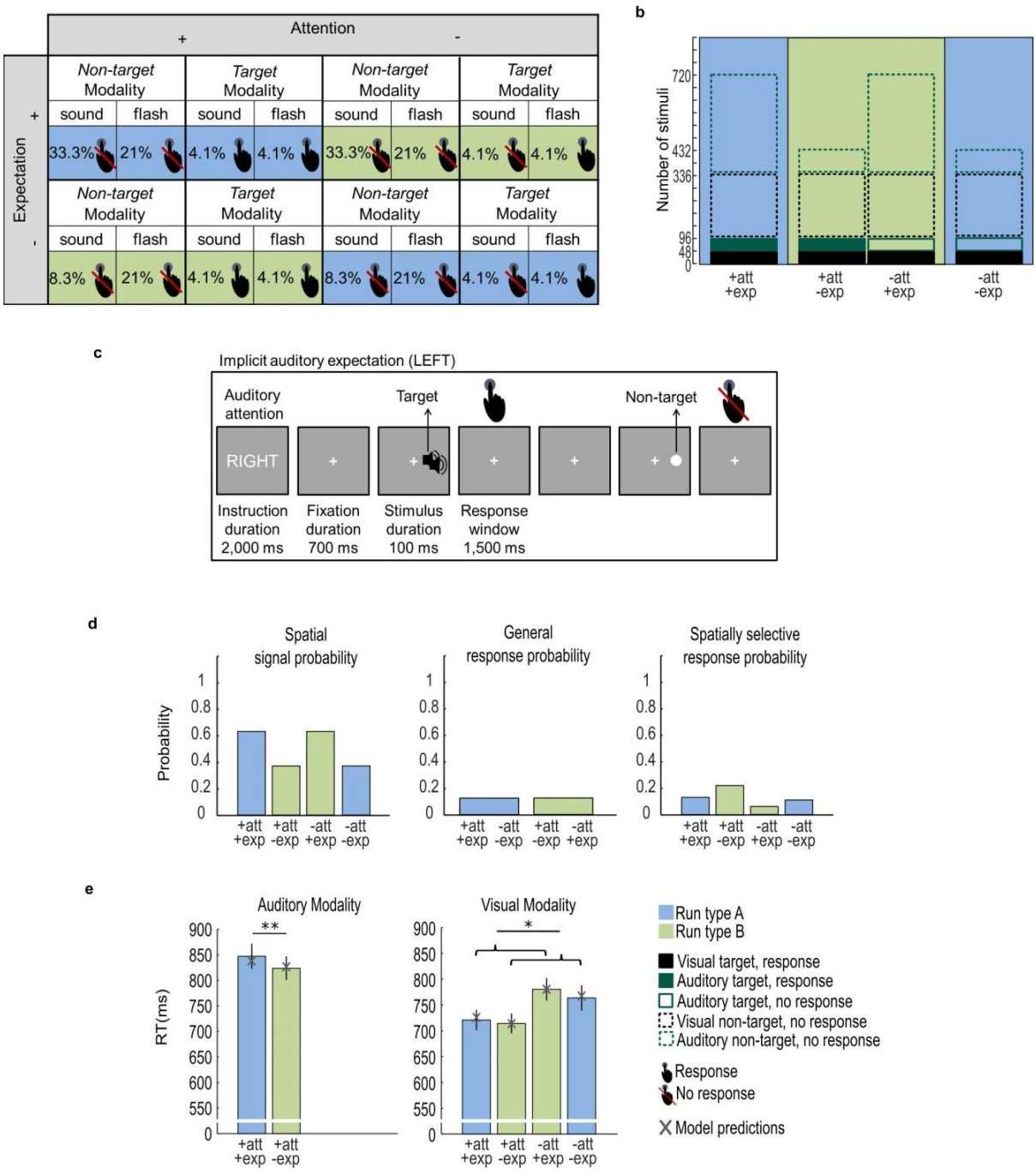


Figure 3.2: Experiment 2: Design, example trial, probabilities, observed and predicted response times.

a. Auditory spatial attention and expectation (i.e. signal probability) were manipulated in a 2 (attended vs. unattended) x 2 (expected vs. unexpected) x 2 (auditory vs. visual) x 2 (target vs. non-target) factorial design. Colours indicate run type (blue: attention and expectation are congruent; green: attention and expectation are incongruent). Presence vs. absence of response requirement is indicated by the hand symbol.

b. Number of auditory (dark green) and visual (black) trials in the 2 (attended vs. unattended) x 2 (expected vs. unexpected) conditions. The bar plots are filled (i.e. response required) or not filled (i.e. no response required). Their contours indicate target vs. non-target trials (solid = target; dotted = non-target). The fraction of filled area for each bar represents the spatially selective response probability, which is greater in the unexpected than expected conditions.

c. At the beginning of each run, a cue informed participants whether to attend and respond to auditory targets (i.e. double noise burst) selectively in their left or right hemifield throughout the entire run. On each trial participants were presented with an auditory or visual stimulus (100 ms duration) either in their left or right hemifield. They were instructed to respond to auditory targets (i.e. double noise burst) only in the 'attended' hemifield and to all visual targets (i.e. double flash) irrespective of hemifield as fast and accurately as possible with the same finger. The visual or auditory non-targets never required any response. The response window was limited to 1500 ms. Participants were not explicitly informed that auditory signals were more likely in one of the two hemifields. Instead, spatial expectation (i.e. spatial signal probability) was implicitly learnt over runs.

d. The bar plots show i. spatial signal probability: the probability that a signal (pooled over visual and auditory modalities) was presented in a particular hemifield, ii. general response probability: the probability that a signal required a response in a particular (green or blue) run type, iii. spatially selective response probability: the probability that a signal required a response conditioned on that it was presented in a particular hemifield.

e. The bar plot shows across subjects' mean (±SEM) response times for each of the six conditions with response requirements. The brackets and stars indicate significance of main effects and interactions. * $p < 0.05$; ** $p < 0.01$. The grey crosses show the condition-specific response times predicted by the 'winning' $\text{Model}_{\text{Log}}$ 2 that includes attention, general response probability and spatially selective response probability as predictors.

Table 3.1: Group mean reaction times (RT) and hit rates for each stimulus modality in each condition for experiment 1 and experiment 2.

D-prime and bias for the visual modality in each condition for experiment 2. Please note that d-prime and bias could be computed neither for experiment 1 (because there were no non-targets included in the paradigm) nor for auditory modality of experiment 2 (because participants did not make any false alarms in the +att-exp condition). Note. Standard errors of the mean (SEM) are given in parentheses.

	Auditory modality		Visual modality			
	+att +exp	+att -exp	+att +exp	+att -exp	-att +exp	-att -exp
Experiment 1						
RT (ms) (SEM)	599.9 (22.8)	626.8 (22.8)	502.5 (16.1)	514.2 (14.4)	567.3 (20.6)	553.2 (21)
Hit rates (%) (SEM)	99.4 (0.2)	98.4 (0.6)	99.5 (0.2)	98.9 (0.4)	98.8 (0.4)	99.2 (0.3)
Experiment 2						
RT (ms) (SEM)	847.2 (24.6)	823.8 (23.0)	720.8 (19.8)	714.2 (19.2)	780.4 (21.7)	763.7 (24.5)
Hit rates (%) (SEM)	97.6 (0.8)	98.3 (0.6)	96 (0.9)	97 (0.8)	96.2 (0.9)	96.4 (0.9)
D-prime (SEM)	/	/	4.08 (0.12)	4.07 (0.15)	4.25 (0.12)	4.12 (0.15)
Bias (SEM)	/	/	0.19 (0.07)	0.08 (0.05)	0.24 (0.06)	0.17 (0.05)

In summary, even though both experimental paradigms orthogonally manipulated spatial attention and expectation operationally defined as signal probability, they led us to strikingly different conclusions. Experiment 1 suggests that attention and expectation act in an interactive fashion: under spatial attention, participants responded faster to auditory and visual targets in their expected hemifield, but this effect was reversed for visual targets when attention is diverted to the other hemifield. By contrast, experiment 2 suggests that spatial attention and expectation influence response times in an additive fashion. Here, participants were generally slower when responding to targets in the expected than unexpected hemifield irrespective of attention. Not only does this main effect of expectation in experiment 2 contradict the findings from experiment 1, but it may also be surprising given the vast literature suggesting that expectation facilitates processing[3,8,33]. These contradictory results across the two experiments suggest that modelling the experiments as a 2 (attended vs. unattended) x 2 (expected vs. unexpected) factorial design may not provide a coherent explanatory framework.

Key to resolving these seemingly contradictory results is the realization that spatial attention and signal probability jointly determine two additional probabilities critical for perceptual decision making: 1. the general response probability, i.e. the probability that a response is required on a given trial irrespective of where the stimulus is presented and 2. the spatially selective response probability, i.e. the probability that a response is required conditioned on a stimulus being presented in a particular hemifield (see Figure 3.1d and 3.2d and Supplementary Table 8.1). In the following we therefore investigate whether models that include general and spatially selective response probability outperform the traditional factorial model of attention

and expectation in accounting for observers' response times jointly across both experiments.

3.2.3 Joint models of experiment 1 and experiment 2

We generated a 3 (combination of fixed effects predictors) x 2 (linear vs. log-transform of probability values in the predictors) model space (see Methods section for details and Figure 3.3). All models assume that the observer prioritises processing in hemifields that are explicitly attended via task instructions per se, yet they differ in whether expectation (i.e. spatial signal probability), general response probability or spatially selective response probability are included as additional explanatory variables. Model 1 is the conventional factorial model that allows for additive and interactive effects of attention and signal probability. Model 2 assumes that signal probability per se does not directly affect response times, but only indirectly by co-determining the general and spatially selective response probability. Model 3 is based on Model 2, but it is more complex by allowing signal probability to influence response times not only indirectly, by co-determining general and spatially selective response probability, but also directly. In addition, given previous research showing a non-linear relationship between probabilities and response times[34], we manipulated whether probabilities predict response times in a linear (i.e. $\text{Model}_{\text{Lin}}$) or log-transformed (i.e. $\text{Model}_{\text{Log}}$) fashion.

Figure 3.3 shows the Bayesian Information Criterion values for the six models in a factorial matrix (Bayesian Information criterion (BIC): $\text{Model}_{\text{Lin}}$ 1: -690; $\text{Model}_{\text{Lin}}$ 2: -694; $\text{Model}_{\text{Lin}}$ 3: -691; $\text{Model}_{\text{Log}}$ 1: -690; $\text{Model}_{\text{Log}}$ 2: -701; $\text{Model}_{\text{Log}}$ 3: -697; n.b. a smaller value indicates a better model). First, we show that log-transformed probabilities are better predictors for response times in our target detection task than probabilities per se[34]. Second and more importantly, $\text{Model}_{\text{Log}}$ 2 that combines spatial

attention with general and spatially selective response probability outperformed both the conventional factorial model as well as a more elaborate model that also includes spatial signal probability as a predictor by itself (see Figure 3.3). More specifically, Bayesian model comparison provided strong evidence (i.e. increase in BIC of approx. 10) for the winning $\text{Model}_{\text{Log}}$ 2 relative to the conventional $\text{Model}_{\text{Log}}$ 1 and $\text{Model}_{\text{Log}}$ 3. Figures 3.1e and 3.2e show the response time predictions of the winning $\text{Model}_{\text{Log}}$ 2 as grey crosses together with observed condition-specific response times. In contrast to the conventional Model 1, $\text{Model}_{\text{Log}}$ 2 can flexibly model that response times in the attended hemifield are faster for expected than unexpected stimuli in experiment 1, but slower in experiment 2. Critically, adding spatial signal probability in Model 3 did not increase the model evidence. These results suggest that expectation (i.e. spatial signal probability) influences response times mainly indirectly by co-determining general and spatially selective response probabilities.

One may argue that the spatially selective response probability already accounts for observers' allocation of attention over space. However, a seventh model, i.e. the winning $\text{Model}_{\text{Log}}$ 2 without the attention predictor, obtained a BIC of only -686 suggesting that endogenous attention influences perceptual decisions as an independent predictor above and beyond spatially selective response probability.

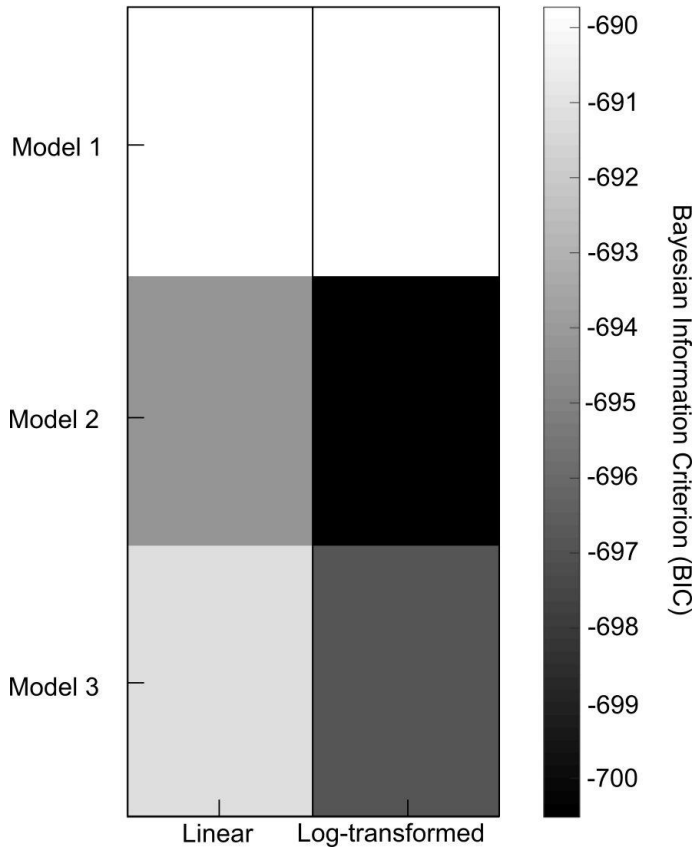


Figure 3.3: Factorial model space and Bayesian Information Criterion (BIC) for the joint analysis of experiment 1 and experiment 2.

3 (combination of fixed effects predictors) x 2 (probability values in predictors: linear vs. log-transformed) model space. Predictors included in i. conventional Model 1: attention, spatial signal probability and attention x spatial signal probability interaction, ii. Model 2: attention, general response probability and spatially selective response probability, iii. Model 3: attention, spatial signal probability, general response probability and spatially selective response probability. To account for potential non-linear relationships between probabilities and response times, the probabilities predicted response times in a linear or a log-transformed fashion. The matrix shows the Bayesian Information Criterion (BIC) of the 6 models (smaller BIC, i.e. darker shade, indicates better model).

3.3 DISCUSSION

Prior expectations and attention are two key determinants controlling perceptual inference. The current study developed a novel multisensory approach to dissociate additive and interactive effects of attention and expectation at the behavioural level. Manipulating attention and signal probability selectively in the auditory modality, we evaluated their effects on signal detection in the auditory and visual modalities. In each experiment, we observed qualitatively equivalent effects in audition and vision for signal probability in the attended hemifield demonstrating the effectiveness of multisensory generalization in our experiments[27,28]. Critically however, even though both experiments manipulated spatial attention and signal probability orthogonally, they revealed either interactive (experiment 1) or additive (experiment 2) effects.

In experiment 1, we observed synergistic interactions between spatial attention and signal probability at the behavioural level consistent with previously reported interactive effects at the neural level in an equivalent experimental paradigm[31]. When stimuli were attended, participants responded faster to signals presented in the expected than unexpected hemifield, which is in line with the general notion that prior expectations facilitate perceptual processing[35]. Yet, when signals were unattended, stimuli presented in the expected hemifield were associated with slower response times than in the unexpected hemifield. This crossover interaction between spatial attention and signal probability can be explained by the effects of general response probability (and spatially selective response probability). As shown in figure 3.1d (middle bar plot) and in Supplementary Table 8.1, in runs (blue) where attention and expectation are directed to the same spatial hemifield, participants are required to respond on 90%

of the trials as compared to 60% of the trials in runs when attention and expectation are directed to different hemifields.

Experiment 2 manipulated spatial signal probability via auditory non-targets that never required any response. As a consequence, the general response probability was constant across conditions (see middle bar plot of Figure 3.2d, Supplementary Table 8.1 and [32] for related design). Further, increasing spatial signal probability via an increase in the number of non-targets for expected relative to unexpected hemifields (see Figure 3.2a and 3.2b) necessarily decreases the spatially selective response probability (right bar plot of Figure 3.2d). Consistent with the profile of the spatially selective response probability, but contrary to conventional views of expectation[36], signals presented in the expected relative to unexpected hemifield were generally associated with slower response times. Put simply, the neural system does not facilitate target detection in the spatial hemifield where many events occur (i.e. high signal probability, left bar plot of Figure 2d), but in the hemifield where a high percentage of those events requires a response (i.e. high spatially selective response probability, right bar plot of Figure 2d).

In line with this qualitative explanation, formal Bayesian model comparison confirmed that the response time profiles across the two experiments can be accounted for parsimoniously by $\text{Model}_{\text{Log}}$ 2 that predicates response times on i. spatial attention, ii. general response probability and iii. spatially selective response probability. Critically, adding spatial signal probability per se as a predictor did not substantially increase the model evidence ($\text{Model}_{\text{Log}}$ 3, Figure 3.3). These results suggest that spatial signal probability influences response times mainly indirectly via general and spatially selective response probabilities rather than as an independent factor.

Our findings indicate that the interactions between spatial attention and signal probability previously reported at the neural level[31] are behaviourally relevant, yet they offer a new perspective: the interactions may represent 'response related' expectations encoding the general and spatially selective response probabilities. Observers concurrently form multiple sorts of expectations such as the expectation of a signal (i.e. signal probability), the expectation to make a response irrespective of (i.e. the general response probability) or dependent on (i.e. the spatial selective response probability) a particular signal location. Collectively, these expectations encode the probabilistic structure of sensory inputs and demands within the entire perception-action loop[37,38]. This sheds new light on the interplay of spatial attention and expectation. Current operational definitions of attention as task-relevance / response requirement and expectation as signal probability[1,35] set up an artificial dichotomy between attention and signal probability. Yet signal probability and task-relevance are intimately linked by co-determining general and spatially selective response probabilities that shape response-related expectations and allocation of attentional resources. In the next step we therefore asked whether observers' spatial allocation of attentional resources (i.e. main effect of attention on response times) is already accounted for by response-related expectations (i.e. general and spatially selective response probability). Using Bayesian model comparison we show that a more parsimonious model without spatial attention as an independent predictor performed substantially worse than the winning model ($\text{Model}_{\text{Log}}$ 2). These results suggest that endogenous spatial attention can at least to some extent influence perceptual decisions above and beyond spatially-selective response probability. In line with this conjecture, observers have been shown to allocate attentional resources based on instructions alone[39,40], even if response requirements are kept constant thereby dissociating

attention and task-relevance. In order to dissociate additive and interactive effects of endogenous spatial attention (independent from task-relevance) and signal-related expectations at the neural level, future studies may manipulate endogenous spatial attention via instructions alone whilst holding the response requirements constant. Critically, our experimental results also reveal that the dichotomy between spatial expectation and spatial attention when operationally defined as response requirement over space may be useful from an initial analytical perspective. However, signal probability and attention as task-relevance are intimately linked by co-determining response-related expectations. Thus, changes in spatial expectation will inherently alter general or spatially selective response probabilities and therefore influence participant's allocation of attentional resources over spatial hemifields.

From a cognitive perspective, changes in general response probability may be associated with increases in alertness, arousal or motor preparation[37,38] leading to response facilitation. By contrast, spatially selective response probability needs to engage spatially selective mechanisms. For instance, the neural system may increase the precision of representations at spatial locations associated with a greater spatially selective response probability. Alternatively, observers may bias their decisional processes by shifting the starting point closer to the decisional boundary or lowering the decisional boundary of the evidence accumulation process[41,42]. In support of the latter, our signal detection theoretic analysis of response choices showed an increase in hits for higher general response probability in experiment 1 and a decrease in bias for higher spatially selective response probability in experiment 2. Critically, as participants always respond with the same effector organ (i.e. index finger of the dominant hand) to all visual and task-relevant auditory signals across both hemifields,

decisional processes for all spatial locations need to finally map onto the same effector organ.

Cast in the predictive coding framework our results suggest that the brain iteratively adjusts its predictions of the sensory inputs at multiple levels across the cortical hierarchy. Critically, spatial attention as task-relevance profoundly modulates the precision of prediction errors and thereby the gain with which they impact higher cortical levels that are critical for response selection: First, the precision or gain of prediction errors is higher for sessions where many stimuli require a response (i.e. sessions with high general response probability). Second, the precision of prediction errors are selectively optimized for spatial locations associated with a high spatially selective response probability, i.e. locations where a high percentage of signals require a response. Guided by current task demands the brain thus adaptively optimizes the precision of spatial representations at locations that are critical for effective interactions.

Future neuroimaging and neurophysiological research will need to investigate whether the facilitatory effects of endogenous attention, signal probability, general and spatially selective response probability are mediated by similar or distinct neural mechanisms. For instance, they may be reflected in pre-stimulus baseline shifts or emerge only during post-stimulus processing (e.g. enhanced precision of spatial representations). Further, from a multisensory perspective we need to determine whether putative pre-stimulus baseline shifts, biases and increases in precision of spatial representations emerge only at higher levels of the cortical hierarchy associated with multisensory representations of space or in primary sensory areas across vision and audition either directly or via feed-back from higher order areas[43,44,45].

In conclusion, our results across two experiments demonstrate that spatial attention and signal probability can influence signal detection either interactively or additively. These seemingly contradictory results can be reconciled by a new model that explains response times parsimoniously by spatial attention, general and spatially selective response probabilities. Our model provides a novel perspective on the intricate interplay of attention, signal- and response-related expectations in perceptual decisions, which is critical for effective interactions with the environment.

3.4 GENERAL METHODS

3.4.1 Apparatus

During the experiment, participants rested their chin on a chinrest with the height held constant across all the participants. Auditory stimuli were presented at approximately 72 dB SPL, via HD 280 PRO headphones (Sennheiser, Germany). Visual stimuli were displayed on a gamma-corrected LCD monitor (2560 x 1600 resolution, 60 Hz refresh rate, 30" Dell UltraSharp U3014, USA), at a viewing distance of approximately 50 cm from the participant's eyes. Stimuli were presented using Psychtoolbox version 3[46,47] (www.psychtoolbox.org), running under Matlab R2014a (Mathworks Inc., Natick, MA, USA) on a Windows machine. Participants responded to all stimuli with the same index finger of their reported dominant hand. Responses were recorded via one key of a small keypad (Targus, USA). Throughout the experiment, participants' eye-movements and fixations were monitored using Tobii Eyex eyetracking system (Tobii, Sweden).

3.4.2 Stimuli

Spatial auditory stimuli (100 ms duration) were created by convolving a burst of white noise (with 5 ms onset and offset ramps) with spatially selective head-related transfer functions (HRTFs) based on the KEMAR dummy head of the MIT Media Lab[48] (http://sound.media.mit.edu/resources/KEMAR.html). Visual stimuli (i.e. the so-called 'flashes') were white discs (radius: 0.88º visual angle, luminance: 196 cd/m2, 100 ms duration) presented on a grey background. Both auditory and visual stimuli were presented at ±10º of visual angle along the azimuth (0º of vertical visual angle). A fixation cross was presented in the centre of the screen throughout the entire experiment. In Experiment 2, two types of stimuli were presented, *target* and *non-targets*. Targets were identical to the non-targets, except that a brief gap (10 ms) was inserted after 45 ms.

Prior to the main experiment, participants were tested for their ability to discriminate left and right auditory stimuli in a brief series of 20 trials in a 2-alternative forced choice task. Participants indicated their spatial discrimination response (i.e. 'left' vs. 'right') via a two choice key press (group mean accuracy was 99.3% ± 1.7% SD in Experiment 1 and 98% ± 2.2% SD in Experiment 2).

3.4.3 Experiment 1

3.4.3.1 Participants

Sixteen healthy subjects (5 males, 11 females; mean age = 20.06 years; range 18-27 years; 15 right-handed) participated in experiment 1. The sample size was determined based on previous studies investigating attention/expectation[2-4,6,8-10,31,32] and/or multisensory integration[27,28]. All participants had normal or corrected to normal vision and reported normal hearing. In experiment 1, one participant was excluded

post-hoc from the analysis because the overall performance accuracy on the target detection task was lower than 2 SD of the group mean (i.e. the across subjects' mean = 98.2% correct). All participants provided written informed consent and were naïve to the aim of the experiment. The study was approved by the local ethics committee of the University of Birmingham (Science, Technology, Mathematics and Engineering (STEM) Ethical Review Committee) and the experiment was conducted in accordance with these guidelines and regulations.

3.4.3.2 Design and procedure

Experiment 1 investigated the effect of auditory spatial attention and expectation on detection of auditory and visual targets using a 2 (*auditory spatial attention*: left vs. right hemifield) x 2 (*auditory spatial expectation*: left vs. right hemifield) x 2 (*stimulus modality*: auditory vs. visual) x 2 (*stimulus location*: left vs. right hemifield) factorial design (see Figure 3.1a). For the design figure, analysis and results, we pooled over the factor 'stimulus location' to provide a more succinct 2 (*attention*: attended vs. unattended hemifield) x 2 (*expectation*: expected vs. unexpected hemifield) x 2 (*stimulus modality*: auditory vs. visual) design. Auditory spatial expectation was manipulated as auditory spatial signal probability in the left and right hemifields across experimental sessions that were performed on different days. Observers were not informed about those probabilities, but learnt them implicitly. Auditory spatial attention was manipulated as 'task-relevance', i.e. the requirement to respond to an auditory target in the left vs. right hemifield over runs of 80 trials. Prior to each run a cue (duration: 2000 ms) informed the observer whether to respond to auditory targets either in the left or right hemifield. Critically, spatial attention and expectation were manipulated only in audition but not vision. Hence, observers needed to respond to all visual targets that were presented with equal probabilities in either spatial hemifield.

Each trial (SOA: 2300 ms) included three time windows (see Figure 3.1c): i. fixation cross alone (700 ms duration), ii. brief flash or sound (stimulus duration: 100 ms) and iii. fixation cross alone (1500 ms as response interval). Observers responded to the auditory stimuli in the attended hemifield and to the visual stimuli via key press with the same index finger (i.e. the same response for all auditory and visual stimuli) as fast and accurately as possible. They fixated the cross in the centre of the screen that was presented throughout the entire experiment with their fixation performance monitored via eye tracking.

Two sessions (i.e. spatial expectation left vs. right on different days) included 12 attention runs with 80 trials each. Runs were of two types: in run type A (coded in blue in Figure 3.1a and 3.1b) spatial attention and expectation were congruent (i.e. spatial attention was directed to the hemifield with higher auditory target frequency); in run type B (coded in green) spatial attention and expectation were incongruent (i.e. attention was directed to the hemifield with less frequent auditory stimuli). Thus, in total the experiment included 80 trials x 12 attention runs (6 runs of type A and 6 runs of type B) x 2 expectation sessions = 1920 trials in total. Specifically, each run type overall included 384 auditory targets for the expected hemifield (pooled over left and right) and 96 auditory targets for the unexpected hemifield (pooled over left and right). Each run type also included 240 visual targets for the expected hemifield and 240 visual targets for the unexpected hemifield (pooled over left and right). For further details see Figure 3.1b which shows the absolute number of trials for each condition and run type.

The order of expectation (i.e. left vs. right) sessions was counterbalanced across participants; the order of attention runs was counterbalanced within and across participants and the order of stimulus location and stimulus modality were pseudo-

randomized within each participant. Brief breaks were included after every two attention runs to provide feedback to participants about their performance accuracy (averaged across all conditions) in the target detection task and about their eye-movements (i.e. fixation maintenance).

Prior to each session, participants were familiarized with the stimuli in brief practice runs (with equal spatial signal probability) and trained on target detection performance and fixation (i.e. a warning signal was shown when the disparity between the central fixation cross and the eye-data samples exceeded 2.5 degrees). After the final session participants indicated in a questionnaire whether they thought the sound or the flash were presented more frequently in one of the two spatial hemifields. Fifteen out of the total 16 participants correctly reported that the auditory stimuli were more frequent in one hemifield and all participants reported the visual stimuli to be equally frequent across the two hemifields, suggesting that most participants were aware of the manipulation of signal probability in experiment 1.

3.4.4 Experiment 2

3.4.4.1 Participants

Twenty-four new healthy subjects (5 males, 19 females; mean age = 20.54; range 18-40 years; 20 right-handed) took part in experiment 2. The sample size was increased compared to experiment 1, because the number of targets (i.e. trials requiring a response) in the condition with the minimal number of targets was half the number of targets in the same condition of experiment 1 (48 versus 96, compare Figure 3.1b and 3.2b). All participants had normal or corrected to normal vision and reported normal hearing. Three participants were excluded post-hoc from the analysis because their overall performance accuracy on the target detection task was lower than 2 SD of the group mean (i.e. the across subjects' mean = 98% correct). All participants provided

written informed consent and were naïve to the aim of the experiment. The study was approved by the local ethics committee of the University of Birmingham (Science, Technology, Mathematics and Engineering (STEM) Ethical Review Committee) and the experiment was conducted in accordance with these guidelines and regulations.

3.4.4.2 Design and procedure

Experiment 2 again investigated the effect of auditory spatial attention and expectation on detection of auditory and visual targets using a 2 (*auditory spatial attention*: left vs. right hemifield) x 2 (*auditory spatial expectation*: left vs. right hemifield) x 2 (*stimulus modality*: auditory vs. visual) x 2 (*stimulus location*: left vs. right hemifield) x 2 (*stimulus type*: target vs. non-target) factorial design (see Figure 3.2a). For the design figure, analysis and results, we pooled over the factor 'stimulus location' to provide a more succinct 2 (*attention*: attended vs. unattended hemifield) x 2 (*expectation*: expected vs. unexpected hemifield) x 2 (*stimulus modality*: auditory vs. visual) x 2 (*stimulus type*: target vs. non-target) design.

As in experiment 1, we manipulated spatial attention and expectation in audition alone and investigated its effect on response times to auditory and visual stimuli. Critically, in contrast to experiment 1, experiment 2 manipulated spatial expectation not directly via the frequency of auditory targets, but indirectly via the frequency of auditory non-targets across the two hemifields on different days. Thus, experiment 2 included two sets of stimuli for both auditory and visual modalities: targets and non-targets. The non-targets were identical to the stimuli in experiment 1. The targets were flashes or sounds with a 10 ms gap introduced after 45 ms, but otherwise identical to the non-targets. The non-targets never required a response in any sensory modality irrespective of hemifield. They were introduced to manipulate auditory stimulus expectation, i.e. auditory stimulus probability across the two hemifields, such that the

general response probability was held constant across all conditions[32] (see Figure 3.2a and below for further explanation). Again observers were not informed about those probabilities, but learnt them implicitly. Auditory spatial attention was manipulated as 'task-relevance', i.e. the requirement to respond to auditory targets in the left vs. right hemifield over runs of 96 trials. Prior to each run a cue (duration: 2000 ms) informed observers whether to respond to auditory targets either in the left or right hemifield. Critically, spatial attention and expectation were manipulated only in audition but not vision. Hence, observers needed to respond to all visual targets that were presented with equal probabilities in both spatial hemifields.

As shown in Figure 3.2c, the sequence and timing of a trial were identical to experiment 1. In experiment 2, observers responded only to the auditory targets (i.e. double sound) in the attended hemifield and the visual targets (i.e. double flash) via key press with the same index finger (i.e. the same response for all auditory and visual targets) as fast and accurately as possible.

As in experiment 1, two sessions (i.e. spatial expectation left vs. right on different days) included 12 attention runs with 96 trials each. Runs were of two types: in A (coded in blue in Figure 3.2a and 3.2b) spatial attention and expectation were congruent (i.e. spatial attention was directed to the hemifield with higher auditory stimuli frequency), whereas in run B (coded in green) spatial attention and expectation were incongruent (i.e. attention was directed to the hemifield with less frequent auditory stimuli). Thus, in total the experiment included 96 trials x 12 attention runs (6 run of type A and 6 runs of type B) x 2 expectation sessions = 2304 trials in total. Specifically, each run type overall included 384 auditory non-targets for the expected hemifield (pooled over left and right) and 96 auditory non-targets for the unexpected hemifield (pooled over left and right). Each run type also included 240 visual non-

targets for the expected hemifield and 240 visual non-targets for the unexpected hemifield (pooled over left and right). Further, each run type included 48 auditory and 48 visual targets for the expected hemifield and 48 auditory and 48 visual targets for the unexpected hemifield. For further details see Figure 3.2b which shows the absolute number of trials for each condition and run.

The procedural details (e.g. counterbalancing, eye tracking, post-questionnaire) were otherwise comparable to experiment 1. Eight out of 24 participants correctly reported that the auditory stimuli were more frequent in one hemifield and 23 out of 24 participants reported the visual stimuli to be equally frequent across the two hemifields, suggesting that most participants were not explicitly aware of the manipulation of signal probability in experiment 2. Most likely, participants were less aware of the expectation manipulation in experiment 2 than experiment 1 because signal probability was manipulated only for the task-irrelevant non-targets. Further, the inclusion of target and non-target stimuli in the same paradigm increased the stimulus variability making the expectation manipulation less apparent.

3.4.5 Experiments 1 and 2: Spatial signal, general response and spatially selective response probability

Both experiments 1 and 2 orthogonally manipulated spatial attention (i.e. response requirement) and expectation (i.e. spatial signal probability). Yet, while experiment 1 manipulated spatial signal probability (i.e. the probability of a signal irrespective of sensory modality in the left or right hemifield) directly via the probability of task-relevant auditory targets across hemifields (see left bar plot of Figure 3.1d and Supplementary Table 8.1), experiment 2 manipulated it indirectly via additional task-irrelevant auditory non-targets that never required a response (see left bar plot of Figure 3.2d and Supplementary Table 8.1). As a result, the two experiments

were associated with different profiles of i. the general response probability (i.e. the probability that the observer needs to make a response irrespective of the hemifield in which the signal is presented, see middle bar plot of Figure 3.1d and 3.2d and Supplementary Table 8.1) and ii. the spatially selective response probability (i.e. the probability that the observer needs to make a response conditioned on that the signal is presented in a particular hemifield, see right bar plot of Figure 3.1d and 3.2d and Supplementary Table 8.1).

In experiment 1, the general response probability is greater in the run type A (coded in blue in the middle bar plot of Figure 3.1d), where spatial attention and expectation are congruent (i.e. spatial attention is directed to the hemifield with higher auditory target frequency), than in run type B (coded in green), where spatial attention and expectation are incongruent (i.e. attention is directed to the hemifield with less frequent auditory stimuli). Likewise, the spatially selective response probability changes across conditions co-determined by auditory spatial attention and spatial signal probability (right bar plot of Figure 3.1d).

Experiment 2 manipulated auditory spatial expectation via additional non-targets that never require a response. As a consequence, the general response probability is held constant throughout the entire experiment (see middle bar plot of Figure 3.2d). Further, the increase in spatial signal probability in the expected hemifield inherently decreases the spatially selective response probability (see right bar plot of Figure 3.2d).

3.4.6 Experiment 1 and 2: Data analysis

3.4.6.1 Eye movement analysis

We excluded trials where participants did not successfully fixate the central cross based on a dispersion criterion (i.e. distance of fixation from subject's centre of fixation as defined in calibration trials > 1.3 degrees for three subsequent samples[49])

(an additional analysis including all trials yielded basically equivalent results). Our eye tracking data confirmed that participants successfully maintained fixation in both experiments with only a small number of trials to be excluded (experiment 1: excluded auditory response trials 2.9% ± 1% [across subjects mean ± SEM]; excluded visual response trials 2.8% ± 1% [across subjects mean ± SEM]; experiment 2: excluded auditory response trials 4.4% ± 1.6% [across subjects mean ± SEM]; excluded visual trials 4.6% ± 1.5% [across subjects mean ± SEM]).

3.4.6.2 Response choice and time analysis

Using signal detection theory we reported hit rates for experiment 1 and experiment 2 and d-prime (Z(hit rate) - Z(false alarm rate)) as well as bias (-0.5[Z(false alarm rate) + Z(hit rate)]) for the visual modality for experiment 2. D-prime and bias could not be computed for experiment 1 (because non-target trials were not included in the paradigm) and for auditory modality of experiment 2 (because participants did not make any false alarms in the +att-exp condition). Response time analysis was limited to correct trials only and response times within the range of participant- and condition-specific mean ± two SD and <1500 ms.

General linear models fitted independently for each experiment: Inference was initially made separately for each experiment at the random effects level to allow for generalization to the population. For auditory targets in the attended hemifield, condition-specific hit rate (experiment 1), d-prime and bias (experiment 2), and median response times (experiment 1 and 2) for each subject were entered into a two-sided paired-sample t-tests with *expectation* (expected, unexpected) as factor. For visual targets, condition-specific hit rate (experiment 1), d-prime and bias (experiment 2), and median response times (experiment 1 and 2) for each subject were entered into a 2 (*attention*: attended, unattended) x 2 (*expectation*: expected, unexpected) repeated

measures analysis of variance (ANOVA). Please note that the computation of d-prime and bias was not possible for experiment 1, because non-target trials were not included in the paradigm. Unless otherwise indicated, we only report effects that are significant at $p < 0.05$.

Generalized mixed effects model fitted jointly to both experiments: To identify a model that parsimoniously explains the data jointly from both experiments, we used generalized mixed effects models across both experiments and Bayesian model comparison. We specified six generalized mixed effects models that were organized in a 3 (combination of *predictors,* as in Figure 3.1d and 3.2d and Supplementary Table 8.1) x 2 (*probabilities*: linear vs. log-transformed) factorial model space. The first factor specified the set of fixed effects predictors:

Traditional Model 1: i. spatial attention: categorical dummy variable encoding whether or not stimulus is presented in the attended hemifield, ii. spatial expectation: signal probability of the hemifield where the stimulus is presented, iii. interaction between spatial attention and expectation.

Model 2: i. spatial attention: categorical dummy variable encoding whether or not stimulus is presented in attended hemifield, ii. general response probability: probability that the participants need to respond to the stimulus prior to knowing where the stimulus is presented, i.e. this probability does not depend on the spatial location of the stimulus, iii. spatially selective response probability: probability that the participants need to respond given that the stimulus is presented in a particular hemifield, i.e. this probability depends on spatial location.

Model 3: Model 2 with the added regressor iv. spatial expectation: signal probability of the hemifield where the stimulus was presented.

The second factor specified whether the probabilities encoded in the regressor predicted response times in a linear or log-transformed fashion to accommodate previous evidence that response times may be non-linearly related to stimulus probability[34].

All six models also included a fixed effects regressor encoding the sensory modality (i.e. auditory vs. visual) of the stimulus. Furthermore, they modelled subject and experiment as random effects to account for the higher order organization into two experiments.

We fitted these six models jointly to the data from experiment 1 and 2 using maximum likelihood estimation and compared the non-nested models using the Bayesian information criterion (BIC) as an approximation to the model evidence. We fitted the winning model again to the data from both experiments using restricted maximum likelihood estimation to obtain unbiased estimates for generating the model predictions shown as grey crosses in Figure 3.1e and 3.2e (see guidelines in [50])

.

REFERENCES

1. Summerfield, C. & Egner, T. Expectation (and attention) in visual cognition. *Trends Cogn. Sci.* **13,** 403–409 (2009).

2. Coull, J.T. & Nobre, A.C. Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *J. Neurosci* **18,** 7426–7435 (1998).

3. Geng, J.J. & Behrmann, M. Probability cuing of target location facilitates visual search implicitly in normal participants and patients with hemispatial neglect. *Psychol. Sci.* **13(6),** 520–525 (2002).

4. Geng, J.J. & Behrmann, M. Spatial probability as an attentional cue in visual search. *Percept. Psychophys*. **67(7),** 1252–1268 (2005).

5. Doherty, J.R., Rao, A., Mesulam, M.M., & Nobre, A.C. Synergistic effect of combined temporal and spatial expectations on visual attention. *J. Neurosci.* **25(36),** 8259–8266 (2005).

6. Vossel, S., Thiel, C.M. & Fink, G.R. Cue validity modulates the neural correlates of covert endogenous orienting of attention in parietal and frontal cortex. *Neuroimage* **32,** 1257–1264 (2006).

7. Carrasco, M. Visual attention: the past 25 years. *Vision. Res.* **51,** 1484–1525 (2011).

8. van Ede, F., de Lange, F.P., & Maris, E. Attentional cues affect accuracy and reaction time via different cognitive and neural processes. *J. Neurosci.* **32(30),** 10408–10412 (2012).

9. Wyart, V., Nobre, A.C., & Summerfield, C. Dissociable prior influences of signal probability and relevance on visual contrast sensitivity. *P. Natl. Acad. Sci. U.S.A.*, **109(9),** 3593–3598 (2012).

10. Rohenkohl, G., Gould, I.C., Pessoa, J., & Nobre, A.C. Combining spatial and temporal expectations to improve visual perception. *J. Vis.* **14(4),** 1–13 (2014).

11. Kastner, S., Pinsk, M.A., De Weerd, P., Desimone, R. & Ungerleider, L.G. Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron* **22,** 751–761 (1999).

12. Corbetta, M. & Shulman, G.L. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* **3,** 201–215 (2002).

13. Silver, M.A., Ress, D. & Heeger, D.J. Neural correlates of sustained spatial attention in human early visual cortex. *J. Neurophysiol.* **97,** 229–237 (2007).

14. Stokes, M., Thompson, R., Nobre, A.C. & Duncan, J. Shape-specific preparatory activity mediates attention to targets in human visual cortex. *Proc. Natl. Acad. Sci.* **106,** 19569–19574 (2009).

15. Summerfield, C., Monti, J.M.P., Trittschuh, E.H., Mesulam, M.M. & Egner, T. Neural repetition suppression reflects fulfilled perceptual expectations. *Nat. Neurosci.* **11,** 1004–1006 (2008).

16. den Ouden, H.E.M., Friston, K.J., Daw, N.D., McIntosh, A.R. & Stephan, K.E. A dual role for prediction error in associative learning. *Cereb. Cortex.* **19,** 1175–1185 (2009).

17. Alink, A., Schwiedrzik, C.M., Kohler, A., Singer, W. & Muckli, L. Stimulus predictability reduces responses in primary visual cortex. *J. Neurosci*. **30,** 2960–2966 (2010).

18. Todorovic, A., van Ede, F., Maris, E. & de Lange, F.P. Prior expectation mediates neural adaptation to repeated sounds in the auditory cortex: an MEG Study. *J. Neurosci.* **31,** 9118–9123 (2011).

19. Schiffer, A.M., Ahlheim, C., Wurm, M.F., & Schubotz, R.I. Surprised at all the entropy: hippocampal, caudate and midbrain contributions to learning from prediction errors. *PLoS One* **7(5),** e36445 (2012).

20. Schröger, E., Kotz, S.A. & SanMiguel, I. Bridging prediction and attention in current research on perception and action. *Brain Res*. **1626,** 1–13 (2015).

21. Rao, B.P.N. & Ballard, D.H. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci*. **2,** 79–87 (1999).

22. Friston, K.J. A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **360,** 815–836 (2005).

23. Friston, K.J. The free-energy principle: a unified brain theory?. *Nat. Rev. Neurosci.* **11(2),** 127–138 (2010).

24. Feldman, H. & Friston, K.J. Attention, uncertainty, and free-energy. *Front. Hum. Neurosci.* **4,** 1–23 (2010).

25. Auksztulewicz, R. & Friston, K.J. Repetition suppression and its contextual determinants in predictive coding. *Cortex* **80,** 125–140 (2016).

26. Auksztulewicz, R. & Friston, K.J. Attentional enhancement of auditory mismatch responses: a DCM/MEG study. *Cereb. Cortex* **25,** 4273–4283 (2015).

27. Spence, C. & Driver, J. Audiovisual links in endogenous covert spatial attention. *J. Exp. Psychol. Hum. Percept. Perform.* **22,** 1005–1030 (1996).

28. Spence, C. & Driver, J. Audiovisual links in exogenous covert spatial orienting. *Percept. Psychophys*. **59,** 1–22 (1997).

29. Posner, M.I. Orienting of attention. *Q. J. Exp. Psychol.* **32(1),** 3–25 (1980).

30. Hayward, D.A. & Ristic, J. Measuring attention using the Posner cuing paradigm: the role of across and within trial target probabilities. *Front. Hum. Neurosci.* **7,** 1–11 (2013).

31. Kok, P., Rahnev, D., Jehee, J.F.M., Lau, H.C. & de Lange, F.P. Attention reverses the effect of prediction in silencing sensory signals. *Cereb. Cortex* **22,** 2197–2206 (2012).

32. Jiang, J., Summerfield, C. & Egner, T. Attention sharpens the distinction between expected and unexpected percepts in the visual brain. *J. Neurosci.* **33(47),** 18438–18447 (2013).

33. Beck, M.R., Hong, S.L., van Lamsweerde, A.E. & Ericson, J.M. The effects of incidentally learned temporal and spatial predictability on response times and visual fixations during target detection and discrimination. *PLoS One* **9(4),** e94539 (2014).

34. Gordon, I.E. Stimulus probability and simple reaction time. *Nature* **215,** 895–896 (1967).

35. Summerfield, C. & de Lange, F.P. Expectation in perceptual decision making: neural and computational mechanisms. *Nat. Rev. Neurosci.* **15,** 745–756 (2014).

36. Wiggs, C.L., & Martin, A. Properties and mechanisms of perceptual priming. *Curr. Opin. Neurobiol.* **8**, 227–233 (1998).

37. Mars, R.B., Bestmann, S., Rothwell, J.C. & Haggard, P. Effects of motor preparation and spatial attention on corticospinal excitability in a delayed-response paradigm. *Exp. Brain Res*. **182,** 125–129 (2007).

38. Bestmann, S. et al. Influence of uncertainty and surprise on human corticospinal excitability during preparation for action. *Curr. Biol.* **18,** 775–780 (2008).

39. Heinze, H.J. et al. Combined spatial and temporal imaging of brain activity during visual selective attention in humans. *Nature* **372,** 543–546 (1994).

40. Yi, D.J., Kelley, T.A., Marois, R., & Chun, M.M. Attentional modulation of repetition attenuation is anatomically dissociable for scenes and faces. *Brain Res.* **1080,** 53–62 (2006).

41. Gold, J.I., & Shadlen, M.N. Neural computations that underlie decisions about sensory stimuli. *Trends Cogn. Sci.* **5,** 10–16 (2001).

42. Gold, J.I., & Shadlen, M.N. The neural basis of decision making. *Annu. Rev. Neurosci.* **30,** 535–574 (2007).

43. Macaluso, E., & Driver, J. Multisensory spatial interactions: a window onto functional integration in the human brain. *Trends Neurosci.* **28(5),** 264–271 (2005).

44. Koelewijn, T., Bronkhorst, A., & Theeuwes, J. Attention and the multiple stages of multisensory integration: a review of audiovisual studies. *Acta Psychol*. **134(3),** 372–384 (2010).

45. Talsma, D. Predictive coding and multisensory integration: an attentional account of the multisensory mind. *Front. Integr. Neurosci*. **9,** 1–13 (2015).

46. Brainard, D.H. The psychophysics toolbox. *Spat. Vis*. **10,** 433–436 (1997).

47. Kleiner, M. et al. What's new in Psychtoolbox-3? *Perception*, **36,** 1–16 (2007).

48. Gardner, W.G. & Martin, K.D. HRTF measurements of a KEMAR. *J. Acoust. Soc. Am*. **97(6),** 3907–3908 (1995).

49. Blignaut, P. Fixation identification: the optimum threshold for a dispersion algorithm. *Att. Perc. Psychophys.* **71(4),** 881–895 (2009).

50. West, B.T., Welch, K.B., & Galecki A.T. *Linear-mixed models - a practical guide using statistical software, Second Edition* (CRC Press, 2015).

Chapter 4:

Modality-specific and crossmodal effects of response probability

Arianna Zuanazzi[1], Uta Noppeney[1]

[1]Computational Cognitive Neuroimaging Laboratory, Computational Neuroscience and Cognitive Robotics Centre, University of Birmingham, B15 2TT Birmingham, UK

Author contributions:

Conceived and designed the experiments: **Arianna Zuanazzi**, Uta Noppeney.

Performed the experiments: **Arianna Zuanazzi**.

Analysed the data: **Arianna Zuanazzi**.

Wrote the manuscript: **Arianna Zuanazzi**.

ABSTRACT

In our multisensory environment, our brain needs to combine signals in multiple sensory modalities into a coherent percept. Spatial attention is a top-down mechanism which contributes to perceptual decisions by prioritizing processing of sensory signals that are task-relevant. Previous studies showed that signals in different sensory modalities which are in spatial proximity share attentional resources. However, by manipulating attention via signal probability over space, these studies confounded attention and expectation.

In two experiments within-subjects we orthogonally manipulated spatial attention (i.e. task-relevance) and expectation (i.e. stimulus probability) selectively in audition (experiment 1) or vision (experiment 2) and evaluated their modality-specific effects (i.e., in audition in experiment 1) and their crossmodal effects (i.e., in vision in experiment 1). Synergistic behavioural effects of attention and expectation (i.e., response probability effects) were found crossmodally in both audition and vision as secondary modalities. Further, we showed that modality-specific effects of response probability in audition emerged later in time than in vision. We discussed this result in terms of sensory reliability of the target stimuli.

Keywords:

modality-specific · crossmodal · audition · vision · spatial attention · spatial expectation · response probability

4.1 INTRODUCTION

Our brain is constantly exposed to auditory and visual signals. For efficient interactions with the environment, it therefore needs to create a coherent representation of such multisensory information. Critically, attention is a top-down mechanism which contributes to selection of information that is task relevant, by increasing the precision of perceptual inference (Rao, 2005; Carrasco, 2011). Consequently, attention facilitates perception (e.g., faster reaction times, improved accuracy, van Ede et al., 2012) of: i. signals in the attended sensory modality (i.e., modality specific attention, Odegaard et al., 2016), ii. attended objects or features (i.e., object and feature-based attention, Saenz et al., 2002; Soto & Blanco, 2004), iii. signals presented at the attended location (i.e., spatial attention, Macaluso & Doricchi, 2013). The question arises as to whether different attentional resources are available for different sensory modalities (i.e., modality-specific effects of attention). Previous studies showed that, when space becomes relevant and signals are in spatial proximity, attentional resources are consistently shared across sensory modalities (Eimer & Driver, 2001; Wahn & König 2015, 2017). For instance, Spence & Driver (1996) manipulated spatial attention (i.e. signals were presented in the attended or unattended hemifield) in one modality only (e.g., audition) and showed behavioural facilitation for attended signals not only for the modality where attention was directly manipulated (i.e., primary modality) but also for the secondary modality where attention was not directly manipulated (e.g., vision). Their results demonstrate that the auditory and visual spatial attention systems are linked. This conclusion is further confirmed by neuroimaging studies showing increased activations in frontoparietal and sensory cortices for

attention orienting irrespective of stimulus modality (Macaluso et al. 2000, 2002a, 2002b; Macaluso & Driver 2001, 2005; Calvert, 2001).

Importantly, a fully supramodal attention system would entail qualitatively and quantitatively identical effects across different sensory modalities. Conversely, previous studies revealed that the magnitude of spatial attentional effects is larger in the primary than in the secondary modality (Spence & Driver, 1996; Mondor & Amirault, 1998). These results are at odds with the hypothesis that the attention system may be fully supramodal, and suggest that attention effects spread across sensory modalities in an attenuated fashion (Driver & Spence, 1998).

Critically, if spatial attention is not fully supramodal and modality-specific effects can be observed in perception and decision-making, attention may differently interact with the sensory-specific properties of the signal and asymmetric crossmodal behavioural effects may be observed, e.g., generalization of attention effects from vision to audition but not vice versa (Ward at al., 2000). In particular, audition and vision differ in terms of spatial resolution. The tonotopic auditory system is characterized by a poorer spatial resolution than the retinotopic visual system (Mondor & Zatorre, 1995), thus modality-specific and crossmodal effects may potentially decrease when spatial attention is manipulated in audition. Consequently, asymmetric effects may not depend on attentional processing per se but rather on the interaction between attention and different sensory systems (for similar conclusions, Spence & Driver, 1997).

One further aspect deserves consideration. Previous studies investigating crossmodal properties of attention often conflated attention with expectation by manipulating spatial attention via signal probability (i.e., by employing the Posner task, Posner, 1980) (Spence and Driver, 1996, 1997). Similarly to attention,

expectation (or signal probability) facilitates perception (Rohenkohl et al., 2014) by encoding the statistical structure of the environment (Summerfield & Egner, 2009). When these two processes are operationally confounded, modality-specific and crossmodal behavioural effects of spatial attention cannot be separated from that of spatial expectation and conclusions about each single process cannot be drawn.

In this study we orthogonally manipulated spatial attention (i.e. task-relevance) and expectation (i.e. signal probability) across the two hemifields and we investigated modality-specific and crossmodal behavioural effects of such manipulation. In particular, in two (within participants) experiments we (1) manipulated spatial attention and expectation in the auditory modality and evaluated their effects on detection of auditory and visual signals (experiment 1) and (2) manipulated spatial attention and expectation in the visual modality and evaluated their effects on detection of visual and auditory signals (experiment 2). This approach allowed us to separately analyse modality-specific and crossmodal effects of spatial attention and expectation. By comparing behavioural effects of our two experiments, we observed qualitatively similar synergistic effects of attention and expectation in audition and vision. However, we showed that such effects differed quantitatively across sensory modalities. We demonstrated that this difference can be accounted for in terms of sensory-specific probability learning over time.

4.2 MATERIALS AND METHODS

4.2.1 Participants

Twenty-eight healthy subjects (19 females; mean age = 25.57 years; range 18–60 years, 24 right-handed) participated in the study (experiment 1 and experiment 2, within participants). A power analysis (G*, Faul et al., 2009) showed that at least 14

and 26 subjects were needed to detect effect sizes (i.e., Cohen's d_{av}) of 0.18 and 0.20 respectively, with a power (1-β) of 0.8 and α = 0.05. These effect sizes were estimated for the significant simple main effects in a previous study that we conducted with the same experimental design (Zuanazzi & Noppeney, 2018). Thus, data collection was terminated as soon as 28 participants had undergone the study. All participants had normal or corrected to normal vision and reported normal hearing. All participants provided written informed consent and were naïve to the aim of the study. The study was approved by the local ethics committee of the University of Birmingham (Science, Technology, Mathematics and Engineering (STEM) Ethical Review Committee) and the experiment was conducted in accordance with these guidelines and regulations.

4.2.2 Stimuli and Apparatus

Spatial auditory stimuli of 100 ms duration were created by convolving five different bursts of white noise (with 5 ms onset and offset ramps) with spatially selective head-related transfer functions (HRTFs) based on the KEMAR dummy head of the MIT Media Lab (http://sound.media.mit.edu/resources/KEMAR.html). Visual stimuli ('flashes') were white discs (radius: 0.88º visual angle, luminance: 196 cd/m2) of 100 ms duration presented on a grey background. Both auditory and visual stimuli were presented at ±10º of horizontal visual angle along the azimuth (0º of vertical visual angle). Throughout the entire experiment, a fixation cross was presented in the centre of the screen.

Prior to the beginning of the study, participants were tested for their ability to discriminate left and right auditory stimuli on a brief series of 20 trials. They indicated their spatial discrimination response (i.e., 'left' vs 'right') via a two-choice key press (group mean accuracy was 99% ± 0.4% [across subjects mean ± SEM]).

During the experiment, participants rested their chin on a chinrest with the height held constant across all the participants. Auditory stimuli were presented at approximately 72 dB SPL, via HD 280 PRO headphones (Sennheiser, Germany). Visual stimuli were displayed on a gamma-corrected LCD monitor (2560 x 1600 resolution, 60 Hz refresh rate, 30" Dell UltraSharp U3014, USA), at a viewing distance of approximately 50 cm from the participant's eyes. Stimuli were presented using Psychtoolbox version 3 (Brainard, 1997; www.psychtoolbox.org), running under Matlab R2014a (Mathworks Inc., Natick, MA, USA) on a Windows machine. Participants responded to all stimuli with the same index finger of their dominant hand. Responses were recorded via one key of a small keypad (Targus, USA). Throughout the study, participants' eye-movements and fixation were monitored using Tobii Eyex eyetracking system (Tobii, Sweden).

4.2.3 Design and Procedure

Experiment 1 investigated the effect of auditory spatial attention and expectation on detection of auditory (i.e., primary modality) and visual (i.e., secondary modality) targets using a 2 (*auditory spatial attention*: left vs right) x 2 (*auditory spatial expectation*: left vs right) x 2 (*stimulus modality*: auditory vs visual) x 2 (*stimulus location*: left hemifield vs right hemifield) factorial design; experiment 2 investigated the effect of visual spatial attention and expectation on detection of visual (i.e., primary modality) and auditory (i.e., secondary modality) targets using a 2 (*visual spatial attention*: left vs right) x 2 (*visual spatial expectation*: left vs right) x 2 (*stimulus modality*: visual vs auditory) x 2 (*stimulus location*: left hemifield vs right hemifield) factorial design.

Spatial attention was manipulated for the primary modality as task-relevance, i.e. the requirement to respond to an auditory (experiment 1) or a visual (experiment

2) target in the left vs right hemifield over runs of 80 trials. Prior to each run a cue (duration: 2000 ms) informed the observer whether to respond to targets in either their left or right hemifield. Spatial expectation was manipulated as spatial signal probability for signals in the primary modality across experimental sessions that were performed on different days. Auditory (experiment 1) or visual (experiment 2) signals were presented with a ratio of 2.33/1 (i.e., 70%, 30%) in the expected/unexpected hemifield. Observers were not informed about those probabilities, but learnt them implicitly. Conversely, spatial attention and expectation were not directly manipulated in the secondary modality. Thus, participants needed to respond to all visual targets (experiment 1) or auditory targets (experiment 2) presented with equal probabilities in their spatial hemifields (i.e., ratio 1/1 in the expected/unexpected hemifields) (Fig. 4.1A and 4.1B). Overall participants' accuracy in target detection (i.e., percentage of hits + correct rejections) was: experiment 1: 98.3% ± 0.2% [across subjects mean ± SEM]; experiment 2: 98.6% ± 0.2% [across subjects mean ± SEM].

Each trial (SOA: 2300 ms) included three time windows: i. fixation cross alone (700 ms duration), ii. brief sound or flash (stimulus duration: 100 ms) and iii. fixation cross alone (1500 ms as response interval). Participants responded only to the targets in the primary modality presented in the attended hemifield and to targets in the secondary modality in either hemifield via key press with the same index finger (i.e., the same response for all auditory and visual stimuli) as fast and accurately as possible (Fig. 4.1C). Their fixation performance was monitored via eye tracking.

Each experiment included two sessions (i.e. spatial expectation left vs right on different days), for an overall of 4 sessions. Each session included 12 attention runs with 80 trials (duration: 3 mins/run). Runs were of two types: in run type A (Fig. 4.1A and 4.1B) spatial attention and expectation were congruent (i.e. spatial attention was

directed to the hemifield with higher stimulus frequency); in run type B spatial attention and expectation were incongruent (i.e. spatial attention was directed to the hemifield with less frequent stimuli).

Overall, each experiment included 80 trials x 12 attention runs (6 runs of type A and 6 runs of type B) x 2 expectation sessions = 1920 trials in total (and 3840 for the whole study). Specifically, each run type included, for the primary modality, 336 targets presented in the expected hemifield (pooled over left and right) and 144 targets in the unexpected hemifield (pooled over left and right). Each run type also included, for the secondary modality, 240 targets presented in the expected hemifield and 240 targets in the unexpected hemifield (pooled over left and right). For further details see Fig. 4.1B which shows the absolute number of trials for each condition and run type.

The order of experiment 1 and 2 and of expectation (i.e. left vs right) sessions was counterbalanced across participants; the order of attention runs was counterbalanced within and across participants and the order of stimulus location and stimulus modality was pseudo-randomized within each participant. Brief breaks were included after every two attention runs to provide feedback to participants about their performance accuracy (averaged across all conditions) in the target detection task and about their eye-movements (i.e. fixation maintenance). Prior to each session, participants were familiarized with the stimuli in brief practice blocks (with equal spatial signal probability) and trained on target detection performance and fixation (i.e. a warning signal was shown when the disparity between the central fixation cross and the eye-data samples exceeded 2.5 degrees). After the final session (for half of the participants it corresponded to the second session of experiment 1 and for the other half to the second session of experiment 2) participants indicated in a questionnaire whether they thought the stimuli in the primary modality were presented more

frequently in one of the two spatial hemifields. 10 out of 14 participants in experiment 1 and 11 out of 14 participants in experiment 2 correctly identified the expectation manipulation in the primary modality. Moreover, 13 out of 14 participants in experiment 1 and 13 out of 14 in experiment 2 correctly reported that stimuli in the secondary modality were presented with equal probabilities across the two hemifields. These data suggest that most participants were aware of th manipulation of signal probability in both experiments.

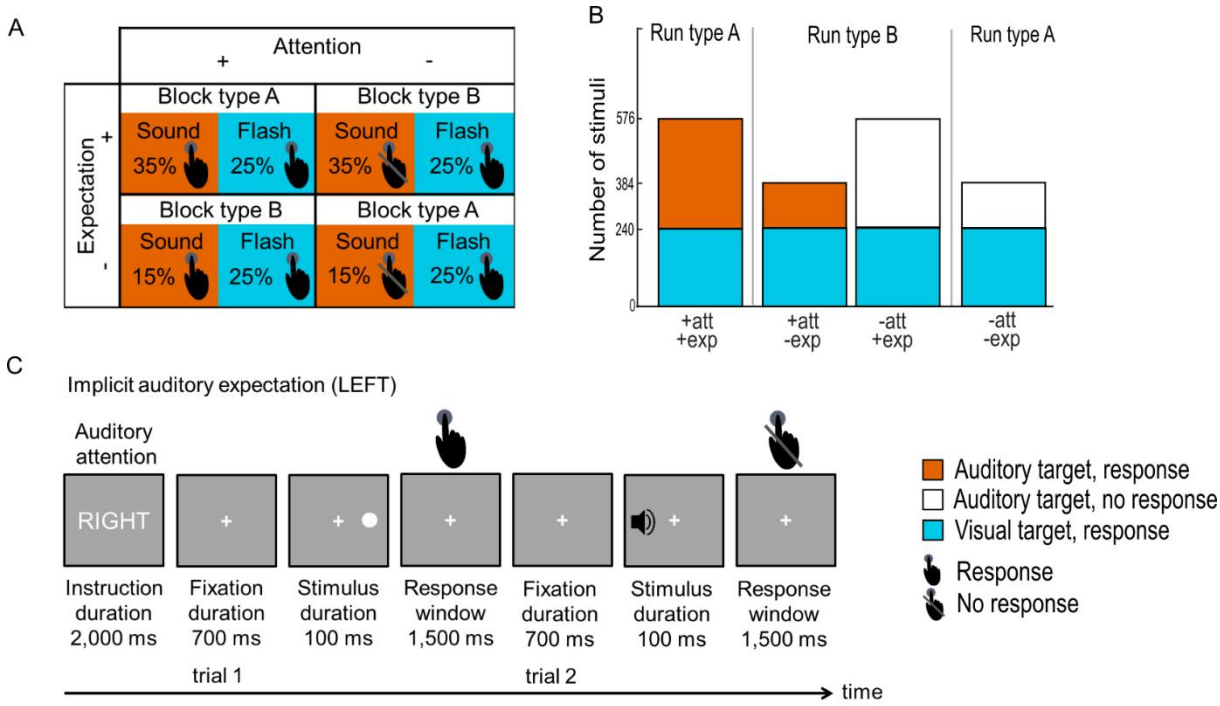


Figure 4.1: Experiment 1: Design and example trials.

A. Auditory spatial attention and expectation (i.e. signal probability) were manipulated in a 2 (auditory modality – orange, vs visual modality – blue) x 2 (attended hemifield vs unattended hemifield) x 2 (expected hemifield vs unexpected hemifield) factorial design. For illustration purposes, stimulus locations (left/right) were collapsed. Presence vs absence of response requirement is indicated by the hand symbol, probability manipulation is indicated by the %. **B.** Number of auditory (orange) and visual (blue) trials in the 2 (attended vs unattended) x 2 (expected vs unexpected) design. The bar plots are filled (i.e. a response is required) or not filled (i.e. no response is required). The fraction of filled area for each bar represents the probability that the observer needs to respond conditioned on that the signal is presented in a particular hemifield (i.e., 'spatially selective response probability'). The fraction of filled area pooled over the two bars of one particular run type (e.g. run type A) represents the probability that the observer needs to respond irrespective of the hemifield in which the signal is presented (i.e., 'general response probability'); general response probability is greater for run type A, where attention and expectation are congruent. **C.** Example of two consecutive trials of Experiment 1: at the beginning of each block, a cue informed participants whether to attend and respond to auditory signals selectively in their left or right hemifield throughout the entire block. On each trial

participants were presented with an auditory or visual stimulus (100 ms duration) either in their left or right hemifield. They were instructed to respond to auditory stimuli only in the 'attended' hemifield and to all visual stimuli irrespective of hemifield as fast and accurately as possible with the same finger. The response window was limited to 1500 ms. Participants were not explicitly informed that auditory signals were more likely to appear in one of the two hemifields. Instead, spatial expectation was implicitly learnt over runs. In experiment 1 audition was the primary modality and vision was the secondary modality.

Note. Design and procedure of experiment 2 were identical to that of experiment 1, except that attention and expectation were manipulated in vision, thus vision is the primary modality and audition was the secondary modality.

4.2.4 Data analysis

4.2.4.1 Eye movement: exclusion criteria

We excluded trials where participants did not successfully fixate the central cross based on a dispersion criterion (i.e. distance of fixation from subject's centre of fixation, as defined in calibration trials, > 1.3 degrees for three subsequent samples, Blignaut, 2009). Our eye tracking data confirmed that participants successfully maintained fixation in both experiments with only a small number of trials to be excluded (experiment 1: excluded auditory response trials 1.8% ± 0.5% [across subjects mean ± SEM]; excluded visual response trials 1.7% ± 0.5% [across subjects mean ± SEM]; experiment 2: excluded auditory response trials 2.7% ± 0.9% [across subjects mean ± SEM]; excluded visual response trials 2.7% ± 1% [across subjects mean ± SEM]).

4.2.4.2 Response time analysis

Response time (RT) analysis was limited to trials where RT did not exceed 1500 ms. To evaluate the effects of spatial attention and expectation and their crossmodal generalization and to qualitatively compare the two experiments, separately for each experiment median RT for attended auditory (experiment 1) or visual (experiment 2) targets for each subject were entered into a two-sided paired-sample t-test with *spatial expectation* (expected vs unexpected stimulus) as factor. Moreover, separately for each experiment median RT for visual (experiment 1) or auditory (experiment 2) targets for each subject were entered into a 2 (*spatial attention*: attended vs unattended stimulus) x 2 (*spatial expectation*: expected vs unexpected stimulus) repeated measures analysis of variance (rmANOVA). Unless otherwise indicated, we only report effects that are significant at $p < 0.05$.

To quantitatively evaluate the magnitude of modality-specific vs crossmodal effects of expectation in function of sensory modality (n.b. given our design where participants responded only to attended targets in the primary modality, the effect of modality-specific attention could not be quantitatively evaluated for the primary modality), we combined the two experiments: for each subject the difference of median RT (ΔRT) between unexpected and expected targets presented in the attended hemifield were entered into a 2 (*effect type*: modality-specific, crossmodal) x 2 (*stimulus modality*: audition, vision) rmANOVA.

Finally, modality-specific effects in the primary modality were further analysed from a temporal perspective. For each subject and for each primary modality, the difference of ΔRT between unexpected and expected targets presented in the attended hemifield was calculated over six chunks of 160 trials and entered in a rmANOVA with *stimulus modality* (audition, vision) and *temporal chunks* (1 to 6 in temporal order) as factors. Temporal chunks included stimuli presented between breaks, pooled over sessions 1 and 2.

4.3 RESULTS

4.3.1 Modality-specific and crossmodal effects of attention and expectation

In experiment 1, participants responded to auditory targets presented in their attended hemifield and to all visual targets. In experiment 2, participants responded to visual targets presented in their attended hemifield and to all auditory targets. We observed qualitatively similar modality-specific and crossmodal effects in experiment 1 and 2. Table 4.1 displays RT (across participants' mean ± SEM) for the auditory and visual modality for the two experiments. For the primary modality, the two-sided paired-sample t-tests showed significantly faster RT in the attended hemifield, when

this hemifield was expected than unexpected (experiment 1, auditory modality: $t(27) = -2.83$, $p = 0.009$, Cohen's d_{av} [95% CI] = -0.16 [-0.27 , -0.04]; experiment 2, visual modality: $t(27) = -9.62$, $p < 0.001$, Cohen's d_{av} [95% CI] = -0.35 [-0.47 , -0.23]) (Table 4.1, Fig. 4.2A and 4.2B). For the secondary modality, the 2 (attended vs unattended) x 2 (expected vs unexpected) rmANOVAs revealed a significant main effect of *attention* (experiment 1, visual modality: $F(1, 27) = 72.08$, $p < 0.001$, η_p^2 [90% CI] = 0.73 [0.55, 0.80]; experiment 2, auditory modality: $F(1, 27) = 36.91$, $p < 0.001$, η_p^2 [90% CI] = 0.58 [0.34, 0.70]). Participants responded faster to targets in their attended than unattended hemifields, suggesting that the effect of attention generalized crossmodally, as in the secondary modality attention was not directly manipulated. Moreover, a significant crossover interaction between *attention* and *expectation* was observed (experiment 1, visual modality: $F(1, 27) = 10.09$, $p = 0.004$, η_p^2 [90% CI] = 0.27 [0.06, 0.46]; experiment 2, auditory modality: $F(1, 27) = 44.10$, $p < 0.001$, η_p^2 [90% CI] = 0.62 [0.40, 0.73]). The simple main effects showed that participants responded significantly faster to targets in the attended hemifield when this hemifield was expected than unexpected (experiment 1, visual modality: $t(27) = -2.81$, $p = 0.009$, Cohen's d_{av} [95% CI] = -0.08 [-0.15 , -0.02]; experiment 2, auditory modality: $t(27) = -5.96$, $p < 0.001$, Cohen's d_{av} [95% CI] = -0.14 [-0.19 , -0.08]). These results correspond to the response profile we observed for targets in the primary modalities, suggesting that also the effect of expectation generalized crossmodally. By contrast, participants responded significantly more slowly to targets in the unattended hemifield, when this hemifield was expected than unexpected (experiment 1, visual modality: $t(27) = 2.56$, $p = 0.017$, Cohen's d_{av} [95% CI] = 0.09 [0.02, 0.17]; experiment 2, auditory modality: $t(27) = 5.53$, $p < 0.001$, Cohen's d_{av} [95% CI] = 0.18 [0.10 , 0.26]) (Table 4.1, Fig. 4.2A and 4.2B).

Table 4.1: Group mean reaction times (RT) for each stimulus modality in each condition requiring a response for experiment 1 (primary modality: audition) and experiment 2 (primary modality: vision).
Standard errors (SEM) are given in parentheses.

	Auditory modality		Visual modality			
Experiment 1	+att +exp	+att -exp	+att +exp	+att -exp	-att +exp	-att -exp
RT (ms) (SEM)	599.7 (20.1)	616.2 (19.2)	514.4 (16.9)	522.1 (16.9)	583.8 (18.5)	574.4 (18.9)
	Visual modality		Auditory modality			
Experiment 2	+att +exp	+att -exp	+att +exp	+att -exp	-att +exp	-att -exp
RT (ms) (SEM)	526.3 (17.9)	561.8 (19.9)	527.4 (20.2)	542.1 (20)	605.2 (25.7)	581.3 (24.2)

4.3.2 Magnitude of modality-specific and crossmodal effects of expectation (for attended targets)

In the previous analysis we qualitatively showed that behavioural effects of attention and expectation generalised crossmodally. Here, we further explored these results from a quantitative perspective by directly comparing the magnitude of the crossmodal effects in the two experiments. The effects of attention cannot be compared across primary and secondary modality as in the primary modality unattended targets were not responded to. Thus, we limited this analysis to the effect of expectation for the attended targets. The 2 (*effect type*) x 2 (*stimulus modality*) rmANOVA revealed a significant main effect of *effect type* ($F(1, 27) = 18.03$, $p < .001$, η_p^2 [90% CI] =0 .40

[0.16, 0.56]), showing larger modality-specific effects than crossmodal effects, and a significant crossover interaction between *effect type* and *stimulus modality* ($F(1, 27) = 20.71$, $p < .001$, η_p^2 [90% CI] = 0 .43, [0.19, 0.59]). Post-hoc comparisons showed that (1) modality-specific effects were significantly larger than crossmodal effect when the primary modality was vision ($t(27) = 5.39$, $p < 0.001$, Cohen's d_{av} [95% CI] = 1.25 [0.68 , 1.81]) but only a nonsignificant trend could be observed for audition as primary modality (Fig. 4.2C); (2) modality-specific effects were significantly larger for vision than audition ($t(27) = 3.55$, $p = 0.001$, Cohen's d_{av} [95% CI] = 0.73 [0.28, 1.18]), but no significant difference was found for crossmodal effects between the two modalities (Fig. 4.2C); (3) while for vision modality-specific effects were significantly larger than crossmodal effects ($t(27) = 7.14$, $p < 0.001$, Cohen's d_{av} [95% CI] = 1.62, [0.99, 2.23]), no difference was found for audition (Fig. 4.2C). These results suggest that modality specific effects of expectation were overall larger than crossmodal effects, and this difference was significantly larger when expectation was manipulated in vision as primary modality.

4.3.3 Temporal aspects of modality-specific effects of expectation (for attended targets)

In the previous analysis we showed that modality-specific and crossmodal effects of expectation differed in relation to sensory modality. Here, we investigated whether this difference can be explained in terms of how auditory and visual expectations develop over time. The rmANOVA revealed a significant main effect of *modality* ($F(1, 27) = 13.28$, $p = 0.001$, η_p^2 [90% CI] =0.33 [0.1, 0.51]) and a significant *modality* by *temporal chunks* interaction ($F(5, 135) = 2.57$, $p = 0.03$, η_p^2 [90% CI] =0.09 [0, 0.14]). First, we unpacked the interaction by analysing the temporal chunks effect in each primary modality. For auditory targets in experiment 1, the rmANOVA

with *temporal chunks* as factor revealed a significant main effect of *temporal chunks* ($F(5, 135) = 2.83$, $p = 0.018$, η_p^2 [90% CI] =0.09 [0, 0.15]). Post-hoc pairwise comparisons (Bonferroni-corrected for 15 multiple comparisons) revealed that the ΔRT between unexpected and expected auditory targets was significantly smaller at the beginning of the session than at the end (chunk 1 [mean ΔRT: 1 ms, SEM ΔRT: 9 ms], vs chunk 5 [mean ΔRT: 38 ms, SEM ΔRT: 8 ms]: $p = 0.019$, chunk 2 [mean ΔRT: 3 ms, SEM ΔRT: 8 ms] vs chunk 5 [mean ΔRT: 38 ms, SEM ΔRT: 8 ms]: $p = 0.041$) suggesting that expectation effects for auditory targets emerged only towards the end of the session (Fig. 4.2D, orange line). For visual targets in experiment 2, we did not find a significant main effect of *temporal chunks*. This result suggests that expectation effects for visual targets emerged right from the beginning of the session (further confirmed by ΔRT in chunk 1 [mean ΔRT: 31 ms, SEM ΔRT: 8 ms] significantly larger than 0: $t_1(27) = 3.97$ $p_1 < 0.001$ Cohen's d_1 [95% CI] = 0.75 [0.32, 1.16]) and remained constant throughout the session (further confirmed by ΔRT in chunk 2, 3, 4, 5 and 6 [mean ΔRT: 26 ms, SEM ΔRT: 6 ms] significantly larger than 0: $t_2(27) = 4.77$ $p_2 < 0.001$ Cohen's d_2 [95% CI] = 0.90 [0.45, 1.34], $t_3(27) = 5.01$ $p_3 < 0.001$ Cohen's d_3 [95% CI] = 0.95 [0.50, 1.39], $t_4(27) = 6.28$ $p_4 < 0.001$ Cohen's d_4 [95% CI] = 1.19 [0.70, 1.67], $t_5(27) = 4.94$ $p_5 < 0.001$ Cohen's d_5 [95% CI] = 0.93 [0.48, 1.37], $t_6(27) = 4.67$ $p_6 < 0.001$ Cohen's d_6 [95% CI] = 0.88 [0.44, 1.31]). To further investigate the sustained effects of visual expectation over time (i.e., no difference between e.g. chunk 1 and 6), we ran a Bayesian paired-sample *t*-test (JASP Version 0.8.4.0, using the default Cauchy prior) between chunk 1 and 6. These data were 4.33 more likely to occur under the null hypothesis (i.e., ΔRT in chunk 1 and 6 are the same) (Bayes Factor in favour of the null hypothesis, $BF_{01} = 4.33$, i.e., moderate evidence, Jeffreys, 1961)

relative to the alternative hypothesis (i.e., ΔRT in chunk 1 and 6 are different) (Fig. 4.2D, blue line).

Second, to verify whether modality-specific effects of auditory expectation eventually became similar to that of visual expectation, we further unpacked the *modality* by *temporal chunk* significant interaction by comparing ΔRT of auditory and visual targets for each chunk. Results showed that ΔRT for auditory and visual targets were quantitatively different at the beginning of the session (auditory chunk 2 [mean ΔRT: 3 ms, SEM ΔRT: 8 ms] vs visual chunk 2 [mean ΔRT: 35 ms, SEM ΔRT: 7 ms], $t(27) = -3.41$, $p = 0.012$ Bonferroni-corrected for 6 multiple comparisons), but became comparable at the end of the session (auditory chunk 6 [mean ΔRT: 22 ms, SEM ΔRT: 10 ms] vs visual chunk 6 [mean ΔRT: 26 ms, SEM ΔRT: 6 ms], $p > 0.05$ Bonferroni-corrected for 6 multiple comparisons). For completeness, comparisons for auditory and visual chunks 1, 3, 4 and 5 showed no significant results (Fig. 4.2D, orange vs blue line). This conclusion is further supported by the Bayesian paired-sample t-test between auditory and visual ΔRT for chunk 6. In fact, these data were 4.68 more likely ($BF_{01} = 4.68$, i.e., moderate evidence) to occur under the null hypothesis, i.e., at the end of the session expectation effects for audition and vision were comparable (Fig. 4.2D).

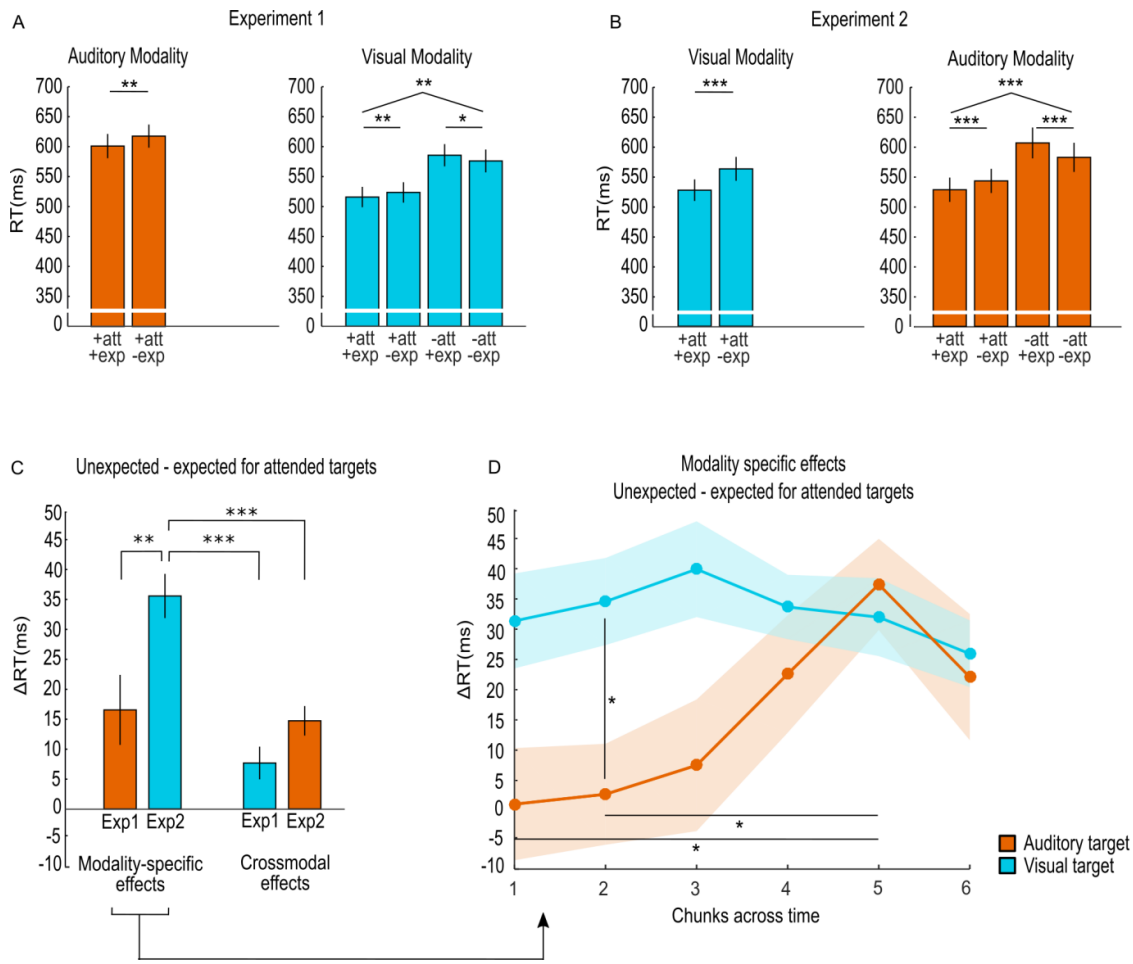


Figure 4.2: Behavioural results.

Bar plots represent across subjects' mean (±SEM) RT for each of the six conditions with response requirements for experiment 1 (**A**) and 2 (**B**). **C.** Bar plots represents across subjects' mean (±SEM) ΔRT for modality-specific and crossmodal effects of spatial expectation (i.e., unexpected – expected) for attended targets in experiment 1 and 2. **D.** Modality specific effects of expectation (across subjects' mean ΔRT) over time (i.e., across chunks) for audition and vision. Shading represents standard error of the mean.

Brackets and stars indicate significance of main effects and interactions. *$p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. Audition: orange; vision: blue.

4.4 DISCUSSION

Previous behavioural and neuroimaging research demonstrated that attentional resources are, to a certain degree, shared across sensory modalities (Spence & Driver, 1996; Macaluso et al., 2002a, 2002b). Nevertheless, these studies operationally conflated spatial attention and expectation making it difficult to disentangle their respective crossmodal effects. The novelty of our study arises from our multisensory paradigm which allowed us to investigate modality-specific and crossmodal effects of spatial attention and expectation when operationally defined as task-relevance and signal probability. In two experiments (within subjects) we manipulated spatial attention and expectation in the primary modality (audition in experiment 1 and vision in experiment 2) and evaluated modality-specific effects and crossmodal effects on detection of auditory and visual targets (i.e., secondary modalities).

In both experiments, when stimuli were presented in the attended hemifield, participants responded faster than when presented in the unattended hemifield. We observed this effect in the secondary modality, where spatial attention was not directly manipulated. This result suggests that attention effects generalized across modalities and is consistent with previously reported facilitation induced by spatial attention regardless of sensory modality (Spence & Driver, 1996; Driver & Spence, 1998). Importantly, in both experiments we observed synergistic interactions between spatial attention and expectation in the secondary modality, which suggests that a generalization of expectation across modalities took place. In particular, for attended stimuli we observed faster RT when signals were expected than unexpected. This pattern reversed when stimuli where unattended. In a previous study (Zuanazzi & Noppeney, 2018) we argued that this RT pattern is explained by attention and

expectation interactively contributing to perceptual decisions by co-determining: general response probability, i.e. the probability to respond regardless of stimulus location, and spatially selective response probability, i.e. the probability to respond conditioned on the stimulus location. Thus, stimulus probability per se did not affect RT. More specifically, in blocks where attention and expectation are directed to the same hemifield (blocks of type A, Fig. 4.1A and 4.1B), participants have to respond to 85% of trials of the entire block, versus 65% in blocks of type B (i.e., general response probability, Fig. 4.1A and 4.1B). Hence, response facilitation may be interpreted as an increase in alertness, arousal or motor preparation (Mars et al., 2007; Bestmann et al., 2008) (Fig. 4.2A and 4.2B). Moreover, when stimuli are presented in the unattended and expected hemifield, participants are slower as they have to respond to only a small proportion of trials presented at that location (i.e. 240 out of 576: 41.7%), versus a larger proportion when presented in the unattended and unexpected hemifield (i.e. 240 out of 384: 62.5%) (i.e., spatially selective response probability, Fig. 4.1A and 4.1B, Fig. 4.2A and 4.2B). Therefore, we concluded that the operational dichotomy where spatial attention and expectation are fully dissociated is not appropriate when investigating perceptual decisions, as we demonstrated that these two mechanisms in fact jointly co-determine response probability. Thus, following this line of reasoning, in the current study we interpreted modality-specific and crossmodal effects co-determined by attention and expectation as effects of response probability.

In Zuanazzi & Noppeney (2018) we investigated crossmodal effects of response probability only for the visual modality. Importantly, our two current experiments not only largely replicate our previous results but enrich our conclusions by showing that response probability had qualitatively equivalent crossmodal effects in vision and audition.

Next, going beyond previous research showing that spatial attention is not fully supramodal due to the fact that behavioural effects in the secondary modality are attenuated (Spence & Driver, 1998) and that there exist asymmetries between crossmodal effects depending on the sensory modality employed as primary modality (Ward at al., 2000. We investigated whether response probability exhibited fully supramodal effects (i.e., same effects irrespective of stimulus modality). Hence, we examined whether crossmodal effects of response probability were quantitatively equivalent to modality-specific effects. Critically, our 2 (*effect type*) x 2 (*stimulus modality*) rmANOVA allowed us to compare (1) modality-specific vs crossmodal effects within the same experiment, (2) modality-specific and crossmodal effects for audition vs vision (3) modality-specific vs crossmodal effects in audition and vision across experiments. This analysis allowed us to disentangle crossmodal effects of response probability from effects deriving from sensory-specific properties of the signal. Interestingly, our results show that the effects of response probability were larger for the primary modality (in particular when the primary modality was vision) and consequently suggest that response probability is not fully supramodal. Furthermore, the magnitude of modality-specific effects in vision was larger than the same effects in audition. These results suggest that some properties of the sensory signal interacted with response probability. This hypothesis was tested in our temporal analysis. We showed that modality-specific effects of response probability differed across modalities only at the beginning of the session but became equivalent at the end, indicating that modality-specific effects of response probability in audition and vision are not quantitatively different per se, but emerge at different times across the session, with effects in audition arising later in time. We interpreted this result as

evidence for a substantial difference between how spatial response probability is built over time for the two modalities.

This difference may derive from the way vision and audition are represented in the brain (Neumann et al., 1986). Vision is primarily encoded retinotopically (i.e., based on space, Sereno et al., 1995) whereas audition is primarily encoded tonotopically (i.e., based on frequency, Lauter et al., 1985). As a consequence, sounds are localized based on head- and ears-related cues such as interaural differences over time (Middlebrooks & Green, 1991). For this reason, vision is more spatially reliable than audition when visual information is located in the centre of the visual field (Dacey et al., 1992; Knudsen & Brainard, 1995; Stephen et al., 2002). Moreover, auditory and visual events are substantially different in nature. While visual information is continuously available, auditory information is transient, event-dependent and broadly distributed in space (Neisser, 1976; Neumann et al., 1986). Past unisensory research demonstrated that sensory-specific properties of the signal can in fact interact with higher level processes such as attention, especially when task-relevant. For instance, bottom-up capture of attention is determined by the most salient stimulus (Treisman & Gelade, 1980; Theeuwes, 2010) and by stimuli with abrupt onsets (Yantis & Jonides, 1990). Furthermore, statistical information about location, colour or shape of an object can (incidentally) be learnt and used to improve performance (Olson et al., 2005; van Lamsweerde & Beck, 2011).

In the current study, the high reliability of the visual stimulus (i.e., a white disc) may have contributed to shortening the time it took for participants to become aware of the manipulation of spatial expectation in vision. Critically, participants' awareness of the expectation manipulation is evidenced by our questionnaires' results. Alternatively, even in absence of explicit awareness, the probability distribution over

space of more reliable visual signals could have been (implicitly) learnt faster than that of less reliable auditory signals. This could have determined the earlier behavioral effect of response probability that we observed for visual signals. (Miller & Pachella, 1973; Jabar & Anderson, 2015. For a comparison with results of chapter 3 and 5 where the same experimental design was employed, see Supplementary Fig. 8.2: interestingly, when pre-experiment runs were used to boost the implicit learning of auditory spatial signal probability, the effect of response probability emerged right from the beginning of the session).

In conclusion, our findings suggest that synergistic behavioral effects of attention and expectation (i.e., effects of response probability) can be observed regardless of sensory modality. Our design allowed us to explicitly investigate how response probability shapes perceptual decisions in primary and secondary sensory modalities. We showed that effects of response probability are qualitatively similar in the primary and secondary modality but quantitatively attenuated in the secondary modality, suggesting that modality-specific mechanisms are involved. This attenuation is less clear when the primary modality is audition. This result is explained by considering how spatial expectation is learned over time for different sensory modalities. Seemingly smaller modality-specific effects of response probability for auditory signals are more appropriately accounted for in terms of speed of probability learning rather than in terms of absolute magnitude of the effect. We suggested that such effect hinges on the spatial reliability of the sensory signal.

REFERENCES

Barker, L., & Shaw, K. (2015). Best (but of t-forgotten) practices: checking assumptions concerning regression residuals. *Am J Clin Nutr*, *102*(3), 533–539. http://doi.org/10.1016/j.cogdev.2010.08.003.Personal

Bestmann, S., Harrison, L. M., Blankenburg, F., Mars, R. B., Haggard, P., Friston, K. J., & Rothwell, J. C. (2008). Influence of uncertainty and surprise on human corticospinal excitability during preparation for action. *Current Biology*, *18*(10), 775–780. http://doi.org/10.1016/j.cub.2008.04.051

Blignaut, P. (2009). Fixation identification: the optimum threshold for a dispersion algorithm. *Attention, Perception & Psychophysics*, *71*(4), 881–895.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. http://doi.org/10.1163/156856897X00357

Calvert, G. (2001). Crossmodal processing in the human brain: insights from functional neuroimaging studies. *Cerebral Cortex*, *11*, 1110–23. http://doi.org/10.1093/cercor/11.12.1110

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, *51*(13), 1484–1525. http://doi.org/10.1016/j.visres.2011.04.012

Dacey, D. M., & Petersen, M. R. (1992). Dendritic field size and morphology of midget and parasol ganglion cells of the human retina. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(20), 9666–9670. http://doi.org/10.1073/pnas.89.20.9666

Driver, J., & Spence, C. (1998). Attention and the crossmodal construction of space. *Trends in Cognitive Sciences*, *2*(7), 254–262. http://doi.org/10.1016/S1364-6613(98)01188-7

Eimer, M., & Driver, J. (2001). Crossmodal links in endogenous and exogenous spatial attention: evidence from event-related brain potential studies. *Neuroscience and Biobehavioral Reviews*, *25*(6), 497–511. http://doi.org/S0149-7634(01)00029-X [pii]

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–60. http://doi.org/10.3758/BRM.41.4.1149

Gardner, W.G. & Martin, K.D. (1995). HRTF measurements of a KEMAR. *The Journal of the Acoustical Society of America*, *97*(6), 3907–3908.

Jabar, S. B., & Anderson, B. (2015). Probability shapes perceptual precision: A study in orientation estimation. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(6), 1666–1679. http://doi.org/10.1037/xhp0000121

Jeffreys, H. (1961). *Theory of Probability*, 3rd ed. Oxford: Oxford University Press.

Knudsen, E., & Brainard, M. S. (1995). Creating a unified representation of visual and auditory space in the brain. *Annual Review of Neuroscience*, *18*(1), 19–43. http://doi.org/10.1146/annurev.neuro.18.1.19

Lauter, J. L., Herscovitch, P., Formby, C., & Raichle, M. E. (1985). Tonotopic organization in human auditory cortex revealed by positron emission tomography. *Hearing Res*, *20*, 199–205.

Macaluso, E., & Doricchi, F. (2013). Attention and predictions: control of spatial attention beyond the endogenous-exogenous dichotomy. *Frontiers in Human Neuroscience*, *7*(October), 685. http://doi.org/10.3389/fnhum.2013.00685

Macaluso, E., & Driver, J. (2001). Spatial attention and crossmodal interactions between vision and touch. *Neuropsychologia*, *39*(12), 1304–1316. http://doi.org/10.1016/S0028-3932(01)00119-1

Macaluso, E., & Driver, J. (2005). Multisensory spatial interactions: A window onto functional integration in the human brain. *Trends in Neurosciences*, *28*(5), 264–271. http://doi.org/10.1016/j.tins.2005.03.008

Macaluso, E., Frith, C. D., & Driver, J. (2002). Directing Attention to Locations and to Sensory Modalities: Multiple Levels of Selective Processing revealed with PET. *Cerebral Cortex*, *12*(4), 357–368. http://doi.org/10.1093/cercor/12.4.357

Macaluso, E., Frith, C. D., & Driver, J. (2002). Supramodal effects of covert spatial orienting triggered by visual or tactile events. *Journal of Cognitive Neuroscience*, *14*(3), 389–401. http://doi.org/10.1162/089892902317361912

Macaluso, E., Frith, C., & Driver, J. (2000). Selective spatial attention in vision and touch: unimodal and multimodal mechanisms revealed by PET. *Journal of Neurophysiology*, *83*(5), 3062–3075.

Mars, R. B., Bestmann, S., Rothwell, J. C., & Haggard, P. (2007). Effects of motor preparation and spatial attention on corticospinal excitability in a delayed-response paradigm. *Experimental Brain Research*, *182*(1), 125–129. http://doi.org/10.1007/s00221-007-1055-4

Middlebrooks, J. C., & Green, D. M. (1991). Sound localization by human listeners. *Annu. Rev. Psychol.*, (42), 135–159.

Miller, J. O., & Pachella, R. G. (1973). Locus of the stimulus probability effect. *Journal of Experimental Psychology*, *101*(2), 227–231. http://doi.org/10.1037/h0035214

Mondor, T. A., & Amirault, K. J. (1998). Effect of same- and different-modality spatial cues on auditory and visual target identification. *Journal of Experimental Psychology*, *24*(3), 745–755.

Mondor, T. A., & Zatorre, R. J. (1995). Shifting and focussing auditory spatial attention. *J Exp Psychol*, *21*(2), 387–409.

Neisser, U. (1977). Cognition and reality. *The American Journal of Psychology*, *90*(3), 541–543.

Neumann, O., Van der Heijden, A. H., & Allport, D. A. (1986). Visual selective attention: Introductory remarks. *Psychological Research*, *48*(4), 185–188. http://doi.org/10.1007/bf00309082

Odegaard, B., Wozny, D. R., & Shams, L. (2016). The effects of selective and divided attention on sensory precision and integration. *Neuroscience Letters*, *614*, 24–28. http://doi.org/10.1016/j.neulet.2015.12.039

Olson, I. R., Jiang, Y., & Moore, K. S. (2005). Associative learning improves visual working memory performance. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(5), 889–900. http://doi.org/10.1037/0096-1523.31.5.889

Posner, M. I. (1980). Orienting of attention. *Q J Exp Psychol*, *32*(1), 3–25. http://doi.org/10.1080/00335558008248231

Rao, R. P. N. (2005). Bayesian inference and attentional modulation in the visual cortex. *NeuroReport*, *16*(16), 1843–1848. http://doi.org/10.1097/01.wnr.0000183900.92901.fc

Rohenkohl, G., Gould, I. C., Pessoa, J., & Nobre, A. C. (2014). Combining spatial and temporal expectations to improve visual perception. *Journal of Vision*, *14*(4), 1–13. http://doi.org/10.1167/14.4.8

Saenz, M., Buracas, G. T., & Boynton, G. M. (2002). Global effects of feature-based attention in human visual cortex. *Nature Neuroscience*, *5*(7), 631–632. http://doi.org/10.1038/nn876

Sereno, M. I., Dale, a M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., … Tootell, R. B. H. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, *268*(May), 889–893. http://doi.org/10.1126/science.7754376

Soto, D., & Blanco, M. J. (2004). Spatial attention and object-based attention: A comparison within a single task. *Vision Research*, *44*(1), 69–81. http://doi.org/10.1016/j.visres.2003.08.013

Spence, C., & Driver, J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(4), 1005–1030. http://doi.org/10.1037/0096-1523.22.4.1005

Spence, C., & Driver, J. (1997). Audiovisual links in exogenous covert spatial orienting. *Perception & Psychophysics*, *59*(1), 1–22. http://doi.org/10.3758/BF03206843

Stephen, J. M., Aine, C. J., Christner, R. F., Ranken, D., Huang, M., & Best, E. (2002). Central versus peripheral visual field stimulation results in timing differences in dorsal stream sources as measured with MEG. *Vision Research*, *42*(28), 3059–3074. http://doi.org/10.1016/S0042-6989(02)00415-7

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*, *13*(9), 403–409. http://doi.org/10.1016/j.tics.2009.06.003

Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica*, *135*(2), 77–99. http://doi.org/10.1016/j.actpsy.2010.02.006

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. http://doi.org/10.1016/0010-0285(80)90005-5

van Ede, F., de Lange, F. P., & Maris, E. (2012). Attentional cues affect accuracy and reaction time via different cognitive and neural processes. *Journal of Neuroscience*, *32*(30), 10408–10412. http://doi.org/10.1523/JNEUROSCI.1337-12.2012

van Lamsweerde, A. E., & Beck, M. R. (2011). The change probability effect: Incidental learning, adaptability, and shared visual working memory resources. *Consciousness and Cognition*, *20*(4), 1676–1689. http://doi.org/10.1016/j.concog.2011.09.003

Wahn, B., & König, P. (2015). Audition and vision share spatial attentional resources, yet attentional load does not disrupt audiovisual integration. *Frontiers in Psychology*, *6*(July), 1–12. http://doi.org/10.3389/fpsyg.2015.01084

Wahn, B., & König, P. (2017). Is attentional resource allocation across sensory modalities task-dependent ? *Advances in Cognitive Psychology*, *13*(1), 83–96.

Ward, L. M., McDonald, J. J., & Lin, D. (2000). On asymmetries in cross-modal spatial attention orienting. *Perception & Psychophysics*, *62*(6), 1258–1264. http://doi.org/10.3758/BF03212127

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(1), 121–134. http://doi.org/10.1037/0096-1523.16.1.121

Zuanazzi, A., & Noppeney, U. (2018). Additive and interactive effects of spatial attention and expectation on perceptual decisions. *Scientific Reports*, 1–12. http://doi.org/10.1038/s41598-018-24703-6

Chapter 5:

Distinct neural mechanisms of spatial attention and expectations guide perceptual inference in a multisensory world

Arianna Zuanazzi[1], Uta Noppeney[1]

[1]Computational Cognitive Neuroimaging Laboratory, Computational Neuroscience and Cognitive Robotics Centre, University of Birmingham, B15 2TT Birmingham, UK

Author contributions:

Conceived and designed the experiments: **Arianna Zuanazzi**, Uta Noppeney.

Performed the experiments: **Arianna Zuanazzi**.

Analysed the data: **Arianna Zuanazzi**.

Wrote the manuscript: **Arianna Zuanazzi**.

Contributed to analysing the data and with writing the manuscript: Uta Noppeney.

ABSTRACT

Spatial attention (i.e. task-relevance) and expectation (i.e. signal probability) are two critical top-down mechanisms guiding perceptual inference. Spatial attention prioritises processing of information at task-relevant locations. Spatial expectations encode the statistical structure of the environment. An unresolved question is how the brain allocates attention and forms expectations in a multisensory environment, where task-relevance and signal probability over space can differ across sensory modalities.
This functional magnetic resonance imaging (fMRI) study investigated whether the brain encodes task-relevance and signal probability over space separately or interactively across sensory modalities. In a novel multisensory paradigm we manipulated spatial attention and expectation selectively in audition and assessed their effects on behavioural and neural responses to visual and auditory stimuli.
Our results show that both auditory and visual stimuli increased activations in a right-lateralized frontoparietal system, when they were presented at locations that were task-irrelevant in audition. Yet, only auditory stimuli increased activations in the medial prefrontal cortex when presented at expected locations and in audiovisual and frontoparietal cortices signaling a prediction error when presented at unexpected locations.
This dissociation in multisensory generalization for attention and expectation effects shows that the brain controls attentional resources interactively across the senses, but encodes the statistical structure of the environment as spatial expectations independently for each sensory system. Our results demonstrate that spatial attention and expectation engage partly overlapping neural systems via distinct mechanisms to guide perceptual inference in a multisensory world.

Keywords:
spatial attention · spatial expectation · perceptual decision making · multisensory · functional magnetic resonance imaging

5.1 INTRODUCTION

The brain needs to combine signals from different sensory modalities to enable robust perception in the multisensory environment. Crucially, perceptual processes are not purely driven by bottom-up sensory signals but are also guided by top-down selective attention and prior expectations (or predictions). While attention prioritizes relevant signals and favours efficient allocation of cognitive resources according to current task-demands, expectation encodes the statistical structure of the environment (Summerfield and Egner, 2009). At the behavioural level, both spatial attention and expectation can facilitate effective interactions with the environment, leading to faster and more accurate responses (Geng and Behrmann, 2002, 2005, Carrasco, 2011; Rohenkohl et al., 2014). At the neural level, while attention increases blood oxygenation level-dependent (BOLD) response and signal-to-noise ratio (Grady et al., 1997; O'Leary et al., 1997, Kastner et al., 1999; Ress et al., 2000, Silver et al, 2006), expectation leads to a reduction in stimulus-induced neural activity (Alink et al., 2010; Todorovic et al., 2011).

Crucially, in our natural environment the brain is exposed to a constant influx of signals furnished by all of our five senses. This raises the critical question of how the brain allocates spatial attention and forms spatial expectations in a multisensory environment. Because observers need to respond to stimuli irrespective of the sense by which they are perceived, attentional resources may be allocated interactively across the senses and form an 'amodal map' that encodes the probability of events. In line with this conjecture, parietal cortices have previously been shown to integrate audiovisual signals weighted by their bottom-up sensory reliabilities and top-down task-relevance into audiovisual spatial priority maps (Rohe and Noppeney, 2016).

Likewise, attentional resources were shown to be allocated interactively across the senses. Previous research has shown that spatial attention and orienting relies to some extent on mechanisms that generalize across sensory modalities (Eimer and Schröger, 1998; Calvert et al. 2000, 2004; Macaluso, 2010). For instance, shifting participants' spatial attention endogenously or exogenously in one sensory modality to a particular location facilitates processing at this spatial location also of stimuli from other sensory modalities (Spence and Driver, 1996, 1997; Eimer and Schröger, 1998; McDonald et al.,, 2000; Spence et al., 2000; Ward et al., 2000). Neuroimaging studies provided further evidence for attentional resources that are at least partially shared across sensory modalities. In particular, increased activations for attentional orienting are observed in frontoparietal cortices (Macaluso et al., 2000, 2002a; Macaluso and Driver, 2001, 2005, Wu et al., 2007) and in sensory cortices (Macaluso et al., 2000, 2002b) irrespective of stimulus modality. It has thus been suggested that attentional mechanisms may rely on forward- and backward-projections between early sensory and higher order association areas or on lateral connections between early sensory cortices (Johnson and Zatorre, 2005; Klemen and Chambers, 2012).

Critically, previous research studying the amodal or modality-specific nature of spatial attention manipulated attention by varying the probability of stimulus occurrence over space (Posner, 1980). In the field of unisensory perception it has recently been argued that attention as task-relevance and expectation as stimulus probability need to be dissociated (Summerfield and Egner, 2009). For instance, by manipulating spatial attention and expectation independently in vision, Kok et al. (2012) showed that spatial attention as task-relevance reverses the decrease of neural activity induced by spatial expectation. These intriguing results raise the question as to whether attention and expectation rely on amodal systems (Macaluso et al. 2000,

2002a, 2000b), when they are orthogonally manipulated. In fact, less is known about how the brain forms spatial expectations across sensory modalities. Because information is initially gathered by distinct sensory organs and enters the brain via parallel pathways, each sensory system may initially encode the probability of signals selectively for its preferred sensory modality. These modality-specific spatial expectations may be reinforced particularly in environments where auditory and visual signals arise from separate sources such as in experiments that present unisensory auditory or visual signals independently (e.g., Spence and Driver, 1996).

Hence, the current study investigated (1) whether spatial attention and expectation operate in an amodal or modality-specific fashion and (2) whether spatial attention and expectation rely on shared or distinct neural systems. In our (i) psychophysics experiment and (ii) fMRI experiment, participants detected a sound or a flash presented in either hemifield. Using a novel multisensory approach we orthogonally manipulated spatial attention (i.e. task-relevance) and expectation (i.e. stimulus probability) across the two hemifields selectively in audition and evaluated their effects on observers' neural and behavioural responses to signals presented in either auditory or visual modalities. Our results demonstrate that the neural processes mediating spatial attention and expectations are partly overlapping, but differ in their engagement of modality-specific or amodal processes.

5.2 MATERIALS AND METHODS

5.2.1 Participants

Thirty-one healthy volunteers (23 females; mean age, 21.4 years; range 18–27 years) participated in the psychophysics experiment. All participants had normal or corrected to normal vision, reported normal hearing and had no history of neurological

or psychiatric illness. All participants were right-handed, according to the Edinburgh Handedness Inventory (Oldfield, 1971) (mean laterality index, 84; range 60–100). A subgroup of 22 participants (17 females; mean age, 21.2 years; range 18-27 years) was selected to take part in the fMRI experiment (see Inclusion criteria). Overall, data collection was terminated when 22 participants had undergone the fMRI study. This sample size was determined based on Thirion et al. (2007). All participants provided written informed consent, as approved by the local ethics committee of the University of Birmingham (Science, Technology, Mathematics and Engineering (STEM) Ethical Review Committee) and the experiment was conducted in accordance with these guidelines and regulations.

5.2.2 Stimuli

Auditory spatialized stimuli (100 ms duration) were created by convolving bursts of white noise (with 5 ms onset and offset ramps) with spatially specific head-related transfer functions (HRTFs) based on the KEMAR dummy head of the MIT Media Lab (http://sound.media.mit.edu/resources/KEMAR.html). Visual stimuli (i.e. the so-called 'flashes') were white discs (100 ms duration. radius: 0.88º visual angle, luminance: 165 cd/m2) presented on a grey background (luminance: 78 cd/m2). Both auditory and visual stimuli were presented at ±10º visual angle along the azimuth (0º visual angle for elevation). A fixation cross was presented in the centre of the screen throughout the entire experiment.

5.2.3 Experimental design

In both the psychophysics and the fMRI experiment, we orthogonally manipulated spatial attention (i.e. task-relevance) and expectation (i.e. stimulus probability) across the two hemifields selectively in audition and evaluated their

effects on observers' neural and behavioural responses to signals presented in either auditory or visual modality. Thus, the 2 x 2 x 2 x 2 design manipulated *auditory spatial attention* (left vs right hemifield), *auditory spatial expectation* (left vs right hemifield), *stimulus location* (left vs right hemifield) and *stimulus modali*ty (auditory vs visual, see Fig. 5.1A). Across days, auditory spatial expectation was manipulated as auditory spatial signal probability, i.e. the probability of auditory stimuli to be presented over the left and right hemifield. During the psychophysics and fMRI experiments, auditory stimuli were presented with a ratio of 4/1 in the expected/unexpected hemifields. Observers were not explicitly informed about those probabilities. To boost the implicit learning of auditory spatial signal probability, observers participated in two additional pre-experiment runs prior to the psychophysics and fMRI experiment (with spatial probability ratio of auditory targets: 9/1 in the expected/unexpected hemifields). Auditory spatial attention was manipulated as 'task-relevance', i.e. the requirement to respond to an auditory target in the left vs right hemifield. Critically, spatial attention and expectation were manipulated only in audition but not in vision, thus participants needed to respond to all visual targets presented with equal probabilities in either spatial hemifield (i.e. 1/1 in the expected/unexpected hemifields) (see Fig. 5.1A and 5.1B). A central fixation cross presented throughout the entire experiment coded in colour whether participants should attended and respond to sounds in either their left or right hemifield. The mapping between colour and task-relevant hemifield was counterbalanced across participants.

5.2.4 Experimental procedure

The current study included two experiments: (i) a psychophysics experiment conducted across two days (i.e. auditory spatial expectation was manipulated across the two days) and (ii) an fMRI experiment conducted across two days (i.e. auditory

spatial expectation was manipulated across the two days). The psychophysics experiment was conducted prior to the fMRI experiment. On each day, the psychophysics and the fMRI experiments were preceded by two pre-experiment runs which were used to boost the implicit learning of auditory spatial signal probability. Each pre-experiment and experimental run (duration: ~8mins/run) included 10 attention blocks with 20 trials each, interleaved with fixation baseline periods of 6 s. Blocks were of two types: in block type A, spatial attention and expectation were congruent (i.e. spatial attention was directed to the hemifield with higher auditory target frequency); in block type B, spatial attention and expectation were incongruent (i.e. attention was directed to the hemifield with less frequent auditory targets). Thus, both psychophysics and fMRI experiment included 20 trials x 10 blocks (attention manipulation: 5 blocks of type A and 5 blocks of type B) x 2 pre-experiment runs x 2 days (expectation manipulation) + 20 trials x 10 blocks (attention manipulation: 5 blocks of type A and 5 blocks of type B) x 5 experimental runs x 2 days (expectation manipulation) = 800 trials (pre-experiment runs) + 2000 trials (experimental runs) in total. In the experimental runs, each block type overall included 400 auditory stimuli for the expected hemifield (pooled over left and right) and 100 auditory stimuli for the unexpected hemifield (pooled over left and right). Each block type also included 250 visual stimuli for the expected hemifield and 250 visual stimuli for the unexpected hemifield (pooled over left and right). For further details see Figure 5.1B which shows the absolute number of trials for each condition and block type and their response requirement for the psychophysics and the fMRI experiment.

The order of expectation days (i.e. left vs right) was counterbalanced across participants; the colour (i.e., pink or green) of the color-coded fixation cross (i.e. attention instruction) was counterbalanced across participants, the order of attention

blocks was counterbalanced within and across participants and the order of stimulus location and stimulus modality were pseudo-randomized within each participant. Brief breaks were included after every run to provide feedback to participants about their performance accuracy (averaged across all conditions) in the target detection task. In the psychophysics experiment participants' fixation performance was monitored via eye tracking thus, during the breaks, participants were also provided with a feedback about their eye-movements (i.e. fixation maintenance).

Each trial (SOA: 2200 ms) included three time windows (see Fig. 5.1C): i. the fixation cross alone (700 ms duration), ii. brief flash or sound (stimulus duration: 100 ms) and iii. the fixation cross alone (1400 ms). Participants responded to the auditory targets in the attended hemifield and to all visual targets via key press with their right index finger (i.e. the same response for all auditory and visual targets) as fast and accurately as possible. They fixated the cross in the centre of the screen which was presented throughout the entire experiment.

On each day, participants were first familiarized with the stimuli in brief practice runs (with equal spatial signal probability) to train them on target detection performance and, only in the psychophysics experiment, also on fixation (i.e. a warning signal was shown when the disparity between the central fixation cross and the eye-data samples exceeded 2.5 degrees). After the final fMRI day, participants indicated in a questionnaire whether they thought the sound or the flash was presented more frequently in one of the two spatial hemifields. Eighteen out of the total 22 participants correctly reported that the auditory stimuli were more frequent in one hemifield and 20 out of 22 participants reported the visual stimuli to be equally frequent across the two hemifields, suggesting that most participants were aware of the manipulation of signal probability.

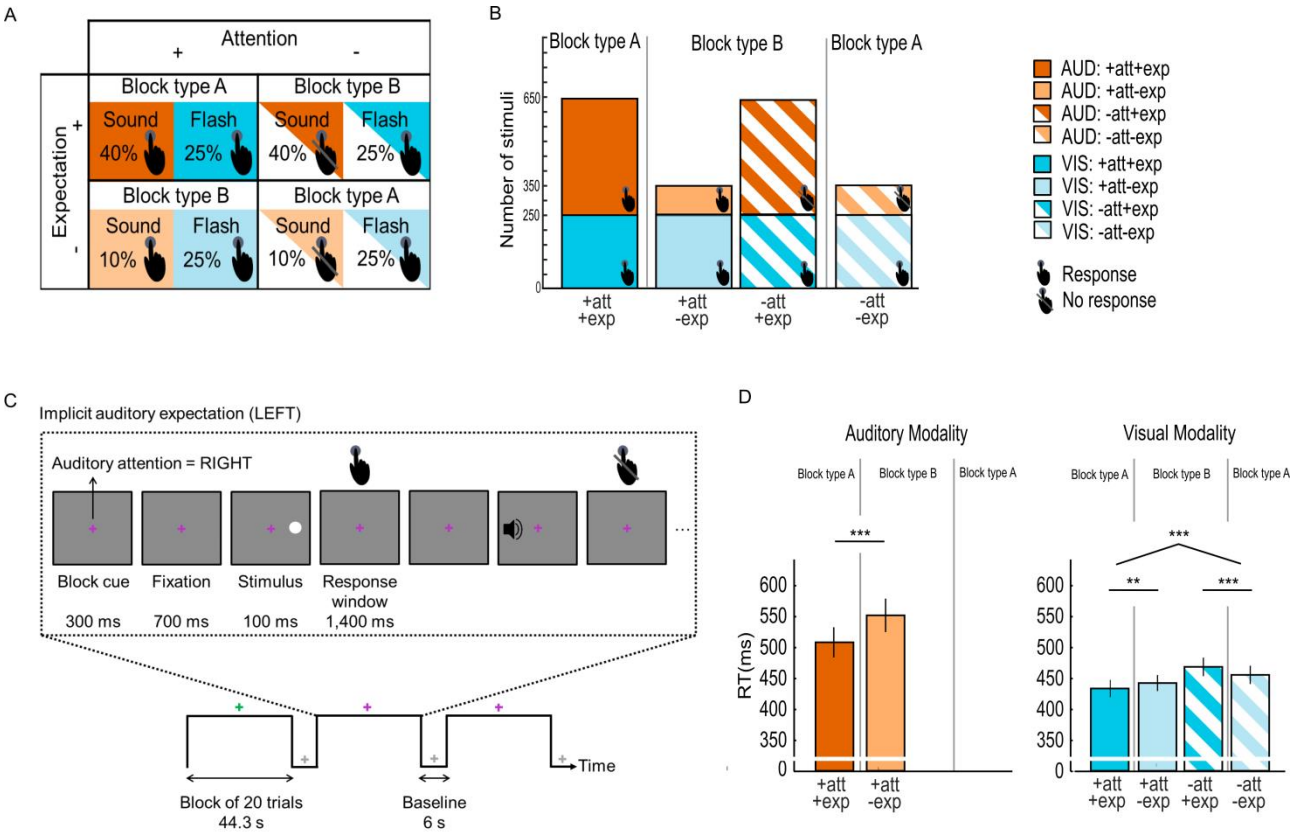


Figure 5.1: Experimental design, example trials and behavioural results of the fMRI experiment.

A. The factorial design manipulated: Auditory spatial attention (attended hemifield – full pattern, vs unattended hemifield – striped pattern), expectation (expected hemifield – dark shade, vs unexpected hemifield– light shade) and stimulus modality (auditory modality – orange, vs visual modality – blue). For illustration purposes, we pooled over stimulus locations (left/right). Presence vs absence of response requirement is indicated by the hand symbol. **B.** Number of auditory (orange) and visual (blue) trials in the 2 (attended vs unattended) x 2 (expected vs unexpected) design. Presence vs absence of response requirement is indicated by the hand symbol. The fraction of trials where a response was required in a particular block type (e.g. block type A), represents the probability that the observer needs to make a response irrespective of the hemifield in which the signal is presented (i.e., 'general response probability'); note that general response probability is greater for block type A where attention and expectation are congruent. **C.** fMRI runs included ten blocks of 20 trials alternating with fixation periods. A fixation cross was presented throughout the entire

run. Its colour indicated: white = fixation period; pink or green = activation period with auditory attention directed to left or right hemifield. On each trial, participants were presented with either an auditory or visual stimulus (100 ms duration) either in their left or right hemifield. They were instructed to respond as fast and accurately as possible with their right finger within a response window of 1400 ms. **D.** Bar plots show response times (across subjects' mean (±SEM)) for each of the six conditions with response requirements for the experimental runs of the fMRI experiment (see Supplementary Table 8.2 and Fig. 8.3 for the results of the psychophysics experiment and for all RT). The brackets and stars indicate significance of main effects and interactions. ** $p < 0.01$; *** $p < 0.001$; audition: orange; vision: blue; attended: full pattern; unattended: striped pattern; expected: dark shade; unexpected: light shade.

5.2.5 Inclusion criteria

A subgroup of 22 participants who had taken part in the psychophysics experiment was selected to take part in the fMRI experiment. Inclusion criteria were participants' accuracy and fixation performance in the psychophysics experiment. Only participants who in the psychophysics experiment produced less than 20 saccades averaged across blocks and showed overall accuracy > 95% (calculated as the percentage of hits + correct rejections pooling over auditory and visual stimuli) were included in the fMRI experiment. For the psychophysics experiment, mean group number of saccades was 22.9 ± 5.2 [across subjects mean ± SEM]) and mean group accuracy was 97% ± 0.2% [across subjects mean ± SEM] for the psychophysics experiment and 97% ± 0.5% [across subjects mean ± SEM] for the fMRI experiment.

5.2.6 Experimental setup

5.2.6.1 Psychophysics experiment

The psychophysics experiment (pre-experiment and experimental runs) was conducted in a darkened room. Participants rested their chin on a chinrest with the height held constant across all the participants. Auditory stimuli were presented at approximately 72 dB SPL, via HD 280 PRO headphones (Sennheiser, Germany). To mimic the scanner environment, the scanner noise was reproduced for the whole duration of the experiment at approximately 80 dB SPL via external loudspeakers. Visual stimuli were displayed on a gamma-corrected LCD monitor (2560 x 1600 resolution, 60 Hz refresh rate, 30" Dell UltraSharp U3014, USA), at a viewing distance of approximately 50 cm from the participant's eyes. Stimuli were presented using Psychtoolbox version 3 (Brainard, 1997; www.psychtoolbox.org), running under Matlab R2014a (Mathworks Inc., Natick, MA, USA) on a Windows machine.

Participants responded to all targets with their right index finger and responses were recorded via one key of a small keypad (Targus, USA). Throughout the study, participants' eye-movements and fixations were monitored using Tobii Eyex eyetracking system (Tobii, Sweden).

5.2.6.2 fMRI experiment.

During the pre-experiment runs, participants lay in a mock scanner, which mimicked all features of the MRI scanner. The scanner noise was reproduced at approximately 80 dB SPL via external loudspeakers. During the experimental runs, participants lay in the MRI scanner. Auditory stimuli were presented at approximately 72 dB SPL using MR-compatible headphones (MR Confon). Visual stimuli were back-projected onto a Plexiglas screen using a BARCO projector (F35). Participants viewed the screen through a mirror mounted on the MR head coil at a viewing distance of 68 cm. Stimuli were presented using Psychtoolbox version 3 (Brainard, 1997; www.psychtoolbox.org), running under Matlab R2014a (Mathworks Inc., Natick, MA, USA) on a MacBook Pro machine. Participants responded to all targets with the same index finger of their right hand and responses were recorded via an MR-compatible keypad (NATA).

5.2.7 fMRI data acquisition

A 3T Philips MRI scanner with 32 channel head coil was used to acquire both T1-weighted anatomical images (TR = 8.4 ms, TE = 3.8 ms, flip angle = 8°, FOV = 288 mm x 232 mm, image matrix = 288 x 232, 175 sagittal slices acquired in ascending direction, voxel size = 1 mm x 1mm x 1 mm) and T2*-weighted axial echoplanar images (EPI) with bold oxygenation level-dependent (BOLD) contrast (TR = 2600 ms, TE = 40 ms, flip angle = 85°, FOV = 240 mm x 240 mm, image matrix 80 x 80, 38 transversal slices acquired in ascending direction, voxel size = 3 x 3 x 3 mm). For each

participant, an overall of 196 volumes x 5 experimental runs x 2 days = 1960 volumes were acquired. The anatomical image volume was acquired at the end of the experiment.

5.2.8 Statistical analysis

5.2.8.1 Behavioural data analysis - psychophysics and fMRI experiments

For the behavioural analysis of the psychophysics experiment, we excluded trials where participants did not successfully fixate the central cross based on a dispersion criterion (i.e. distance of fixation from subject's median of fixation as defined in calibration trials > 1.3 degrees for three subsequent samples, see Blignaut, 2009). Excluded auditory response trials were 1.4% ± 0.4% [across subjects mean ± SEM]; excluded visual response trials were 1.3% ± 0.4% [across subjects mean ± SEM]. For the behavioural analysis of both experiments, response time analysis was limited to correct trials and response times within the range of participant- and condition-specific mean ± two SD and < 1500 ms. For auditory targets in the attended hemifield, median response times for each subject were entered into a two-sided paired-sample t-test with *auditory spatial expectation* (expected vs unexpected stimulus) as factor. For visual targets, median response times for each subject were entered into a 2 (*auditory spatial attention*: attended vs unattended stimulus) x 2 (*auditory spatial expectation*: expected vs unexpected stimulus) repeated measures analysis of variance (rmANOVA). Unless otherwise indicated, we only report effects that are significant at $p < 0.05$.

5.2.8.2 fMRI data analysis

The functional MRI data were analysed with statistical parametric mapping (SPM12; Wellcome Department of Imaging Neuroscience, London; www.fil.ion.ucl.ac.uk/spm, Friston et al., 1995). Scans from each subject were

realigned using the first as a reference, unwarped, slice-time corrected and spatially normalized into MNI standard space using parameters from segmentation of the T1 structural image (Ashburner and Friston, 2005), resampled to a spatial resolution of 2 x 2 x 2 mm^3 and spatially smoothed with a Gaussian kernel of 8 mm full-width at half-maximum. The time series of all voxels were high-pass filtered to 1/128 Hz.

The fMRI experiment was modelled in a mixed event-block related fashion with regressors entered into the day-specific design matrix after convolving each event-related unit impulse with a canonical hemodynamic response function and its first temporal derivative. In addition to modelling the 16 conditions in our 2 (*stimulus modality*: auditory vs visual) x 2 (*auditory spatial attention*: left vs right hemifield) x 2 (*auditory spatial expectation*: left vs right hemifield) x 2 (s*timulus location*: left vs right hemifield) factorial design, the statistical model included the time when the attention cue first appeared (i.e., auditory attention to the left hemifield, auditory attention to the right hemifield) as separate condition. Nuisance covariates included the realignment parameters to account for residual motion artefacts.

Condition-specific effects for each subject were estimated according to the general linear model and passed to a second-level analysis as contrasts. This involved creating 16 contrast images (i.e., each of the 16 conditions relative to fixation, summed over the 10 runs) for each subject and entering them into a second-level ANOVA. Inferences were made at the second level to allow a random-effects analysis and inferences at the population level (Friston et al., 1995).

At the random effects or group level, for each stimulus modality we tested for (i) the main effect of spatial attention (i.e., attended > unattended auditory stimuli and vice versa, attended > unattended visual stimuli and vice versa) and (ii) the main effect of spatial expectation (i.e., expected > unexpected auditory stimuli and vice versa,

expected > unexpected visual stimuli and vice versa). To assess whether these effects of spatial attention and expectation rely on amodal or modality-specific systems, we tested for (i) the effects of attention and expectation that are common for audition and vision (i.e., a logical "AND" conjunction over stimulus modalities) and (ii) the effects that differ between audition and vision (i.e., i. the interaction between attention and stimulus modality and ii. the interaction between expectation and stimulus modality). Next, we investigated whether spatial attention and expectation rely on shared or different neural systems. Separately for each stimulus modality we tested for (i) the effects that are common for attention and expectation (i.e., a logical "AND" conjunction over each attention and expectation main effects) and (ii) the interaction between attention and expectation.

Unless otherwise stated, we report activations at $p < 0.05$ at the cluster level corrected for multiple comparisons within the entire brain using an auxiliary (uncorrected) voxel threshold of $p < 0.001$.

5.2.8.3 ROI analysis

Based on our a priori hypothesis that spatial attention and expectation influence activations in primary sensory cortices, we also tested the effects of auditory spatial attention and expectation selectively within the primary auditory cortex and primary visual cortex. These areas of interests were defined using ROI maps from SPM Anatomy Toolbox (version 2.2b, Eickhoff et al., 2005). The anatomical mask for the primary auditory cortex encompassed 890 voxels in the bilateral cytoarchitectonic maps TE 1.0, TE 1.1 and TE 1.2. The anatomical mask for the primary visual cortex encompassed 2936 voxels in the bilateral cytoarchitectonic maps hOC1. We extracted parameter estimates from each ROI, for each of the 16 conditions relative to fixation and for each subject and we entered them into a 2 (*auditory spatial attention*: attended

vs unattended stimulus) x 2 (*auditory spatial expectation*: expected vs unexpected stimulus) rmANOVA, separately for each stimulus modality and pooling over stimulus location.

5.3 RESULTS

In the following, we report (1) the behavioural results of the psychophysics and the fMRI experiment and (2) the imaging results of the fMRI experiment.

5.3.1 Behavioural results – psychophysics and fMRI experiments

For both psychophysics and fMRI experiments, the two-sided paired-sample t-tests on response times for auditory stimuli in the attended hemifield showed significantly faster responses when this hemifield was expected than unexpected (psychophysics: $t(30) = -4.56$, $p < 0.001$, Cohen's d_{av} [95% CI] = -0.40 [-0.59 , -0.19]; fMRI: $t(21) = -5.06$, $p < 0.001$, Cohen's d_{av} [95% CI] = -0.36 [-0.54, -0.18], left panel of Fig. 5.1D, Supplementary Fig. 8.3 and Table 8.2).

For both psychophysics and fMRI experiments, the 2 (attended vs unattended) x 2 (expected vs unexpected) rmANOVA on response times for visual simuli revealed a significant main effect of attention (psychophysics: $F(1, 30) = 109.88$, $p < 0.001$, η_p^2 [90% CI] = 0.79 [0.64, 0.84]; fMRI: $F(1, 21) = 78.69$, $p < 0.001$, η_p^2 [90% CI] = 0.79 [0.61, 0.85]). Participants responded faster to visual stimuli in their attended than unattended hemifield. Moreover, a significant crossover interaction between attention and expectation was observed (psychophysics: $F(1, 30) = 41.59$, $p < 0.001$, η_p^2 [90% CI] = 0.58 [0.36, 0.69]; fMRI: $F(1, 21) = 49.29$, $p < 0.001$, η_p^2 [90% CI] = 0.70 [0.47, 0.79]). The simple main effects showed that participants responded significantly faster (psychophysics: $t(30) = -5.46$, $p < 0.001$, Cohen's d_{av} [95% CI] = -0.20 [-0.30, -0.11];

fMRI: $t(21) = -3.94$, $p = 0.001$, Cohen's d_{av} [95% CI] = -0.14 [-0.22, -0.06]) to visual targets in the attended hemifield when this hemifield was expected than unexpected (right panel of Fig. 5.1D, Supplementary Fig. 8.3 and Table 8.2). By contrast, they responded significantly more slowly (psychophysics: $t(30) = 5.44$, $p < 0.001$, Cohen's d_{av} [95% CI] = 0.22 [0.12, 0.32]; fMRI: $t(21) = 5.79$, $p < 0.001$, Cohen's d_{av} [95% CI] = 0.18 [0.09, 0.26]) to visual targets in the unattended hemifield, when this hemifield was expected than unexpected (right panel of Fig. 5.1D, Supplementary Fig. 8.3 and Table 8.2). Importantly, we observed equivalent response time results in the psychophysics and the fMRI experiment.

For completeness, we also analysed the behavioural results of the pre-experiment runs, i.e., runs prior to psychophysics and fMRI experiments used to boost the implicit learning of auditory spatial signal probability, where auditory stimuli were presented with a ratio of 9/1 in the expected/unexpected hemifields and visual stimuli with a ratio of 1/1 in the expected/unexpected hemifields. Results showed the same significant patterns of the experimental runs, suggesting that the behavioural effects of spatial attention and expectation are qualitatively similar when the probability ratio is 0.9/0.1 or 0.4/0.1.

5.3.2 fMRI results

5.3.2.1 Effects of auditory attention separately for auditory and visual stimuli

We first evaluated the main effect of spatial attention, separately for each stimulus modality. For auditory stimuli, spatial attention increased activations in bilateral thalami, caudates, hippocampi, left frontoparietal operculum, left putamen and in a motor network encompassing left central sulcus and right cerebellum (Supplementary Fig. 8.4 and Table 8.3). The increased activations for auditory stimuli in motor areas can be explained by the motor responses that were given only in the

attended hemifield. Conversely, because visual stimuli required a motor response in both hemifields, no attentional effects were observed in the motor network for visual stimuli.

Spatial reorienting or shifting attention to an auditory stimulus presented in the unattended hemifield (i.e., unattended vs attended auditory stimuli) induced activations in a predominantly right lateralized dorsal frontoparietal system encompassing the bilateral superior frontal gyri (SFG) and sulci (SFS) extending into the precentral sulci and significant parietal activations in the right postcentral sulcus extending into the right intraparietal sulcus (IPS) and into the right inferior parietal lobule (IPL). Significant activation increases were also observed in the right anterior cingulate cortex (ACC), right middle frontal gyrus (MFG) and bilateral insulae (INS) (orange and yellow on the inflated brain in Fig. 5.2 and Supplementary Table 8.3). Likewise, shifting attention to a visual stimulus presented in the unattended hemifield (i.e., unattended vs attended visual stimuli) increased activations in a more bilateral frontoparietal network including bilateral SFG, bilateral SFS extending into the precentral sulci, bilateral postcentral sulci extending into IPS, bilateral ACC and right anterior INS (blue and yellow on the inflated brain in Fig. 5.2 and Supplementary Table 8.3). For completeness, no significant main effects of auditory attention for auditory or visual stimuli were found in primary auditory or visual cortices even in our selective ROI analysis.

5.3.2.2 Effects of auditory attention: commonalities and differences between auditory and visual stimuli

Next, we investigated the extent to which the neural systems engaged by attention shifts are common (i.e. amodal) or distinct (i.e. modality-specific) for auditory and visual stimuli. The conjunction analysis over sensory modalities showed

increased activations for attention shifts (i.e., ([unattended > attended auditory stimuli] ∩ [unattended > attended visual stimuli]) in the bilateral SFG and SFS, right ACC, right postcentral sulci extending into IPS, and right anterior INS (Table 5.1 and Fig. 5.2). Testing for the interaction between attention and stimulus modality (i.e. ([unattended > attended auditory stimuli] > [unattended > attended visual stimuli]) revealed additional activations in the right anterior INS that was also part of the spatial reorienting system shared across sensory modalities. Table 5.1 and Figure 5.2 show the significant clusters of the conjunction analysis: (i) right postcentral sulcus/IPS (yellow outline on the inflated SPM template, highlighted in green in the axial slice) and (ii) right SFG (yellow outline on the inflated SPM template, highlighted in green in the axial slice). The interaction between attention and stimulus modality in the right INS are displayed in Table 5.1 and in Fig. 5.2 (white outline on the inflated SPM template, highlighted in green in the axial slice).

In summary, in line with our behavioural results, our fMRI analysis suggests that the effect of auditory spatial attention generalizes to processing stimuli in vision, where spatial attention was not explicitly manipulated. Critically, spatial reorienting to both auditory and visual stimuli increased activations in a widespread right lateralized system encompassing frontoparietal cortices, INS and ACC (Corbetta and Shulman, 2002; Shomstein and Yantis, 2006; Corbetta et al., 2008; Shulman et al., 2009; Banerjee et al., 2011). While the right INS exhibited a stronger effect to spatial reorienting to auditory stimuli, we did not observe activations that were truly selective for stimuli from either the visual or auditory modality. Collectively, these results suggest that spatial attention and reorienting relies predominantly on amodal neural processes that are more engaged for stimuli from the sensory modality where attention is actively manipulated, but generalize across sensory modalities.

Table 5.1: 'Amodal' and modality dependent mechanisms of auditory spatial attention. *p*-values are FWE corrected at the cluster level for multiple comparisons within the entire brain. Auxiliary uncorrected voxel threshold of $p < 0.001$. L, Left; R, right.

	MNI coordinates					
Brain regions	***x***	***y***	***z***	**z-score (peak)**	**Cluster size**	***p*- value (cluster)**
[unattended > attended auditory stimuli] ∩ [unattended > attended visual stimuli]						
R superior frontal gyrus	18	-4	64	4.09	731	0.000
R superior frontal sulcus	28	-6	46	4.38		
R anterior cingulate gyrus	10	18	36	3.67		
L superior frontal gyrus	-14	-10	64	3.69	268	0.005
L superior frontal sulcus	-30	-8	48	4.07		
R postcentral sulcus/ R intraparietal sulcus	42	-32	40	3.74	304	0.003
R anterior insula	30	20	6	4.41	185	0.027
[unattended > attended auditory stimuli] > [unattended > attended visual stimuli]						
R anterior insula	38	16	-12	4.32	209	0.016

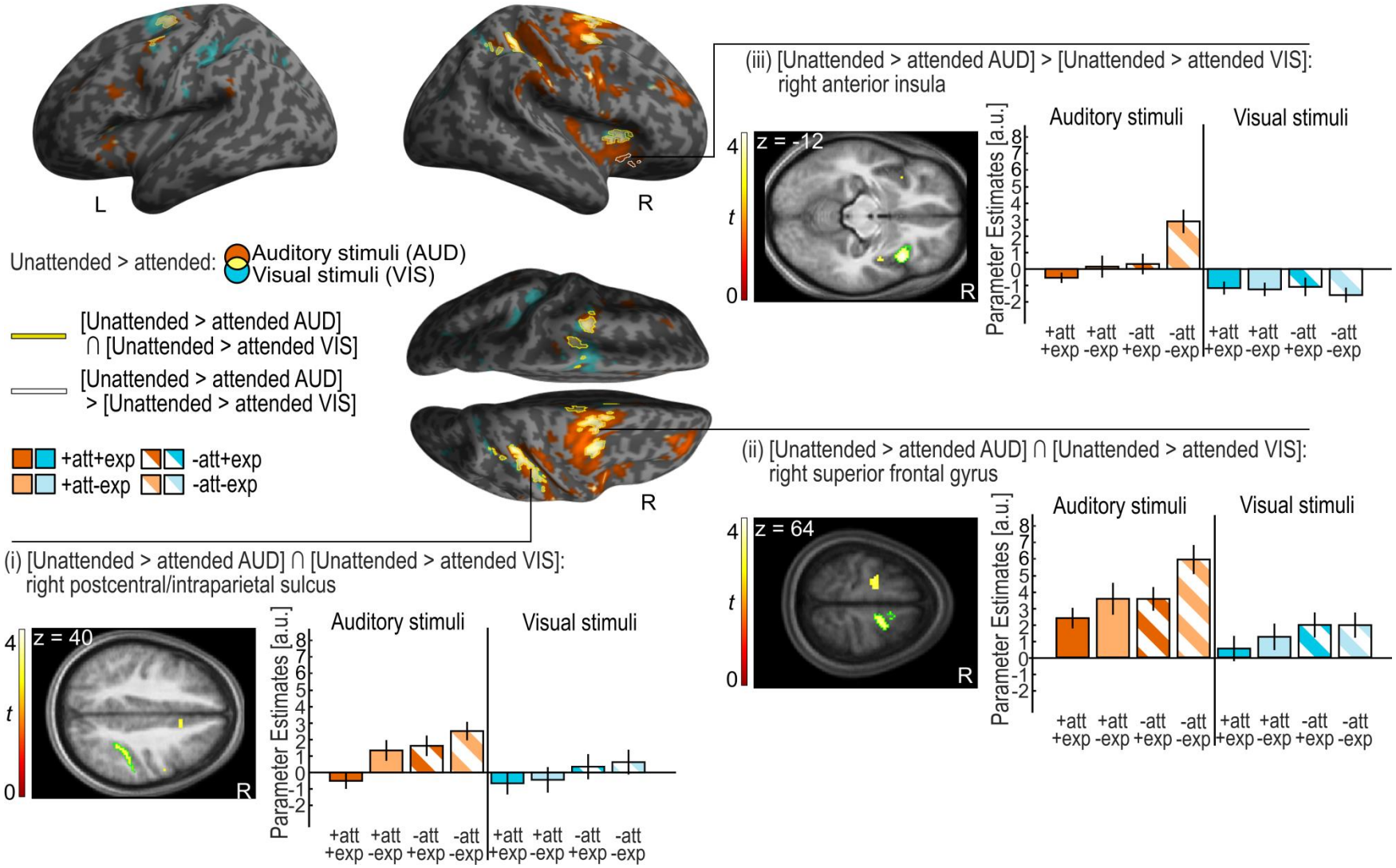

L
R
Unattended > attended:
Auditory stimuli (AUD)
Visual stimuli (VIS)
[Unattended > attended AUD]
∩ [Unattended > attended VIS]
[Unattended > attended AUD]
> [Unattended > attended VIS]
+att+exp
-att+exp
+att-exp
-att-exp
R
(i) [Unattended > attended AUD] ∩ [Unattended > attended VIS]:
right postcentral/intraparietal sulcus
z = 40
4
t
0
R
Parameter Estimates [a.u.]
Auditory stimuli
Visual stimuli
+att +exp
+att -exp
-att +exp
-att -exp
(iii) [Unattended > attended AUD] > [Unattended > attended VIS]:
right anterior insula
z = -12
(ii) [Unattended > attended AUD] ∩ [Unattended > attended VIS]:
right superior frontal gyrus
z = 64

Figure 5.2: Unattended > attended for auditory and visual stimuli.

Activation increases for unattended > attended stimuli for auditory (AUD, orange, height threshold: $p < 0.001$, uncorr.) and visual (VIS, blue, height threshold: $p < 0.001$, uncorr.) stimuli (overlap: yellow) are rendered on an inflated canonical brain. The conjunction of unattended > attended for auditory and visual stimuli is encircled in yellow (height threshold: $p < 0.001$, uncorr.). Activation increases for unattended > attended that are greater for auditory than visual stimuli (i.e. interaction) are encircled in white (height threshold: $p < 0.001$, uncorr). Bar plots show the parameter estimates (across participants mean ± SEM, averaged across all voxels in the green encircled cluster) in the (i) right poscentral/intraparietal sulcus, (ii) right superior frontal gyrus and (iii) right anterior insula that are displayed on axial slices of a mean image created by averaging the subjects' normalized structural images. The bar graphs represent the size of the effect pertaining to BOLD magnitude in nondimensional unit (corresponding to percentage whole-brain mean). Audition: orange; vision: blue; attended: full pattern; unattended: striped pattern.

5.3.2.3 Effects of auditory spatial expectation separately for auditory and visual stimuli

Auditory stimuli in the expected relative to unexpected hemifield elicited significantly greater activation in the bilateral anterior portions of the SFG (i.e., medial prefrontal cortices, mPFC) and the bilateral precunei extending into the posterior cingulate gyri (Table 5.2, Fig. 5.3, see Summerfield et al., 2006).

Table 5.2: Main effects of auditory spatial expectation for auditory stimuli. *p*-values are FWE corrected at the cluster level for multiple comparisons within the entire brain. Auxiliary uncorrected voxel threshold of $p < 0.001$. L, Left; R, right.

	MNI coordinate					
Brain regions	***x***	***y***	***z***	**z-score (peak)**	**Cluster size**	***p*- value (cluster)**
Expected > unexpected auditory stimuli						
R superior frontal gyrus	8	54	18	4.91	1458	0.000
L superior frontal gyrus	-6	54	36	5.80		
R precuneus	-4	-52	26	3.90	260	0.006
L precuneus	6	-56	26	3.28		
Unexpected > expected auditory stimuli						
R superior temporal gyrus	60	-44	16	7.47	18305	0.000
L superior temporal gyrus	-62	-34	14	5.44		
R postcentral sulcus/ R intraparietal sulcus	34	-58	46	5.93		
L postcentral sulcus/ L intraparietal sulcus	-38	-46	42	5.79		
R precuneus	4	-54	54	6.49		
L precuneus	-8	-54	54	6.48		
R anterior insula	38	16	2	7.48		
L anterior insula	-32	16	2	6.94		
R posterior cingulate gyrus/ L posterior cingulate gyrus	4	-28	26	5.18	339	0.001
R anterior cingulate gyrus	8	22	32	5.40	4222	0.000
R superior frontal gyrus	18	2	66	4.14		
L superior frontal gyrus	-26	-8	70	4.23		
L precentral sulcus	-38	0	38	5.08		
R precentral sulcus	40	6	30	4.97	2186	0.000
R middle frontal gyrus	40	34	36	4.42		
L middle frontal gyrus	-34	46	24	4.56	810	0.000
R calcarine cortex	12	-84	8	3.75	680	0.000
L calcarine cortex	-12	-84	6	3.71		

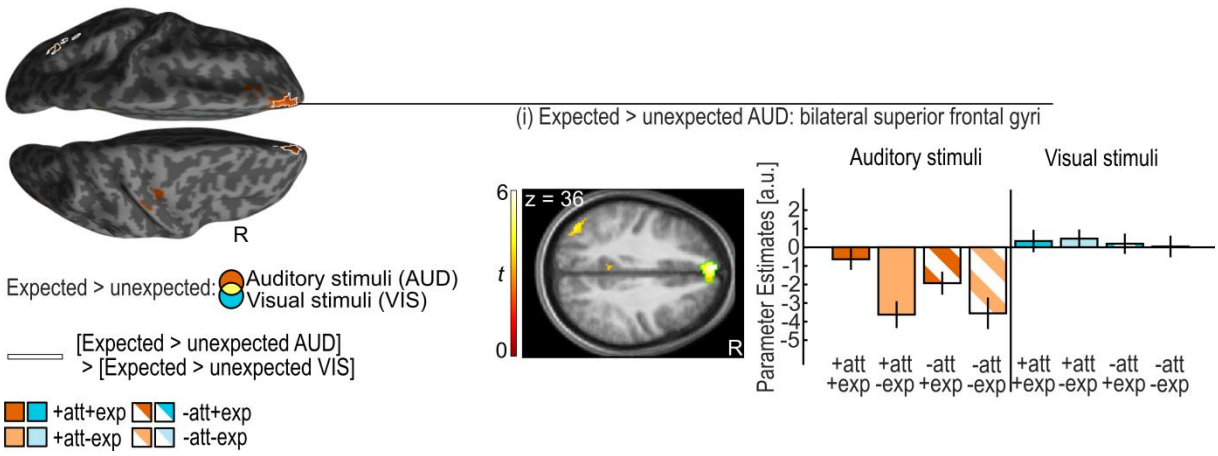


Figure 5.3: Expected > unexpected for auditory and visual stimuli.
Activation increases for expected > unexpected stimuli for auditory stimuli (orange) are rendered on an inflated canonical brain; they are encircled in white, if they are significantly greater for auditory than visual stimuli (i.e. interaction). Height threshold of $p < 0.001$, uncorrected; extent threshold k > 0 voxels.
Bar plots show the parameter estimates (across participants mean ± SEM, averaged across all voxels in the green encircled cluster) in superior frontal gyri that are displayed on axial slices of a mean image created by averaging the subjects' normalized structural images; the bar graphs represent the size of the effect in nondimensional unit (corresponding to percentage whole-brain mean). Audition: orange; vision: blue; attended: full pattern; unattended: striped pattern; expected: dark shade; unexpected: light shade.

By contrast, auditory stimuli in the unexpected relative to expected hemifield increased activations in a widespread frontoparietal system encompassing bilateral postcentral sulci extending into bilateral IPS and to the bilateral IPL, bilateral precuneus, bilateral anterior INS, bilateral anterior and posterior cingulate gyri, bilateral SFG and precentral sulci, bilateral MFG extending into bilateral superior and inferior frontal sulci (Table 5.2, Fig. 5.4). Moreover, we observed pronounced activations for unexpected relative to expected auditory stimuli in the bilateral plana temporalia in the superior temporal gyri (STG), which were previously shown to be

involved in spatial processing (Griffiths and Warren, 2002; Brunetti et al., 2005; Ahveninen et al., 2006), and surprisingly also in the bilateral calcarine cortices (CC). Our selective ROI analysis also revealed higher activations for unexpected relative to expected auditory stimuli in primary auditory and visual cortices (main effect of expectation in the primary auditory cortex: $F(1, 21) = 12.67$, $p = 0.002$, η_p^2 [90% CI] = 0.38 [0.11, 0.56] and the primary visual cortex: $F(1, 21) = 7.21$, $p = 0.014$, η_p^2 [90% CI] = 0.26 [0.03, 0.46]). Surprisingly, neither whole-brain nor ROI analysis revealed any significant effects of spatial expectation for visual stimuli.

5.3.2.4 Effects of auditory expectation: commonalities and differences between auditory and visual stimuli.

Our results suggest that a widespread neural system forms spatial expectations selectively for stimuli from the auditory modality where signal probability was manipulated. Indeed, this was confirmed by the significant interaction between expectation and stimulus modality ([unexpected > expected auditory stimuli] > [unexpected > expected visual stimuli]) that was observed in large parts of the neural system showing expectation effects for auditory stimuli (see areas on the inflated brain with white outline in Fig. 5.4). By contrast, the conjunction analyses over stimulus modality did not reveal any significant effects of auditory expectation that were common to auditory and visual stimuli.

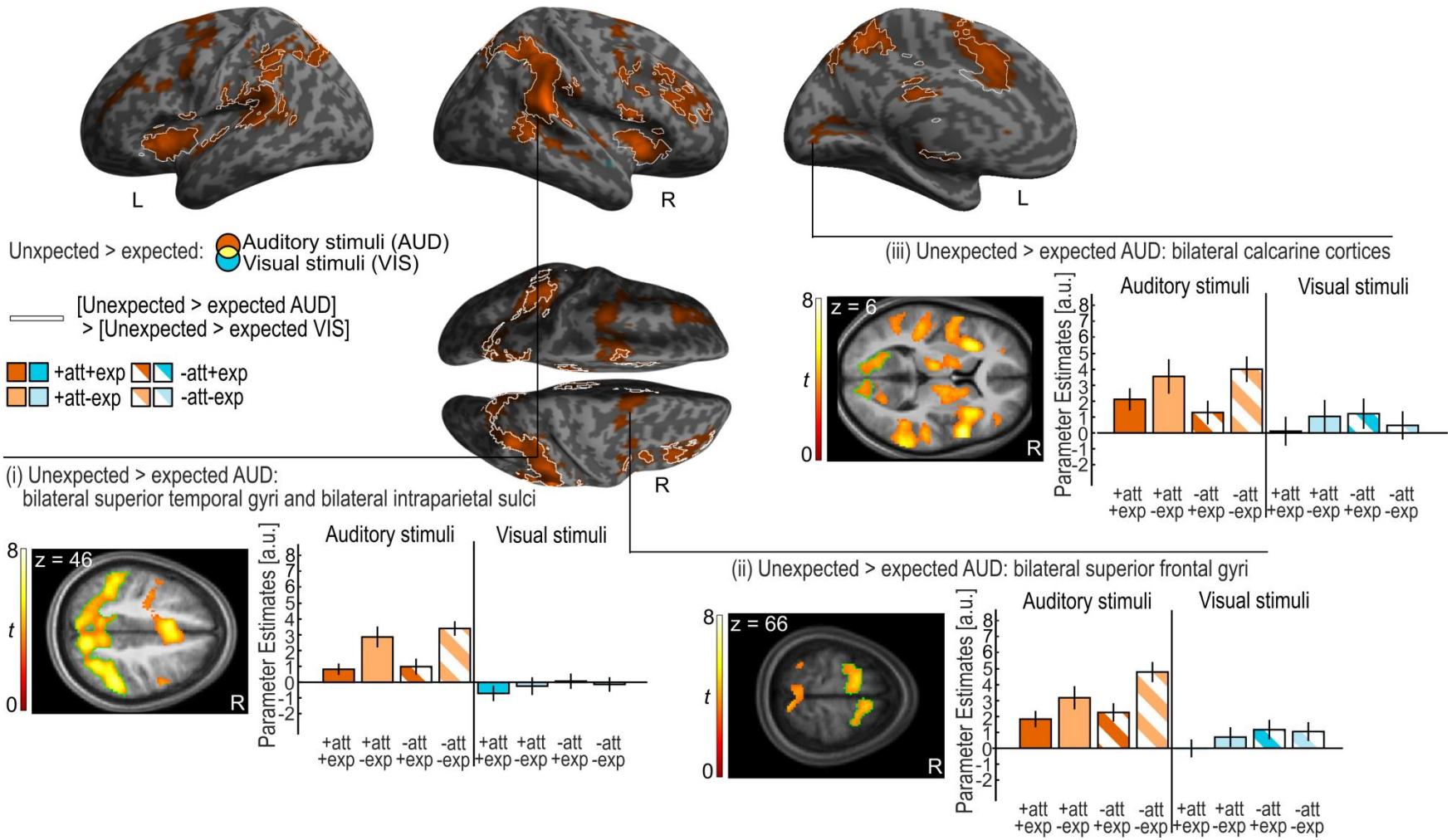

L
R
L
Unxpected > expected:
Auditory stimuli (AUD)
Visual stimuli (VIS)
[Unexpected > expected AUD]
> [Unexpected > expected VIS]
+att+exp
-att+exp
+att-exp
-att-exp
R
(iii) Unexpected > expected AUD: bilateral calcarine cortices
z = 6
t
8
0
R
Parameter Estimates [a.u.]
Auditory stimuli
Visual stimuli
+att +exp
+att -exp
-att +exp
-att -exp
(i) Unexpected > expected AUD:
bilateral superior temporal gyri and bilateral intraparietal sulci
z = 46
(ii) Unexpected > expected AUD: bilateral superior frontal gyri
z = 66

Figure 5.4: Unexpected > expected for auditory and visual stimuli.

Activation increases for unexpected > expected stimuli for auditory stimuli (orange) are rendered on an inflated canonical brain; they are encircled in white, if they are significantly greater for auditory than visual stimuli (i.e. interaction). Height threshold of $p < 0.001$, uncorrected; extent threshold k > 0 voxels.

Bar plots show the parameter estimates (across participants mean ± SEM, averaged across all voxels in the green encircled cluster) in (i) bilateral superior temporal gyri and bilateral intraparietal sulci, (ii) bilateral superior frontal gyri and (iii) bilateral calcarine cortices that are displayed on axial slices of a mean image created by averaging the subjects' normalized structural images. The bar graphs represent the size of the effect in nondimensional unit (corresponding to percentage whole-brain mean). Audition: orange; vision: blue; attended: full pattern; unattended: striped pattern; expected: dark shade; unexpected: light shade.

5.3.2.5 Additive and interactive effects of spatial attention and expectation - separately for auditory and visual stimuli

Finally we investigated separately for auditory and visual stimuli whether attention and expectation effects engage common neural systems in an additive (i.e. conjunction over attention and expectation) or interactive fashion (i.e. interaction between attention and expectation).

For auditory stimuli, neither the whole-brain nor the selective ROI analysis revealed any significant interaction between attention and expectation. By contrast, the conjunction analysis over attention and expectation revealed activation increases jointly for unattended > attended and unexpected > expected in a predominantly right-lateralized frontoparietal system including the bilateral SFG, right ACC, right MFG extending into right inferior frontal sulcus, right postcentral sulcus extending into right IPS and right IPL and bilateral INS (Fig. 5.5A).

For visual stimuli, the whole-brain analysis did not reveal any significant additive or interactive effects for expectation and attention. The ROI analysis revealed a significant interaction between attention and expectation in the primary visual cortex ($F(1, 21) = 5.06$, $p = 0.035$, η_p^2 [90% CI] = 0.19 [0.01, 0.41]), with greater activations for unexpected than expected visual stimuli in the attended hemifield, but greater activation for expected than unexpected visual stimuli in the unattended hemifield (Fig. 5.5B).

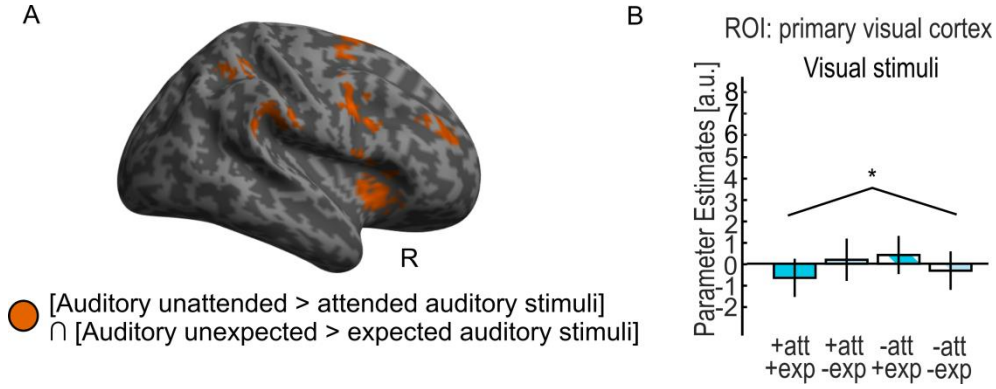


Figure 5.5: Additive and interactive effects of auditory attention and expectation in audition and vision for auditory and visual stimuli.

A. Activation increases for the conjunction of attention and expectation main effects in the auditory modality ([unattended > attended auditory stimuli] ∩ [unexpected > expected auditory stimuli]) are rendered in orange on an inflated canonical brain; height threshold of $p < 0.001$, uncorrected; extent threshold k > 0 voxels. **B.** Bar plots show the parameter estimates (across participants mean ± SEM) averaged across all voxels in the primary visual cortex. The bar graphs represent the size of the effect in nondimensional unit (corresponding to percentage whole-brain mean); Vision: blue; attended: full pattern; unattended: striped pattern; expected: dark shade; unexpected: light shade. The brackets and stars indicate significance of the interaction. * $p < 0.05$.

5.4 DISCUSSION

The current study was designed to investigate whether spatial attention and expectation rely on neural systems that are specific or shared across sensory modalities. Moving beyond previous research, we developed a novel multisensory paradigm that orthogonally manipulated spatial attention as response requirement and expectation as stimulus probability over space selectively in audition. We assessed their effects on neural and behavioural responses to auditory and visual stimuli.

Behaviourally, our results show that even though spatial attention and expectation were manipulated selectively in audition, their effects generalized to the visual modality. With regard to spatial attention, participants were significantly faster to respond to auditory and visual stimuli that were presented in the hemifield where auditory stimuli were task-relevant than irrelevant. In other words, directing observers' spatial attention to one hemifield selectively in audition impacted participants' behavioural responses to auditory and visual stimuli pointing towards attentional resources that are at least partly shared across sensory modalities (Spence and Driver, 1996, 1997). Likewise, at the neural level auditory and visual responses depended on whether or not the stimuli were presented in the hemifield that was task-relevant in audition. More specifically, auditory and visual stimuli in the task-irrelevant as compared to task-relevant hemifield induced activation increases in a widespread right-lateralized system encompassing the bilateral SFG and SFS, right postcentral sulcus extending into IPS and inferior parietal cortices, right ACC and right anterior INS. These results dovetail nicely with previous research showing dorsal frontoparietal activations and right ventral parietal / temporoparietal activations for attentional shifts, spatial reorienting and contextual updating (Nobre et at., 2000; Macaluso et al., 2002a;

Kincade et al. 2005; Bressler et al., 2008; Smith et al., 2010; Langner et al., 2011). Yet, those previous studies did not operationally dissociate spatial attention from stimulus probability. By contrast, by orthogonally manipulating stimulus probability and task-relevance, the current study allowed us to attribute the frontoparietal activations selectively to shifts in spatial attention rather than stimulus probability. Critically, even though unattended visual stimuli increased dorsal parietal activations more bilaterally and unattended auditory stimuli evoked more right temporoparietal activations, these differences between stimulus modalities were subtle and mostly not significant when formally testing for an interaction between attention and stimulus modality. Only the right anterior insula increased activations for stimuli in the task-irrelevant than relevant hemifield significantly more in audition than vision. This enhancement of attentional effects for auditory than visual stimuli most likely results from the fact that attention was selectively manipulated in audition and only generalized to vision.

Collectively, our behavioural and neural results suggest that the brain relies on attentional resources and systems that are largely shared across sensory modalities. Visual stimuli evoke frontoparietal activations when presented in the task-irrelevant hemifield, even though both hemifields are equally task-relevant in vision. Conversely, auditory stimuli evoke frontoparietal activations when presented in the task-irrelevant hemifield, even though they never require any response and should hence be ignored. Thus, our results corroborate the hypothesis that the brain has only limited abilities to define spatial locations as task-relevant or irrelevant independently for audition and vision (Macaluso et al., 2000, 2002a; Macaluso and Driver, 2005; Wu et al., 2007; Krumbholz et al., 2009; Banerjee et al., 2011). The (at least to some extent) 'amodal' definition of spatial task-relevance may also explain the extensive activations that we

observed for ‘unattended’ stimuli not only in the ventral, but also the dorsal attentional network that is typically more associated with focused attention. Greater focused attention may be required for the ‘auditory unattended’ hemifield, because the brain needs to decide whether to respond (i.e. when a visual stimulus is presented) or not to respond (i.e. when an auditory stimulus is presented) (Indovina and Macaluso, 2007; Santangelo et al., 2009). In summary, our behavioural and neuroimaging finding suggest that spatial attention when defined as task-relevance operates interactively across the senses.

Next, we asked whether spatial expectations are formed in a modality-specific fashion or jointly across sensory systems. Behaviorally, we observed faster responses to expected than unexpected stimuli irrespective of sensory modality in the task-relevant hemifield. Yet, surprisingly we observed faster responses for unexpected than expected visual stimuli in the task-irrelevant hemifield (n.b. auditory stimuli do not require a response in the task-irrelevant hemifield). Hence, we observed a significant interaction between attention and expectation for visual response time. As discussed in Zuanazzi and Noppeney (2018) this interaction for visual response times results most likely from the differences in the overall response probability across conditions. Thus, the response probability is greater when attention and expectation are congruent and directed to the same hemifield (90% of the trials in blocks of type A) than when they are directed to different hemifields (60% of the trials in blocks of type B, Fig. 5.1A and 5.1B).

Critically, fMRI allows us to move beyond response time and track neural processes across the entire auditory and visual processing hierarchy from sensory to higher order decisional processes for both visual and auditory stimuli and even for auditory stimuli in task-irrelevant hemifields where no behavioural response is

obtained. This provides us with the opportunity to investigate whether lower cortical areas generate a prediction error signal for spatially unexpected stimuli in a more automatic fashion irrespective of spatial attention. Indeed, in line with the notion of predictive coding (Rao and Ballard, 1999; Friston, 2005), spatially unexpected relative to expected auditory stimuli increased activations as a prediction error signal in the plana temporalia and primary visual cortices as well as higher order frontoparietal areas. These results suggest that the planum temporale forms spatial expectations and prediction error signals selectively for audition. These prediction errors may then propagate up the hierarchy (Friston, 2005) into frontoparietal areas. Surprisingly, we observed activation increases signaling a prediction error only for spatially unexpected stimuli in audition. In fact, even the visual cortex showed activation increases only for unexpected auditory stimuli potentially mediated via direct connectivity from auditory areas or top-down modulation from parietal cortices.

Likewise, activation increases for spatially expected stimuli were observed selectively for the auditory sense in the mPFC that has previously been implicated in forming a representation that matches one's expectations (Summerfield et al., 2006). Hence, in line with the notion of predictive coding, higher order areas such as mPFC form representations when stimuli match our expectations, while sensory and parietal areas signal a prediction error when stimuli violate our spatial expectations (Rao and Ballard, 1999; Friston, 2005). Critically, spatial expectations and prediction error signals were formed in a modality-specific fashion selectively for the sensory modality where stimulus probability was explicitly manipulated. In fact, we did not observe any significant positive or negative effects of spatial expectations for visual stimuli anywhere in the brain. These results suggest that the neural systems form and compute

spatial expectations that encode the probability of stimulus occurrence in a modality-specific fashion.

The results discussed so far suggest that the neural systems mediating spatial attention and expectation are sustained by partly overlapping frontoparietal networks, but differ substantially in their modality-specificity. Hence, we finally asked separately for audition and vision whether spatial attention and expectations influence these neural responses in an additive or interactive fashion. In the auditory modality, spatial attention and expectation did not interact, but influenced neural responses in this system in an additive fashion (Shulman et al., 2009). Our conjunction analysis over spatial attention and expectation revealed a dorsal and ventral frontoparietal network that was jointly recruited by spatial reorienting as well as expectation violations in audition. By contrast, in primary visual areas we observed a significant interaction between spatial attention and expectation selectively for visual stimuli. Activations for visual stimuli were greater when attention and expectation were directed to different than same hemifields. This activation profile mimics the pattern we observed for behavioural response times and is thus most likely mediated by top-down influences from response selection processes onto sensory cortices (Supplementary Fig. 8.4; Roelfsema and de Lange, 2016).

In conclusion, our results demonstrate that spatial attention and expectation engaged partly overlapping neural systems, yet differed in their modality-specificity. Attentional resources were controlled interactively across audition and vision within a widespread right-lateralized frontoparietal system. By contrast, spatial expectations and prediction error signals were formed in audiovisual and frontoparietal cortices selectively for auditory stimuli where stimulus probability was explicitly manipulated. Collectively, our results suggest that spatial attention and expectations are intimately

linked in perception. Yet, while attention engages neural systems interactively across sensory modalities, spatial stimulus probability is mostly encoded and processed in a modality-specific fashion.

REFERENCES

Ahveninen J, Jaaskelainen IP, Raij T, Bonmassar G, Devore S, Hamalainen M, Levanen S, Lin F-HH, Sams M, Shinn-Cunningham BG, Witzel T, Belliveau JW (2006) Task-modulated “what” and “where” pathways in human auditory cortex. Proc Natl Acad Sci U S A 103:14608–14613.

Alink A, Schwiedrzik CM, Kohler A, Singer W, Muckli L (2010) Stimulus predictability reduces responses in primary visual cortex. J Neurosci 30:2960–2966.

Ashburner J, Friston KJ (2005) Unified segmentation. Neuroimage 26:839–851.

Banerjee S, Snyder AC, Molholm S, Foxe JJ (2011) Oscillatory alpha-band mechanisms and the deployment of spatial attention to anticipated auditory and visual target locations: supramodal or sensory-specific control mechanisms? J Neurosci 31:9923–9932.

Blignaut P (2009) Fixation identification: the optimum threshold for a dispersion algorithm. Atten Percept Psychophys 71:881–895.

Brainard DH (1997) The Psychophysics Toolbox. Spat Vis 10:433–436.

Bressler SL, Tang W, Sylvester CM, Shulman GL, Corbetta M (2008) Top-down control of human visual cortex by frontal and parietal cortex in anticipatory visual spatial attention. JNeurosci 28:10056–10061.

Brunetti M, Belardinelli P, Caulo M, Del Gratta C, Della Penna S, Ferretti A, Lucci G, Moretti A, Pizzella V, Tartaro A, Torquati K, Olivetti Belardinelli M, Romani GL (2005) Human brain activation during passive listening to sounds from different locations: An fMRI and MEG study. Hum Brain Mapp 26:251–261.

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

Carrasco M (2011) Visual attention: The past 25 years. Vision Res 51:1484–1525.

Corbetta M, Patel G, Shulman GL (2008) The reorienting system of the human brain: from environment to theory of mind. Neuron 58:306–324.

Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci 3:201–215.

Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K (2005) A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage 25:1325–1335.

Eimer M, Schroger E (1998) ERP effects of intermodal attention and cross-modal links in spatial attention. Psychophysiology 35:313–27.

Friston K (2005) A theory of cortical responses. Philos Trans R Soc Lond B Biol Sci 360:815–836.

Friston KJ, Holmes AP, Worsley KJ, Poline J-P, Frith CD, Frackowiak RSJ (1995) Statistical parametric maps in functional imaging : A general linear approach. Hum Brain Mapp 2:189–210.

Geng JJ, Behrmann M (2002) Probability cuing of target location facilitates visual search implicitly in normal participants and patients with hemispatial neglect. Psychol Sci 13:520–525.

Geng JJ, Behrmann M (2005) Spatial probability as an attentional cue in visual search. Percept Psychophys 67:1252–1268.

Grady CL, Van Meter JW, Maisog JM, Pietrini P, Krasuski J, Rauschecker JP (1997) Attention-related modulation of activity in primary and secondary auditory cortex. Neuroreport 8:2511–2516.

Griffiths TD, Warren JD (2002) The planum temporale as a computational hub. Trends Neurosci 25:348–353.

Indovina I, Macaluso E (2007) Dissociation of stimulus relevance and saliency factors during shifts of visuospatial attention. Cereb Cortex 17:1701–1711.

Johnson JA, Zatorre RJ (2005) Attention to simultaneous unrelated auditory and visual events: Behavioral and neural correlates. Cereb Cortex 15:1609–1620.

Kastner S, Pinsk MA, De Weerd P, Desimone R, Ungerleider LG (1999) Increased activity in human visual cortex during directed attention in the absence of visual stimulation. Neuron 22:751–761.

Kincade JM, Abrams RA, Astafiev S V., Shulman GL, Corbetta M (2005) An event-related functional magnetic resonance imaging study of voluntary and stimulus-driven orienting of attention. J Neurosci 25:4593–4604.

Klemen J, Chambers CD (2012) Current perspectives and methods in studying neural mechanisms of multisensory interactions. Neurosci Biobehav Rev 36:111–133.

Kok P, Rahnev D, Jehee JFM, Lau HC, de Lange FP (2012) Attention reverses the effect of prediction in silencing sensory signals. Cereb Cortex 22:2197–2206.

Krumbholz K, Nobis EA, Weatheritt RJ, Fink GR (2009) Executive control of spatial attention shifts in the auditory compared to the visual modality. Hum Brain Mapp 30:1457–1469.

Langner R, Kellermann T, Boers F, Sturm W, Willmes K, Eickhoff SB (2011) Modality-specific perceptual expectations selectively modulate baseline activity in auditory, somatosensory, and visual cortices. Cereb Cortex 21:2850–2862.

Macaluso E (2010) Orienting of spatial attention and the interplay between the senses. Cortex 46:282–297.

Macaluso E, Driver J (2001) Spatial attention and crossmodal interactions between vision and touch. Neuropsychologia 39:1304–1316.

Macaluso E, Frith C, Driver J (2000) Selective spatial attention in vision and touch: unimodal and multimodal mechanisms revealed by PET. J Neurophysiol 83:3062–3075.

Macaluso E, Frith CD, Driver J (2002a) Supramodal effects of covert spatial orienting triggered by visual or tactile events. J Cogn Neurosci 14:389–401.

Macaluso E, Frith CD, Driver J (2002b) Crossmodal spatial influences of touch on extrastriate visual areas take current gaze direction into account. Neuron 34:647–658.

McDonald JJ, Teder-Sälejärvi W a, Hillyard S a (2000) Involuntary orienting to sound improves visual perception. Nature 407:906–908.

Nobre AC, Gitelman DR, Dias EC, Mesulam MM (2000) Covert visual spatial orienting and saccades: overlapping neural systems. Neuroimage 11:210–216.

O'Leary DS, Andreasen NC, Hurtig RR, Torres IJ, Flashman LA, Kesler ML, Arndt S V., Cizadlo TJ, Ponto LLB, Watkins GL, Hichwa RD (1997) Auditory and visual attention assessed with PET. Hum Brain Mapp 5:422–436.

Oldfield R.C. (1971) The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9:97–113.

Posner MI (1980) Orienting of attention. Q J Exp Psychol 32:3–25.

Rao RPN, Ballard DH (1999) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat Neurosci 2:79–87.

Ress D, Backus BT, Heeger DJ (2000) Activity in primary visual cortex predicts performance in a visual detection task. Nat Neurosci 3:940–945.

Roelfsema PR, de Lange FP (2016) Early visual cortex as a multiscale cognitive blackboard. Annu Rev Vis Sci 2:131–151.

Rohe T, Noppeney U (2016) Distinct computational principles govern multisensory integration in primary sensory and association cortices. Curr Biol 26:509–514.

Rohenkohl G, Gould IC, Pessoa J, Nobre AC (2014) Combining spatial and temporal expectations to improve visual perception. J Vis 14:1–13.

Santangelo V, Olivetti Belardinelli M, Spence C, Macaluso E (2009) Interaction between voluntary and stimulus-driven spatial attention mechanism across sensory modalities. J Cogn Neurosci 21:2384–2397.

Shomstein S, Yantis S (2006) Parietal cortex mediates voluntary control of spatial and nonspatial auditory attention. J Neurosci 26:435–439.

Shulman GL, Astafiev S V, Franke D, Pope DLW, Abraham Z, Mcavoy MP, Corbetta M (2009) Interaction of stimulus-driven reorienting and expectation in ventral and dorsal fronto-parietal and basal ganglia-cortical networks. J Neurosci 29:4392–4407.

Silver MA, Ress D, Heeger DJ (2007) Neural correlates of sustained spatial attention in human early visual cortex. J Neurophysiol 97:229–237.

Smith D V, Davis B, Niu K, Healy EW, Bonilha L, Fridriksson J, Morgan PS, Rorden C (2010) Spatial attention evokes similar activation patterns for visual and auditory stimuli. J Cogn Neurosci 22:347–361.

Spence C, Driver J (1996) Audiovisual links in endogenous covert spatial attention. J Exp Psychol Hum Percept Perform 22:1005–1030.

Spence C, Driver J (1997) Audiovisual links in exogenous covert spatial orienting. Percept Psychophys 59:1–22.

Spence C, Pavani F, Driver J (2000) Crossmodal links between vision and touch in covert endogenous spatial attention. J Exp Psychol Hum Percept Perform 26:1298–1319.

Summerfield C, Egner T (2009) Expectation (and attention) in visual cognition. Trends Cogn Sci 13:403–409.

Summerfield C, Egner T, Greene M, Koechlin E, Mangels J, Hirsch J (2006) Predictive codes for forthcoming perception in the frontal cortex. Science (80-) 314:1311–1314.

Thirion B, Pinel P, Mériaux S, Roche A, Dehaene S, Poline J-B (2007) Analysis of a large fMRI cohort: Statistical and methodological issues for group analyses. Neuroimage 35:105–120.

Todorovic A, van Ede F, Maris E, de Lange FP (2011) Prior expectation mediates neural adaptation to repeated sounds in the auditory cortex: an MEG study. J Neurosci 31:9118–9123.

Ward LM, McDonald JJ, Lin D (2000) On asymmetries in cross-modal spatial attention orienting. Percept Psychophys 62:1258–1264.

Wu CT, Weissman DH, Roberts KC, Woldorff MG (2007) The neural circuitry underlying the executive control of auditory spatial attention. Brain Res 1134:187–198.

Zuanazzi A, Noppeney U (2018) Additive and interactive effects of spatial attention and expectation on perceptual decisions. Sci Rep:1–12.

Chapter 6:

No evidence that transcranial magnetic stimulation of right anterior intraparietal sulcus alters signal reliability or ventriloquist effect

Arianna Zuanazzi[1], Axel Thielscher[2,3], Uta Noppeney[1]

[1]Computational Cognitive Neuroimaging Laboratory, Computational Neuroscience and Cognitive Robotics Centre, University of Birmingham, B15 2TT Birmingham, UK; [2]Department of Electrical Engineering, Technical University of Denmark, 2800 Lyngby, Denmark; [3]DRCMR, Copenhagen University Hospital Hvidovre, 2650 Hvidovre, Denmark

Author contributions:

Conceived and designed the experiment: **Arianna Zuanazzi**, Uta Noppeney.

Contributed with designing the experiment: Axel Thielscher.

Performed the experiment: **Arianna Zuanazzi**.

Analysed the data: **Arianna Zuanazzi**.

Wrote the manuscript: **Arianna Zuanazzi**.

ABSTRACT

The involvement of posterior parietal cortex in attention orienting and multisensory processes is supported by neuroimaging research. In particular, the ventriloquist effect paradigm has been employed to demonstrate that the anterior portion of the intraparietal sulcus (IPS) is involved in combining sensory signals taking into account the causal structure of the environment. However, neuroimaging data cannot prove the causal involvement of the parietal cortex in multisensory processes and previous neurostimulation studies showed contrasting results.

During the after-effect of offline low-frequency transcranial magnetic stimulation (TMS) applied to the right anterior IPS, participants reported the location of auditory or visual signals in (1) unisensory, (2) spatially congruent or (3) spatially incongruent audiovisual trials. We evaluated whether, by altering the cortical excitability of right anterior IPS, behavioural changes in terms of unisensory reliability or ventriloquist effect could be observed.

We replicated previous results emphasising the central role of sensory reliability, spatial proximity and attentional mechanisms in multisensory processing. However our results showed that neither sensory reliability nor audiovisual binding was affected by IPS-TMS. We interpret these results in light of our experimental design and analysis and with regard to the stimulation protocol that we employed.

Keywords:

transcranial magnetic stimulation · intraparietal sulcus · sensory reliability · modality-specific attention · ventriloquist effect · audiovisual integration

6.1 INTRODUCTION

Perceptual decisions are influenced by bottom up information coming from sensory inputs. In a multisensory world, the brain maps sensory signals in time and space and binds signals which derive from the same event into a more reliable unified sensory percept (Wallace et al., 2004). This has multiple consequences at the behavioural level. Facilitation (e.g., shorter reaction times, increased accuracy and sensitivity) is observed for detection and localization of multisensory signals (Gielen et al., 1983; Vroomen & de Gelder, 2000; Frassinetti et al., 2002; Teder-Sälejärvi et al., 2005). Importantly, these effects strictly depend on the physical or perceptual spatial and temporal correspondence of such multisensory signals (Wallace et al., 2004; Bolognini et al., 2007). Moreover, they depend on sensory reliability, with stronger facilitation effects for congruent but less reliable sensory stimuli (i.e., Inverse Effectiveness, Meredith & Stein, 1983; Bolognini et al., 2007; Charbonneau et al., 2013).

When signals are presented simultaneously but are spatially discrepant, multisensory interactions take place in the form of illusions such as the *ventriloquist effect* (VE), where the location of a sound is perceived as shifted towards a simultaneously presented but uninformative visual signal (Bertelson, et al., 1998; Alais & Burr, 2004; Vroomen & de Gelder, 2004). When the location of the visual signal has to be reported, a visual shift towards the uninformative sound may also be observed, although the magnitude of such effect is smaller than that observed for auditory signals towards visual signals (Bertelson & Radeau, 1981). Asymmetric effects for audition and vision can be attributed to the general dominance of vision over audition (Alais et al., 2010). Furthermore, the magnitude of the observed

displacement depends on the spatial and temporal disparity between sensory signals (Wallace et al., 2004) and on their reliability and salience (Alais & Burr, 2004; Bolognini et al., 2007; Charbonneau et al., 2013; Rohe & Noppeney, 2015b). In fact, stimuli that are close in space generally originate from a common source and, since vision provides very reliable spatial information, visual signals are favoured and a perceptual shift of the auditory signal (i.e., VE) is hence produced (Welch & Warren, 1980; Knudsen & Brainard, 1995). Moreover, it has been shown that information coming from different sensory modalities is integrated in a near-optimal fashion and the perceived location is the result of the weighted sum of sensory information which takes into account sensory reliabilities (Alais & Burr, 2004). Top-down influence of attention also plays a role in multisensory integration. Previous studies showed that modality-specific attention (i.e., to attend and report only one signal during bimodal stimulation) can in fact increase the weight of the attended sensory signal (Vercillo & Gori, 2015; Rohe & Noppeney, 2018).

Critically, it would be detrimental if the brain estimated the location of signals exclusively under the assumption that they originate from a common source. Conversely, it has been suggested that the brain more efficiently combines spatial estimates taking into account the causal structure of the environment (i.e., whether spatial estimates rely on integration or segregation, Bayesian Causal Inference model, Körding et al, 2007). A recent study showed that full segregation, forced fusion (i.e., common source assumption) and integration or segregation based on the causal structure of the environment are performed at the level of sensory cortices, posterior and anterior IPS, respectively (Rohe & Noppeney, 2015a). The general involvement of the parietal cortex in multisensory processes is widely supported by neuroimaging studies. Parietal cortex activations (especially IPS) are observed for auditory, visual

and tactile stimulation in isolation (Bremmer et al., 2001; Klemen & Chambers, 2012), are enhanced when stimuli are spatially and temporally congruent (Calvert et al., 2001) and are related to supramodal modulation (i.e., regardless of stimulus modality) of spatial attention (Macaluso et al. 2002; Santangelo et al., 2009; Smith et al., 2010). However, despite being involved in multisensory processes, the parietal cortex hosts neural representations for vision that are more reliable than those for other sensory modalities (Swisher et al., 2007; Rohe & Noppeney, 2018).

A key question is whether the parietal cortex is causally involved in multisensory processes, given that fMRI studies can only provide correlational information about the involvement of a given brain area in a cognitive task. TMS can be employed to provide causal evidence for such involvement by directly manipulating cortical excitability (Pascual-Leone et al., 1999; Wagner et al., 2007). Previous studies showed altered multisensory enhancement (Ro et al., 2004; Bolognini & Maravita, 2007, Pasalar et al., 2010), sound-induced flash illusion (Kamke et al., 2012) and synesthetic congruency effects (Bien et al., 2012) following TMS application to the parietal cortex.

Given the recent importance attributed to anterior IPS with regard to integration and segregation of sensory signals in function of the causal structure of the environment (Rohe and Noppeney, 2015a), this study employed TMS to investigate the causal involvement of right anterior IPS in unisensory and multisensory processes in relation to task-relevance and audiovisual disparity. During the after-effect of offline low-frequency TMS, we instructed participants to report the location of an auditory or visual signal (i.e., task-relevance condition) in the following conditions: (1) presented in isolation (2) presented with a temporally and spatially congruent signal in a second sensory modality (i.e., vision or audition) (3) presented with a temporally congruent but spatially incongruent signal in a second sensory modality (i.e., vision or audition).

We aimed to test whether IPS-TMS altered (i) the reliability of unisensory stimuli and (ii) the magnitude of the VE.

6.2 MATERIALS AND METHODS

6.2.1 Participants

Twenty-five healthy subjects (14 females; mean age = 25.24 years; range 19-44 years) participated in the study. All participants had normal or corrected to normal vision, reported normal hearing, had no history of neurological or psychiatric illness and had no contraindications to TMS as highlighted in the safety questionnaire. All participants were right-handed, according to the Edinburgh Handedness Inventory (Oldfield, 1971) (mean laterality index, 92; range 70-100).

Twenty-two participants (13 females; mean age = 24.23 years; range 19-32 years) were selected to take part in the TMS experiment (see Design and Procedure). Overall, data collection was terminated when 22 participants had undergone the TMS experiment. The sample size was determined based on previous neurostimulation studies investigating multisensory processes (Chambers et al., 2007; Bolognini et al., 2009; Bertini et al., 2010; Pasalar et al., 2010; Bien et al., 2012; Kamke et al., 2012). All participants provided written informed consent, as approved by the local ethics committee of the University of Birmingham (Science, Technology, Mathematics and Engineering (STEM) Ethical Review Committee) and the experiment was conducted in accordance with these guidelines and regulations.

6.2.2 Stimuli and Apparatus

During the screening session and the TMS experiment, participants sat in front of two LCD monitors (1920 x 1080 resolution, 60 Hz refresh rate, 24" Iiyama ProLite

B2481HS), resting their chin on a chinrest with the height held constant. To better resemble a semicircular stimulus displacement, the two monitors were tilted of -25º (monitor on the left) and 25º (monitor on the right) with respect to the axial midline parallel to the participants' eye position. Throughout the entire experiment, a fixation light (0.8º diameter) was presented in the middle of the two monitors at 0º of vertical visual angle at a viewing distance of approximately 50 cm (Fig. 6.1B).

Auditory stimuli were bursts of white noise of 40 ms duration (with 5 ms onset and offset ramps) presented at approximately 70 dB SPL via four pairs of external speakers (MEMTEQ, Shenzhen, China) aligned vertically above and below the left monitor. The four pairs of speakers were located at -8°, -16°, -24° and -32° of the participants' midline (i.e., on the left of the fixation light). Visual stimuli ('flashes') were white discs (radius: 2.5º visual angle) of 40 ms duration presented on a black background. They were presented at -8°, -16°, -24° and -32° of the participants' midline (0º of vertical visual angle) in correspondence to the vertical locations of the speakers. We limited the presentation of the auditory and visual stimuli to the left hemifield because previous research showed that right parietal TMS elicits different behavioural effects for targets presented on the contralateral or ipsilateral hemifields (Hilgetag et al., 2001; Thut et al., 2005; Kim et al., 2005; Zuanazzi & Cattaneo, 2017).

Stimuli were presented using Psychtoolbox version 3 (Brainard, 1997; www.psychtoolbox.org), running under Matlab R2017a (Mathworks Inc., Natick, MA, USA) on a Windows machine. Participants responded by pressing one of four keys of a keypad with four fingers (index, medium, ring and little finger) of their right hand (i.e., dominant hand). The order of the keys corresponded to the order of the stimuli location, from left to right. Throughout the study, participants' eye-movements and

fixation were monitored using the EyeLink eyetracking system (SR Research, 2007, 2000 Hz sampling rate).

6.2.3 Design and Procedure

6.2.3.1 Practice and screening session

Participants were initially familiarized with the stimuli in three brief practice runs and trained on target localization performance. Sounds and flashes were presented in two unisensory blocks, 6 trials for each of the four spatial locations for an overall of 48 trials for each practice run (24 sounds and 24 flashes) and participants were instructed to localise the auditory or visual spatial signal as accurately as possible, while maintaining fixation. After each response, feedback (green or red light presented at fixation light's location) indicated whether the response was correct. If incorrect, the same stimulus was repeated and participants were given the possibility to change their answer until the correct location was identified. Moreover, participants received verbal feedback about their fixation performance. At the end of each practice run, they were given feedback about their performance accuracy for each modality and location.

During the screening session participants were tested for their ability to correctly localise unisensory auditory and visual stimuli. The screening sessionwas preceeded by one practice session and resembled the practice session except that no feedback was provided. Participants were presented with three unisensory blocks of sounds alternated to three of flashes, with 45 trials for each of the four spatial locations for each modality for an overall of 360 trials. They were instructed to localize the auditory or visual spatial signal as accurately as possible, while maintaining fixation. Participants' localization error was quantified for each sensory modality by computing the root mean square error (RMSE) between participant's response and stimulus location, averaged across locations. Only participants whose RMSE was less than 8°

for sounds and 4° for flashes (thresholds were defined on a previous pilot study with 10 participants) were included in the TMS experiment. For the screening session, mean group RMSE (for 25 participants) was 5.3° ± 0.2° [across subjects mean ± SEM] for sounds and 3.2° ± 0.2° [across subjects mean ± SEM] for flashes.

6.2.3.2 TMS experiment

A subgroup of 22 participants who had taken part in the screening session was selected to take part in the TMS experiment. In a spatial ventriloquist paradigm, participants were presented with spatially congruent or discrepant visual and auditory signals. Temporal congruency was not manipulated, thus auditory and visual signals were always synchronous. Our 4 x 4 x 2 x 2 factorial design manipulated (i) the *location of the auditory stimulus* (-8°, -16°, -24°, -32°) and (ii) the *location of the visual stimulus* (-8°, -16°, -24°, -32°). (iii) *TMS* (IPS or sham-TMS), (iv) *task-relevance* block (report auditory or visual location). On each trial, sounds and flashes were presented on one of four locations on the left of the fixation light, resulting in four levels of audio-visual spatial discrepancy (i.e., 0°, 8°, 16°, 24°) (Fig. 6.1A). Each combination of sound and flash locations was presented with equal probability. In each task-relevance block, sounds or flashes were also randomly presented in isolation at each of the four locations (i.e., unisensory trials). Participants were instructed to selectively localize the auditory or visual spatial signal (i.e., task-relevance manipulation) as accurately as possible, while maintaining fixation. Each trial (SOA: 2040 ms) included three time windows: i. fixation light alone (500 ms duration), ii. brief sound, flash or sound and flash together (stimulus duration: 40 ms) and iii. fixation light alone (1500 ms as response interval). TMS conditions were manipulated across experimental sessions that were performed on different days and each condition included two sessions for an overall of four days. IPS-TMS and sham-TMS order was

alternated, with at least 24 hours between IPS- and sham-TMS sessions and seven days between the two IPS-TMS sessions. Task-relevance, i.e. the requirement to locate sounds or flashes, was manipulated over 12 blocks within each session. Each block contained 60 trials. Prior to each block, a cue (duration: 5000 ms) informed participants which modality to report (Fig. 6.1B). Thus, the experiment included 60 trials x 12 task-relevance blocks (6 blocks where the location of the sound was reported and 6 where the location of the flash was reported) x 4 TMS sessions (2 IPS and 2 sham) = 2880 trials in total. The order of TMS conditions across sessions and of task-relevance within each session was counterbalanced across participants and the order of the combination of auditory and visual signals' location was pseudo-randomized within each participant.

Prior to each session, participants were familiarized with the stimuli in three brief practice runs (same practice runs to that delivered before the screening session). Moreover, they were familiarized with the spatial ventriloquist task in one brief practice run (40 trials overall, 20 for each task-relevance condition). No feedback was provided for the ventriloquist task practice. To measure potential differences between subjective effects of sham and real stimulation, after each session we asked participants to answer the question: *During stimulation, how often did you experience: (1) Neck pain, (2) Headache (3) Pinching sensation where the coil was positioned, (4) Heat sensation where the coil was positioned, (5) Loud clicking noise, (6) Fatigue, (7) Faint feeling, (8) Drowsiness*, by rating each item on a 10-point Likert scale (0: never, 10: the whole time) (Fig. 6.2). To evaluate whether participants became aware of the stimulation difference, after the final session participants completed a questionnaire where (1) they indicated whether they noticed any difference during the stimulation period across sessions and (2) after being informed about the application of two

different types of stimulation, they were asked to indicate (or guess) in which sessions they believed sham-TMS was applied. A comparison of the responses given after sham-TMS and IPS-TMS (2-tailed Wilcoxon Signed Ranks tests) did not show significant difference between any of the pairs ($p > 0.05$) (across subject means and SEMs are shown in Fig. 6.1C). Nine out of 22 participants reported a difference between stimulation conditions and eight of them correctly identified the two sessions where sham-TMS was applied.

6.2.4 TMS site and procedure

Two TMS conditions were included in the experiment: stimulation of right anterior IPS as experimental condition and sham stimulation as control condition. For each subject, individual stimulation coordinates were determined by (1), inverse transforming the right IPS3 and IPS4 probabilistic maps from Wang et al. (2015) from MNI to native space using the parameters obtained from spatial normalization of individual T1-weighted structural MRI images, (2) identifying the coordinates (in native space) of the voxel with highest probability located in right anterior IPS, defined as the sum of the inverse transformed IPS3 and IPS4 probabilistic maps (mean coordinates in MNI space are x = 20, y = -64, z = 60) and (3) computing the stimulation entry as the intersection between the scalp and the shortest vector connecting the individual coordinates to the scalp.

Preceding the task, offline low-frequency repetitive TMS was applied in one single continuous train for 30 minutes at 1 Hz frequency. This stimulation protocol was shown to decreases cortical excitability (Maeda et al., 2000) with related worsening or improvement of performance in behavioural and psychophysical tasks (Hilgetag et al., 2001; Pell et al., 2011). We choose an offline approach to avoid that TMS-specific effects, such as lateralized clicking noise, could interfere with sensory

task performance. In terms of after-effects duration, studies showed that low-frequency repetitive TMS (i.e., ≤1 Hz) can reduce cortical excitability for up to 50-100% the duration of the stimulation (Eisenegger et al., 2008; see Robertson et al., 2003 and Hoogendam et al., 2010 for a review). For this reason, task duration was matched to stimulation duration.

During stimulation, participants' head movement was minimized using cushioned supports around their head and, to maximize stimulation accuracy, the TMS coil was attached to a mechanical arm. Participants were provided with earplugs, were allowed to close their eyes and asked to relax. The ventriloquist task was performed for 30 minutes after stimulation, within the putative inhibition-window of TMS after-effects. TMS was delivered via a MagStim figure-of-eight coil (MCF-B70) and a MagPro X100 stimulator (MagVenture Company, Denmark) positioned over the stimulation site with the coil handle held posterolaterally at about 45° from the mid-sagittal axis. The positioning of the TMS coil and real time monitoring of coil position was performed using a neuronavigation system (Localite, Bonn, Germany).

Unlike for visual/motor cortex stimulation, there is no direct way (similar to motor-evoked potentials or phosphenes) to measure intensity effects of TMS delivered over IPS. Therefore, following a standard approach, we based the intensity of IPS-TMS on the individual resting motor threshold (RMT) (Rossi et al., 2009). Yet, the existence of correlations between motor cortex and parietal cortex thresholds is controversially discussed (but see Oliveri et al., 2000 and Oliver et al., 2009). Stimulation intensity was set at 100% of individual RMT, determined as the minimum TMS intensity applied over the right motor cortex required to evoke a visible twitch of the participant's left hand in approximately five out of ten consecutive pulses. Mean TMS intensity across participants was 33% stimulator output (range 24%-42%).

Sham-TMS was performed by positioning the coil on the same spot as for IPS-TMS and by tilting it by 90° with the coil wings touching the scalp. Such positioning has been shown to produce similar acoustic and sensational effects to real TMS without inducing biological activity in the cortex (Lisanby et al., 2001). During sham-TMS, stimulation intensity was set to the same level as for IPS-TMS.

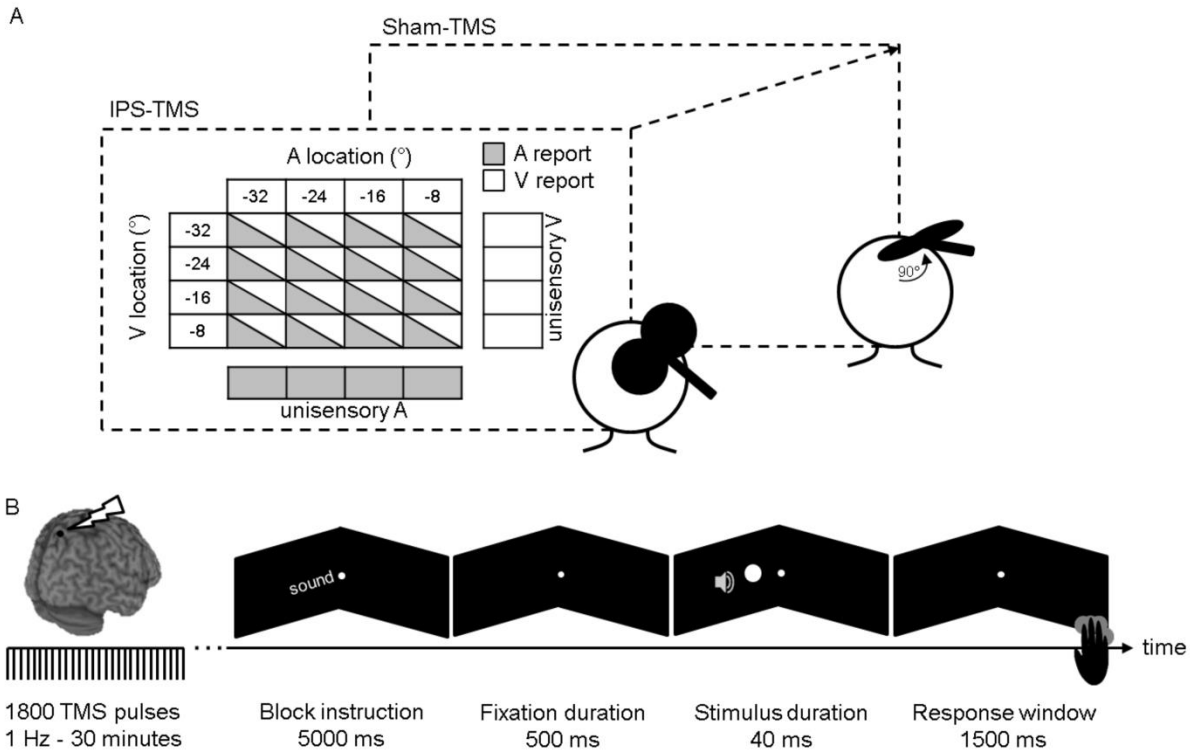


Figure 6.1: Experiment design and example trial.

A. The 4 x 4 x 2 x 2 factorial design manipulated (i) in multisensory conditions, the location of the auditory stimulus (-8°, -16°, -24°, -32°), (ii) in multisensory conditions, the location of the visual stimulus (-8°, -16°, -24°, -32°), (iii) task-relevance (report auditory or visual location) manipulated across blocks, and (iv) TMS application (IPS or sham-TMS) manipulated across sessions which took place on different days. Each task-relevance block included also auditory or visual unisensory trials. Sham-TMS was performed with the coil tilted by 90° with the wings touching the scalp. **B.** The task was preceded by 30 minutes offline stimulation at 1 Hz. During the task, participants were presented with blocks of audiovisual signals and unisensory auditory or visual signals originating from four possible locations. Before each block a cue indicated the task-relevant sensory modality (i.e., whether to locate the sound or the flash). The flash was a white circle. The sound was a burst of white noise presented over loudspeakers. Stimulus duration was 40 ms and response time was limited to 1500 ms.

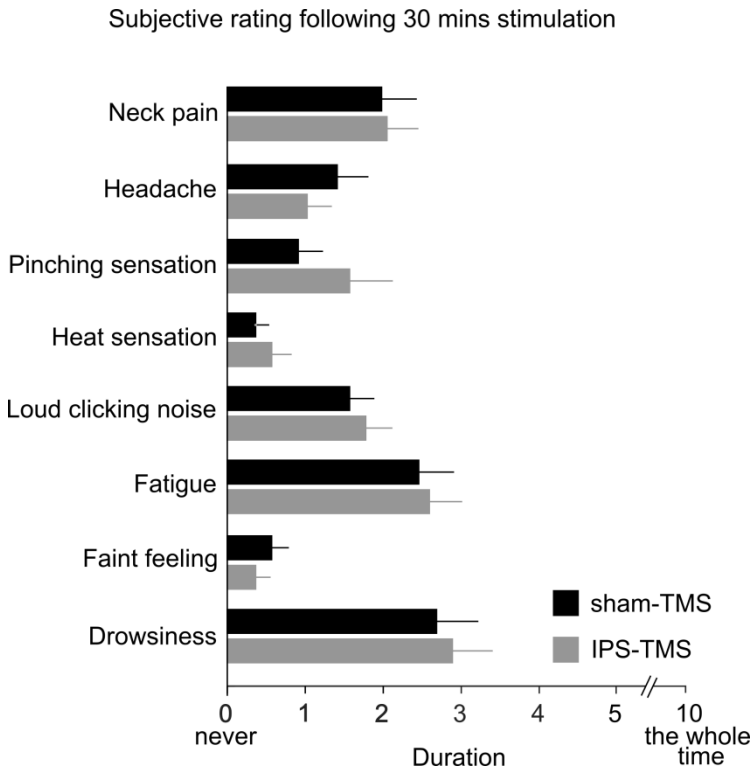


Figure 6.2: Results of the stimulation questionnaire

To measure potential differences between subjective effects of sham-TMS and IPS-TMS, after each session participants were asked to rate how often they experienced the sensations listed in the graph. No significant difference was found for ratings after sham-TMS and IPS-TMS in any of the pairs ($p > 0.05$).

6.2.5 Data analysis

Our analysis aimed to test whether IPS-TMS altered (i) the reliability of unisensory stimuli, (ii) the magnitude of the VE. As a follow-up analysis we also tested the effects of modality-specific attention. Sessions for each TMS condition were pooled over to obtain one session for each TMS condition. For completeness, all the following analyses were also run only for the first half of the task duration to account for the possibility that the TMS after-effects lasted less than 100% of the stimulation duration. In this further analysis conclusions regarding the effect of IPS-TMS did not change compared to the full analysis.

6.2.5.1 Effects of IPS-TMS on reliability of unisensory stimuli

Reliability of unisensory auditory and visual stimuli was computed for each participant as the standard deviation (SD) of responses at each spatial location. To evaluate whether IPS-TMS altered the reliability of auditory or visual unisensory stimuli, we entered the SD values into a 2 (*TMS*: sham-TMS vs IPS-TMS) x 2 (*task-relevance*: report auditory location vs report visual location) x 4 (spatial location: -32°, -24°, -16°, -8°) repeated measures analysis of variance (rmANOVA). We repeated the simple contrasts using non-parametric Wilcoxon Signed Ranks tests (2-tailed, correcting for multiple comparisons) to account for potential violation of normality assumptions.

6.2.5.2 Effects of IPS-TMS on the ventriloquist effect

To evaluate the effect of IPS-TMS on the VE we ran a model-free and a model-based analysis. In the model-free analysis, for each participant, for each TMS condition, for each response modality and for each audiovisual disparity (i.e., 8°, 16°

and 24°), we computed the relative visual and auditory bias for the reported stimuli (VE_{Aresp}, VE_{Aresp}) as:

(1) $VE_{Aresp} = (A_{resp} - A_{loc})/(V_{Loc} - A_{Loc})$

(2) $VE_{Vresp} = (V_{resp} - V_{loc})/(A_{Loc} - V_{Loc})$

With A = auditory modality, V = visual modality, A_{resp} = mean auditory localization response, V_{resp} = mean visual localization response, A_{Loc} = location of auditory stimulus, V_{Loc} = location of visual stimulus. We entered VE_{Aresp} and VE_{Vresp} values in a 2 (*TMS*: sham-TMS vs IPS-TMS) x 2 (*task-relevance*: report auditory vs visual location) x 3 (*spatial disparity*: 8°, 16° and 24°) rmANOVA. We repeated the simple contrasts using non-parametric Wilcoxon Signed Ranks tests (2-tailed, correcting for multiple comparisons) to account for potential violation of normality assumptions.

In the model-based analysis, we estimated the influence of the visual and auditory signals separately for the reported auditory and visual estimates by means of two linear regression models. Each regression model included the reported location as dependent variable (Y_A or Y_V) and eight regressors: 2 (*TMS*: sham-TMS or IPS-TMS) x 2 (*spatial disparity*: 0°/8° as small or 16°/24° as large) x 2 (*modality of signal locations*: auditory or visual).

For each TMS and disparity condition, parameter estimates of auditory and visual location (β_A and β_V) were used to compute a summary index for the relative influence of the visual location on perceived sound location (i.e., relative visual weight, RVW) and for the relative influence of the auditory location on perceived visual location (i.e., relative auditory weight, RAW):

(3) RVW = $\beta_V / (\beta_V + \beta_A)$

(4) RAW = $\beta_A / (\beta_A + \beta_V)$

Positive RVW and RAW indicates that perceived auditory or visual locations were shifted towards the location of the flash or the sound, as measured by the VE. Conversely, negative RVW and RAW indicate a repulsion effect, i.e. a shift in the opposite direction. To test for different RVW and RAW in relation to TMS application and audiovisual disparity, we entered these symmary indices into two 2 (*TMS*: sham-TMS vs IPS-TMS) x 2 (*spatial disparity*: small vs large) rmANOVA. We repeated the simple contrasts using non-parametric Wilcoxon Signed Ranks tests (2-tailed, correcting for multiple comparisons) to account for potential violation of normality assumptions.

6.2.5.3 Effects of modality-specific attention

As a follow-up analysis, we verified whether in the congruent condition participants were able to focus their attention selectively on the sensory modality to be reported. If they were in fact responding to the indicated sensory modality, we predicted that reliabilities of auditory and visual reports in congruent conditions should be different (under the assumption that they are different in the unisensory condition, Rohe and Noppeney, 2018). For completeness, we investigated if such modality-specific effect of attention was altered by IPS-TMS. Thus, we entered the SD of responses for congruent trials into a 2 (*TMS*: sham-TMS vs IPS-TMS) x 2 (*task-relevance*: report auditory vs report visual location) x 4 (*spatial location*: -32°, -24°, -16°, -8°) rmANOVA.

Moreover, to further investigate whether the RVW was exclusively determined by sensory reliability of visual stimuli (i.e., RVW is the same regardless of which stimulus modality has to be reported) or whether top-down processes such as modality-specific attention may contribute to the magnitude of visual weight (i.e., RVW is larger when vision is task-relevant), we included all reported auditory and visual estimates as dependent variable (Y) in one single linear regression model with two regressors: 2 (*task-relevance*: report auditory vs report visual location) x 2 (*stimulus modality at a given location*: auditory vs visual). As participants were instructed to attend and report auditory or visual signals in different audiovisual blocks, our design allowed us to test for different RVW in relation to task-relevance. Thus, the summary index RVW was calculated separately when audition was task-relevant (RVW_A) and when vision was task-relevant (RVW_V):

(5) $RVW_A = \beta_V / (\beta_V + \beta_A)$

(6) $RVW_V = \beta_V / (\beta_V + \beta_A)$

where β_A and β_V are the parameter estimates of auditory and visual locations. RVW_A and RVW_V were compared using a non-parametric Wilcoxon Signed Ranks test (2-tailed).

6.2.5.4 Control analysis: eye-movements

Participants' eye-movements were recorded for the 30-minutes duration of the task for each session. Fixation, saccades and blinks were automatically detected by the EyeLink eyetracking system. Fixation was well maintained within a radius of 1.3° (threshold based on Blignaut et al., 2009) for 21 participants (mean ± SD number of left saccades across participants and sessions for the whole duration of the task: 9 ±

13.7). Only for one participant a high number of left saccades was detected (mean across sessions: 456).

6.3 RESULTS

6.3.1 Effects of IPS-TMS on reliability of unisensory stimuli

For each block manipulating task-relevance (i.e., report auditory or visual location) audiovisual targets were presented interleaved with unisensory targets in the sensory modality defined by task-relevance. For unisensory targets, the 2 (*TMS*: sham-TMS vs IPS-TMS) x 2 (*task-relevance*: report auditory vs report visual location) x 4 (*spatial location*: -32°, -24°, -16°, -8°) rmANOVA with SD of response as dependent variable revealed no significant main effect of TMS, no significant interaction between TMS and task-relevance or spatial location and no significant three-way interaction ($p > 0.05$), suggesting that IPS-TMS did not alter unisensory signal reliability. To further support this hypothesis, we ran a Bayesian paired-sample t-test (JASP Version 0.8.4.0, using the default Cauchy prior) to compare sham-TMS and IPS-TMS effects on signal reliability, pooling over task-relevance and spatial location. These data were 3.04 more likely to occur under the null hypothesis (i.e., sham-TMS effects = IPS-TMS effects) (Bayes Factor in favour of the null hypothesis, BF01 = 3.04, i.e., moderate evidence, Jeffreys, 1961) relative to the alternative hypothesis (i.e., sham-TMS effects ≠ IPS-TMS effects) black vs grey lines in Fig. 6.3A, Table 6.1).

Despite the absence of IPS-TMS effects, the rmANOVA revealed a significant main effect of task-relevance ($F(1, 21) = 148.14$, $p < 0.001$, η_p^2 [90% CI] = 0.88 [0.76, 0.91]) and spatial location ($F(3, 63) = 23.75$, $p < 0.001$, η_p^2 [90% CI] = 0.53 [0.36, 0.62]) and a significant interaction between task-relevance and spatial location ($F(3, 63) = 14.95$, $p < 0.001$, η_p^2 [90% CI] = 0.42 [0.24, 0.52]). Thus, we further ran separate

rmANOVAs with spatial location as factor for auditory and visual modality separately, pooling over TMS conditions. While a significant main effect for spatial location was found for the auditory modality ($F(3, 63) = 33.28$, $p < 0.001$, η_p^2 [90% CI] = 0.61 [0.46, 0.68]), this effect was not significant for the visual modality ($p > 0.05$). Post-hoc tests (Bonferroni-corrected for 6 multiple comparisons) revealed that the two middle sound locations were reported with significantly less reliability than the two outer locations (-32 vs -24: $p < 0.001$; -32 vs -16: $p < 0.001$; -32 vs -8: $p > 0.05$; -24 vs -16: $p > 0.05$; -24 vs -8: $p = 0.001$; -16 vs -8: $p < 0.001$; Wilcoxon Signed Ranks tests: -32 vs -24: $p < 0.001$; -32 vs -16: $p < 0.001$; -32 vs -8: $p > 0.05$; -24 vs -16: $p > 0.05$; -24 vs -8: $p = 0.001$; -16 vs -8: $p < 0.001$). Overall, these results suggest that visual signals were localized with higher reliability than auditory signals (solid vs dashed lines in Fig. 6.3A, Table 6.1) and that auditory signal report was more reliable when the sound was presented at -32° and -8° locations. It should be noted that this latter effect may not be purely perceptual but merely be a consequence of the response distribution of outer signals, as more extreme responses were not made available to participants.

6.3.2 Effects of IPS-TMS on the ventriloquist effect

Participants were presented with audiovisual signals randomly sampled from four spatial locations in their left hemifield. Task-relevance was manipulated in blocks by instructing participants to either report the location of the auditory signal or the location of the visual signal. For the audiovisual targets, the 2 (*TMS*: sham-TMS vs IPS-TMS) x 2 (*task-relevance*: report auditory vs visual location) x 3 (*spatial disparity*: 8°, 16° and 24°) rmANOVA with VE as dependent variable revealed no significant main effect of TMS, no significant interaction between TMS and task-relevance or spatial disparity and no significant three-way interaction ($p > 0.05$). We

ran a default Bayesian paired-sample t-test to compare sham-TMS and IPS-TMS effects on VE, pooling over task-relevance and spatial location. Data were 4.35 more likely (BF01 = 4.35, i.e., moderate evidence) to occur under the null hypothesis (i.e., sham-TMS effects = IPS-TMS effects) relative to the alternative hypothesis (sham-TMS effects ≠ IPS-TMS effects), (black vs grey lines in Fig. 6.3B, Table 6.1).

Despite the absence of IPS-TMS effects, the rmANOVA revealed a main effect of task-relevance ($F(1, 21) = 40.40$, $p < 0.001$, η_p^2 [90% CI] = 0.66 [0.41, 0.76]) and disparity ($F(2, 42) = 79.74$, $p < 0.001$, η_p^2 [90% CI] = 0.79 [0.68, 0.84]) and a significant interaction between task-relevance and disparity ($F(2, 42) = 42.03$, $p < 0.001$, η_p^2 [90% CI] = 0.67 [0.50, 0.74]). Thus, we ran separate rmANOVAs with spatial disparity as factor for task-relevance conditions separately, pooling over TMS conditions. For both task-relevance conditions a main effect of disparity was found (auditory report: $F(2, 42) = 68.7$, $p < 0.001$, η_p^2 [90% CI] = 0.77 [0.64, 0.82]; visual report: $F(2, 42) = 32.07$, $p < 0.001$, η_p^2 [90% CI] = 0.60 [0.42, 0.69]). Post-hoc tests (pooling across TMS conditions, Bonferroni-corrected for 3 multiple comparisons) revealed that, for the 'report auditory' condition, when stimuli were presented at 8° and 16° of disparity, the VE was significantly larger than when they were presented at 24° disparity (8° vs 16°: $p > 0.05$; 8° vs 24°: $p < 0.001$; 16° vs 24°: $p < 0.001$; Wilcoxon Signed Ranks tests: 8° vs 16°: $p > 0.05$; 8° vs 24°: $p < 0.001$; 16° vs 24°: $p < 0.001$) (solid lines in Fig. 6.3B, Table 6.1). Conversely, for the 'report visual' condition, the VE was maximum at 8° but was significantly reduced already at 16° (8° vs 16°: $p < 0.001$; 16° vs 24°: $p > 0.05$; 8° vs 24°: $p > 0.05$; Wilcoxon Signed Ranks tests: 8° vs 16°: $p < 0.001$; 16° vs 24°: $p > 0.05$; 8° vs 24°: $p > 0.05$) (dashed lines in Fig. 6.3B, Table 6.1).

Overall, these results suggest that the VE was higher when participants had to report the location of the sound than the flash despite the audiovisual signal being the same (Fig. 6.3B, Table 6.1). Moreover, they show that the VE effect increased with reduced audiovisual disparity and that already at 16° the VE was significantly reduced for the visual report.

Results of our model-based analysis confirm these conclusions. No significant effect of TMS was found ($p > 0.05$; $BF01_{RVW} = 4.59$; $BF01_{RAW} = 3.53$) and RVW and RAW were larger for smaller audiovisual disparities (main effect of disparity: RVW: $F(1, 21) = 5.52$, $p = 0.029$, η_p^2 [90% CI] = 0.21[0.01, 0.42], Wilcoxon Signed Ranks tests pooling across TMS conditions: $p = 0.049$, solid lines in Fig. 6.3C, Table 6.1; RAW: $F(1, 21) = 18.99$, $p < 0.001$, η_p^2 [90% CI] = 0.47 [0.19, 0.63], Wilcoxon Signed Ranks tests pooling across TMS conditions: $p < 0.001$, dashed lines in Fig. 6.3C, Table 6.1).

6.3.3 Effects of modality-specific attention

Results of the rmANOVA on congruent signals in different task-relevance conditions indicated that modality-specific attention had an effect on reliability of congruent estimates, suggesting that during multisensory stimulation participants focused their attention on the modality to be reported (i.e., significant main effect of task-relevance: $F(1, 21) = 94.67$, $p < 0.001$, η_p^2 [90% CI] = 0.82 [0.66, 0.87]). This effect was not altered by IPS-TMS (no significant main effect of TMS or interaction between TMS and task-relevance, $p > 0.05$, BF01 = 5.93 for main effect of TMS for the difference between task-relevance conditions). Furthermore, a significant increase of visual weight when flashes were task relevant was observed ($RVW_V > RVW_A$, two-tailed Wilcoxon Signed Ranks tests: $p < 0.001$).

Table 6.1: Mean reliability (SD), Ventriloquist Effect and Visual and Auditory Weight for each TMS condition, for each location and for each task-relevance condition. Standard errors (SEM) are given in parentheses.

		Degrees	Report auditory	Report visual
Standard deviation (SEM)	Sham-TMS	-32	3.20 (0.33)	2.43 (0.28)
		-24	4.85 (0.28)	2.37 (0.19)
		-16	5.10 (0.26)	2.54 (0.31)
		-8	3.87 (0.21)	1.96 (0.26)
	IPS-TMS	-32	2.94 (0.29)	2.44 (0.25)
		-24	4.94 (0.22)	2.51 (0.18)
		-16	5.08 (0.22)	2.38 (0.28)
		-8	3.49 (0.22)	1.95 (0.21)
Ventriloquist Effect (SEM)	Sham-TMS	8	0.233 (0.031)	0.074 (0.01)
		16	0.224 (0.023)	0.042 (0.006)
		24	0.133 (0.023)	0.036 (0.006)
	IPS-TMS	8	0.237 (0.031)	0.063 (0.009)
		16	0.224 (0.026)	0.041 (0.006)
		24	0.134 (0.023)	0.037 (0.006)
Visual and Auditory Weight (SEM)	Sham-TMS	0-8	0.113 (0.038)	0.009 (0.005)
		16-24	0.090 (0.027)	-0.010 (0.002)
	IPS-TMS	0-8	0.120 (0.037)	0.001 (0.006)
		16-24	0.083 (0.028)	-0.008 (0.004)

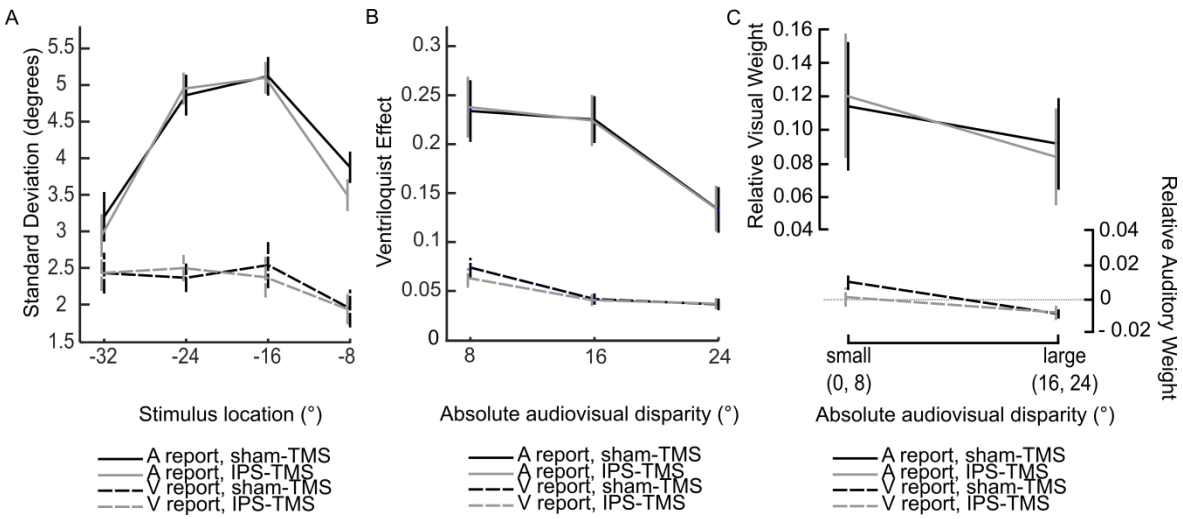


Figure 6.3: Behavioural results: sensory reliability, ventriloquist effect, relative visual and auditory weights.

A. Standard deviation as a measure of reliability for unisensory auditory (solid lines) and visual (dashed lines) report, for the four spatial locations. Auditory signals were localized in space significantly less reliably than visual signals ($p < 0.001$). No significant difference was found between sham-TMS (black line) and IPS-TMS (grey line) conditions ($p > 0.05$). **B.** The VE was computed for the incongruent trials as a function of task-relevance (report auditory location: solid lines, report visual location: dashed lines). The VE was larger for auditory than visual stimuli ($p < 0.001$). For both sensory modalities, the VE significantly increased with audiovisual disparity decreasing ($p < 0.001$). No significant difference was found between sham-TMS (black line) and IPS-TMS (grey line) conditions ($p > 0.05$). **C.** Relative visual (solid lines) and auditory weight (dashed lines) was computed pooling over congruent and incongruent trials as a function of audiovisual disparity (small [≤ 8°] vs large [> 8°]). Visual and auditory weight significantly increased with audiovisual disparity decreasing ($p < 0.05$). No significant difference was found between sham-TMS (black line) and IPS-TMS (grey line) conditions ($p > 0.05$).

6.4 DISCUSSION

The posterior parietal cortex is involved in numerous functions such as visual attention and multisensory integration (Kastner et al., 1999; Ghazanfar & Schroeder, 2006). In particular, previous neuroimaging studies have implicated IPS in attentional control and spatial reorienting (Corbetta et al., 2008) regardless of sensory modality (Macaluso et al., 2002) and have defined it as 'association cortex', i.e. a brain region where multiple sensory modalities converge (Calvert and Thesen, 2004). Functional activations observed in association cortices may derive from areas that are common to different sensory modalities or from coactivation of functionally distinct unisensory areas which are close in space (Klemen & Chambers, 2012).

Importantly, using a VE paradigm, previous functional imaging research has suggested that IPS differently encodes the spatial estimates of sensory signals at different stages: while in IPS1-2 signals are integrated weighted by their sensory reliabilities, in IPS3-4 the full-segregation and the forced-fusion models are combined (Rohe & Noppeney, 2015a). The VE paradigm provides direct information about how signals are integrated or segregated in relation to signal reliability. In particular, the magnitude of visual influence on auditory percepts (i.e., visual weight) increases with increased sensory reliability (Rohe & Noppeney, 2018). Furthermore, previous research has shown that multisensory processes depend on attention orienting, suggesting that also top-down mechanisms such as task-relevance play an important role in shaping multisensory percepts (Talsma & Woldorff, 2005; Vercillo & Gori, 2015).

Using a VE paradigm, the current study aimed to investigate whether changes in terms of sensory reliability and audiovisual binding can be observed following a

decrease of cortical excitability by means of right anterior IPS stimulation. Participants were instructed to report the location of the auditory or visual signal in (1) unisensory, (2) spatially congruent or (3) spatially incongruent audiovisual trials.

Our specific reliability and VE analyses produced three relevant results. First, we found that the reliability of unisensory signals was not altered by IPS-TMS. Second, we found that IPS-TMS did not alter the VE in the model-free analysis or the relative visual or auditory weight in the model-based analysis, indicating that audiovisual binding processes measured with our task did not significantly change following application of IPS-TMS. Third, in our follow-up analysis we found that also modality-specific attentional processes were not altered by IPS-TMS.

Importantly, the absence of TMS effects that we observed cannot be explained by the lack of behavioural effects per se, as we were able to successfully replicate previous behavioural findings: (1) the magnitude of the VE depended on audiovisual disparity (Wallace et al., 2004); (2) vision exerted a stronger influence on audition than audition on vision (Odegaard et al., 2015), which reflected on a greater VE for audition than vision and on a larger relative visual weight than relative auditory weight (Alais et al., 2010); (3) modality-specific attention increased sensory weight (Rohe & Noppeney, 2018), possibly either by increasing reliability of sensory signals (Martinez-Trujillo & Treue, 2004) or by actively biasing multisensory integration processes in favour of the attended signal (Talsma & Woldoroff, 2005 but see Bertelson, et al., 2000). Therefore, collectively we showed that sensory reliability is different in audition and vision and that audiovisual binding can be measured in function of audiovisual disparity, sensory reliability and top-down attentional processes.

Our study is not the first that failed to find effects of IPS-TMS on multisensory processes. Bertini et al. (2010) applied rTMS to the right posterior parietal cortex (rPPC), right temporoparietal cortex (rTPC) and right occipital cortex (rOC) and evaluated its effects on spatially congruent and incongruent auditory and visual stimuli. Inhibition of rTPC reduced multisensory enhancement of congruent trials, whereas inhibition of rOC reduced the VE of incongruent trials. However, no significant effect was found for rPPC. They hypothesised that their null effect would be explained by postulating that temporoparietal regions are responsible for audiovisual integration and posterior parietal activations would only reflect an epiphenomenon. Their conclusions are consistent with previous neurostimulation and neuropsychological studies showing that the parietal lobe is not crucially involved in audiovisual integration (Bolognini et al., 2009; Passamonti et al., 2009). Moreover, in agreement with this hypothesis, further neurostimulation studies brought evidence that temporoparietal cortex stimulation indeed altered multisensory processes (Chambers et al., 2007; Kamke et al., 2012). However, in contrast with this conclusion, other studies targeting the right IPS were able to reveal changes in multisensory processes caused by TMS application (Pasalar et al., 2010; Bien et al., 2012). Given the numerous differences in terms of stimulation site, TMS protocol and experimental task employed by those previous studies, striking conclusions about the causal involvement of posterior parietal cortex in multisensory processes cannot be drawn.

In our study, the stimulation site was selected based on retinotopy maps of a brain region (IPS3-4) which was found to be relevant for multisensory integration measured with fMRI in a similar spatial ventriloquism study (Rohe & Noppeney, 2015a). Further, we performed MRI-guided neuronavigation to accurately localise individual spots and optimise the effect of stimulation (Sack et al., 2008). For these

reasons, we are inclined to exclude the possibility that the cortical region that we targeted in each participant was inadequate for the behavioural outcome that we tested. We showed that our control condition was effective, that only few participants noticed the difference between sham and real stimulation and that the stimulation discomforts reported after real stimulation were not significantly different than those reported after sham stimulation. Moreover, we employed an offline approach to avoid that TMS-specific effects, such as lateralized clicking noise, could interfere with sensory task performance (Duecker & Sack, 2015). Hence, we can safely exclude that any bias could have been introduced by the application of sham-TMS vs IPS-TMS and by related subjective discomforts.

Previous methodological studies showed that the after-effect of low-frequency rTMS lasts for up to 50-100% the duration of the stimulation (Robertson et al., 2003 and Hoogendam et al., 2010). Importantly, only few studies investigated the after-effect duration in brain areas other than the motor cortex (Eisenegger et al., 2008). To evaluate whether stimulation after-effects could have lasted less than stimulation time, we repeated our analyses including only the first 15 minutes of the task, but we did not find significant effects related to TMS application. Critically, we cannot exclude the possibility that longer-lasting stimulation after-effects were completely absent or that the strength of the magnetic field was too weak in the relevant portion of IPS3-4 (the strength of the magnetic field rapidly weakens with depth, Siebner et al., 2009), as we did not make any physiological measurement which could challenge such conclusion.

The magnitude of the VE that we observed deserves some consideration. Despite the visual signal being very reliable (i.e., a white disc on a black background) the mean VE across subjects at the smallest disparity was 0.24. A previous similar study reported higher mean VE across subjects, ranging between 0.5 to 0.9 for auditory report of

audiovisual stimuli with high visual reliability (Rohe & Noppeney, 2015b). Hence, we cannot rule out potential floor effects that could have prevented detectable behavioural changes following IPS-TMS. Moreover, the VE could have potentially been reduced following the unisensory training that our participants underwent before performing the experimental task. The advantage of improving localisation accuracy in the unisensory condition may have come at price of a lower audiovisual binding.

Importantly, our analysis did not simply fail to provide evidence in favour of IPS-TMS altering signal reliability or the VE. By means of Bayesian analyses we were able to provide moderate evidence in favour of the absence of a difference between sham-TMS and IPS-TMS. Nevertheless, in light of recent neuroimaging results, it seems implausible to conclude that IPS3-4 would not be involved in audiovisual binding as more substantial proof is needed to draw such conclusion. Our contribution instead emphasises the complexity of multisensory integration and segregation processes in relation to modality-specific attention and the necessity for more fine-grained investigation on the multifaceted role of IPS in such processes. Previous research combined multisensory paradigms with common-source judgment tasks where participants were asked to report whether the two sensory signals were produced by a common source. Different results were found for signals perceived from same or different sources and localization biases emerged for trials which were reported as to be emanated from a common source (Wallace et al., 2004; Odegaard et al., 2015). Accordingly, reducing the excitability of IPS3-4 may lead to independent behavioural effects in terms of (i) audiovisual bias and (ii) common-source judgments or (iii) their interaction (Wallace et al., 2004; Rohe & Noppeney, 2016). Thus, further studies may compare behavioural and neurophysiological effects of excitation/inhibition of sensory cortices, IPS1-4 and more decision-related areas (e.g., the dorsolateral

prefrontal cortex) during localization and common-source judgment tasks, to disentangle outcomes predicted by full segregation, full integration and Bayesian Causal Inference models.

REFERENCES

Alais, D., & Burr, D. (2004). Ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, *14*(3), 257–262. http://doi.org/10.1016/S0960-9822(04)00043-0

Alais, D., Newell, F. N., & Mamassian, P. (2010). Multisensory processing in review: From physiology to behaviour. *Seeing and Perceiving*, *23*(1), 3–38. http://doi.org/10.1163/187847510X488603

Bertelson, P., & Ascherleben, G. (1998). Automatic visual bias of perceived auditory location PAUL. *Psychonomic Bulletin & Review*, *5*(3), 482–489.

Bertelson, P., Pavani, F., Ladavas, E., Vroomen, J., & De Gelder, B. (2000). Ventriloquism in patients with unilateral visual neglect. *Neuropsychologia*, *38*(12), 1634–1642. http://doi.org/10.1016/S0028-3932(00)00067-1

Bertelson, P., & Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Perception & Psychophysics*, *29*(6), 578–584. http://doi.org/10.3758/BF03207374

Bertini, C., Leo, F., Avenanti, A., & Làdavas, E. (2010). Independent mechanisms for ventriloquism and multisensory integration as revealed by theta-burst stimulation. *European Journal of Neuroscience*, *31*(10), 1791–1799. http://doi.org/10.1111/j.1460-9568.2010.07200.x

Bien, N., ten Oever, S., Goebel, R., & Sack, A. T. (2012). The sound of size. Crossmodal binding in pitch-size synesthesia: A combined TMS, EEG and psychophysics study. *NeuroImage*, *59*(1), 663–672. http://doi.org/10.1016/j.neuroimage.2011.06.095

Blignaut, P. (2009). Fixation identification: the optimum threshold for a dispersion algorithm. *Attention, Perception & Psychophysics*, *71*(4), 881–895.

Bolognini, N., Leo, F., Passamonti, C., Stein, B. E., & Làdavas, E. (2007). Multisensory-mediated auditory localization. *Perception*, *36*(10), 1477–1485. http://doi.org/10.1068/p5846

Bolognini, N., & Maravita, A. (2007). Proprioceptive alignment of visual and somatosensory maps in the posterior parietal cortex. *Current Biology*, *17*(21), 1890–1895. http://doi.org/10.1016/j.cub.2007.09.057

Bolognini, N., Miniussi, C., Savazzi, S., Bricolo, E., & Maravita, A. (2009). TMS modulation of visual and auditory processing in the posterior parietal cortex. *Experimental Brain Research*, *195*(4), 509–517. http://doi.org/10.1007/s00221-009-1820-7

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. http://doi.org/10.1163/156856897X00357

Bremmer, F., Schlack, A., Shah, N. J., Zafiris, O., Kubischik, M., Hoffmann, K.-P., … Fink, G. R. (2001). Polymodal motion processing in posterior parietal and premotor cortex. *Neuron*, *29*, 287–296. http://doi.org/10.1016/S0896-6273(01)00198-2

Calvert, G. (2001). Crossmodal processing in the human brain: insights from functional neuroimaging studies. *Cerebral Cortex*, *11*, 1110–23. http://doi.org/10.1093/cercor/11.12.1110

Calvert, G. A., & Thesen, T. (2004). Multisensory integration: methodological approaches and emerging principles in the human brain. *Journal of Physiology-Paris*, *98*(1–3), 191–205. http://doi.org/10.1016/j.jphysparis.2004.03.018

Chambers, C. D., Payne, J. M., & Mattingley, J. B. (2007). Parietal disruption impairs reflexive spatial attention within and between sensory modalities. *Neuropsychologia*, *45*(8), 1715–1724. http://doi.org/10.1016/j.neuropsychologia.2007.01.001

Charbonneau, G., Veronneau, M., Boudrias-Fournier, C., Lepore, F., & Collignon, O. (2013). The ventriloquist in periphery: Impact of eccentricity-related reliability on audio-visual localization. *Journal of Vision*, *13*(12), 20–20. http://doi.org/10.1167/13.12.20

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron*, *58*(3), 306–324. http://doi.org/10.1016/j.neuron.2008.04.017

Duecker, F., & Sack, A. T. (2015). Rethinking the role of sham TMS. *Frontiers in Psychology*, *6*, 1–5. http://doi.org/10.3389/fpsyg.2015.00210

Eisenegger, C., Treyer, V., Fehr, E., & Knoch, D. (2008). Time-course of "off-line" prefrontal rTMS effects - a PET study. *NeuroImage*, *42*(1), 379–384. http://doi.org/10.1016/j.neuroimage.2008.04.172

Frassinetti, F., Bolognini, N., & Làdavas, E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Experimental Brain Research*, *147*(3), 332–343. http://doi.org/10.1007/s00221-002-1262-y

Ghazanfar, A., & Schroeder, C. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, *10*(6), 278–285. http://doi.org/10.1016/j.tics.2006.04.008

Gielen, S. C. A. M., Schmidt, R. A., & Van Den Heuvel, P. J. M. (1983). On the nature of intersensory facilitation of reaction time. *Perception &Psychophysics*, *34*(2), 161–168. http://doi.org/10.1016/0001-6918(84)90040-4

Hilgetag, C. C., Théoret, H., & Pascual-leone, A. (2001). Enhanced visual spatial attention ipsilateral to rTMS-induced "virtual lesions" of human parietal cortex. *Nature Neuroscience*, *4*(9), 953–957. http://doi.org/10.1038/nn0901-953

Hoogendam, J. M., Ramakers, G. M. J., & Di Lazzaro, V. (2010). Physiology of repetitive transcranial magnetic stimulation of the human brain. *Brain Stimulation*, *3*(2), 95–118. http://doi.org/10.1016/j.brs.2009.10.005

Jeffreys, H. (1961). *Theory of Probability*, 3rd ed. Oxford: Oxford University Press.

Kamke, M. R., Vieth, H. E., Cottrell, D., & Mattingley, J. B. (2012). Parietal disruption alters audiovisual binding in the sound-induced flash illusion. *NeuroImage*, *62*(3), 1334–1341. http://doi.org/10.1016/j.neuroimage.2012.05.063

Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron*, *22*, 751–761.

Kim, Y. H., Min, S. J., Ko, M. H., Park, J. W., Sung, H. J., & Lee, P. K. W. (2005). Facilitating visuospatial attention for the contralateral hemifield by repetitive TMS on the posterior parietal cortex. *Neuroscience Letters*, *382*(3), 280–285. http://doi.org/10.1016/j.neulet.2005.03.043

Klemen, J., & Chambers, C. D. (2012). Current perspectives and methods in studying neural mechanisms of multisensory interactions. *Neuroscience and Biobehavioral Reviews*, *36*(1), 111–133. http://doi.org/10.1016/j.neubiorev.2011.04.015

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal inference in multisensory perception. *PLoS ONE*, *2*(9). http://doi.org/10.1371/journal.pone.0000943

Lisanby, S. H., Gutman, D., Luber, B., Schroeder, C., & Sackeim, H. A. (2001). Sham TMS: Intracerebral measurement of the induced electrical field and the induction of motor-evoked potentials. *Biological Psychiatry*, *49*(5), 460–463. http://doi.org/10.1016/S0006-3223(00)01110-0

Macaluso, E., Frith, C. D., & Driver, J. (2002). Supramodal effects of covert spatial orienting triggered by visual or tactile events. *Journal of Cognitive Neuroscience*, *14*(3), 389–401. http://doi.org/10.1162/089892902317361912

Maeda, F., Keenan, J. P., Tormos, J. M., Topka, H., & Pascual-Leone, A. (2000). Interindividual variability of the modulatory effects of repetitive transcranial magnetic stimulation on cortical excitability. *Experimental Brain Research*, *133*(4), 425–430. http://doi.org/10.1007/s002210000432

Martinez-Trujillo, J. C., & Treue, S. (2004). Feature-based attention increases the selectivity of population responses in primate visual cortex. *Current Biology*, *14*, 744–751. http://doi.org/10.1016/j

Meredith, M. A., & Stein, B. E. (1983). Interactions among converging sensory inputs in the superior colliculus, *221*, 389–391.

Odegaard, B., Wozny, D. R., & Shams, L. (2015). Biases in visual, auditory, and audiovisual perception of space. *PLoS Computational Biology*, *11*(12), 1–23. http://doi.org/10.1371/journal.pcbi.1004649

Oldfield R.C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*, 97–113.

Oliver, R., Bjoertomt, O., Driver, J., Greenwood, R., & Rothwell, J. (2009). Novel “hunting” method using transcranial magnetic stimulation over parietal cortex disrupts visuospatial sensitivity in relation to motor thresholds.

Neuropsychologia, *47*(14), 3152–61. http://doi.org/10.1016/j.neuropsychologia.2009.07.017

Oliveri, M., Caltagirone, C., Filippi, M. M., Traversa, R., Cicinelli, P., Pasqualetti, P., & Rossini, P. M. (2000). Paired transcranial magnetic stimulation protocols reveal a pattern of inhibition and facilitation in the human parietal cortex. *Journal of Physiology*, *529*(2), 461–468. http://doi.org/10.1111/j.1469-7793.2000.00461.x

Pasalar, S., Ro, T., & Beauchamp, M. S. (2010). TMS of posterior parietal cortex disrupts visual tactile multisensory integration. *European Journal of Neuroscience*, *31*(10), 1783–1790. http://doi.org/10.1111/j.1460-9568.2010.07193.x

Passamonti, C., Frissen, I., & Làdavas, E. (2009). Visual recalibration of auditory spatial perception: Two separate neural circuits for perceptual learning. *European Journal of Neuroscience*, *30*(6), 1141–1150. http://doi.org/10.1111/j.1460-9568.2009.06910.x

Pell, G. S., Roth, Y., & Zangen, A. (2011). Modulation of cortical excitability induced by repetitive transcranial magnetic stimulation: Influence of timing and geometrical parameters and underlying mechanisms. *Progress in Neurobiology*, *93*(1), 59–98. http://doi.org/10.1016/j.pneurobio.2010.10.003

Ro, T., Wallace, R., Hagedorn, J., Farné, A., & Pienkos, E. (2004). Visual enhancing of tactile perception in the posterior parietal cortex. *Journal of Cognitive Neuroscience*, *16*(1), 24–30. http://doi.org/10.1162/089892904322755520

Robertson, E. M., Théoret, H., & Pascual-Leone, A. (2003). Studies in cognition: the problems solved and created by transcranial magnetic stimulation. *Journal of Cognitive Neuroscience*, *15*(7), 948–960. http://doi.org/10.1162/089892903770007344

Rohe, T., & Noppeney, U. (2015a). Cortical hierarchies perform Bayesian Causal Inference in multisensory perception. *PLOS Biology*, *13*(2), e1002073. http://doi.org/10.1371/journal.pbio.1002073

Rohe, T., & Noppeney, U. (2015b). Sensory reliability shapes Bayesian Causal Inference in perception via two mechanisms. *Journal of Vision*, *15*, 1–38. http://doi.org/10.1167/15.5.22.doi

Rohe, T., & Noppeney, U. (2016). Distinct computational principles govern multisensory integration in primary sensory and association cortices. *Current Biology*, *26*(4), 509–514. http://doi.org/10.1016/j.cub.2015.12.056

Rohe, T., & Noppeney, U. (2018). Reliability-weighted integration of audiovisual signals can be modulated by top-down control. *Eneuro*, *5*, 1–20. http://doi.org/10.1523/ENEURO.0315-17.2018

Rossi, S., Hallett, M., Rossini, P. M., & Pascual-Leone, A. (2009). Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clinical Neurophysiology*, *120*(12), 2008–2039. http://doi.org/10.1016/j.clinph.2009.08.016

Sack, A. T., Kadosh, R. C., Schuhmann, T., Moerel, M., Walsh, V., & Goebel, R. (2008). Optimizing functional accuracy of TMS in cognitive studies: A comparison of methods. *Journal of Cognitive Neuroscience*, *21*(2), 207–221. http://doi.org/10.1162/jocn.2009.21126

Santangelo, V., Olivetti Belardinelli, M., Spence, C., & Macaluso, E. (2009). Interaction between voluntary and stimulus-driven spatial attention mechanism across sensory modalities. *Journal of Cognitive Neuroscience*, *21*(12), 2384–2397. http://doi.org/10.1162/jocn.2008.21178

Siebner, H. R., Hartwigsen, G., Kassuba, T., & Rothwell, J. (2009). How does transcranial magnetic stimulation modify neuronal activity in the brain ? - Implications for studies of cognition. *Cortex*, *45*(9), 1035–1042. http://doi.org/10.1016/j.cortex.2009.02.007.How

Smith, D. V, Davis, B., Niu, K., Healy, E. W., Bonilha, L., Fridriksson, J., … Rorden, C. (2010). Spatial attention evokes similar activation patterns for visual and auditory stimuli. *Journal of Cognitive Neuroscience*, *22*(2), 347–61. http://doi.org/10.1162/jocn.2009.21241

Swisher, J. D., Halko, M. A., Merabet, L. B., McMains, S. A., & Somers, D. C. (2007). Visual topography of human intraparietal sulcus. *Journal of Neuroscience*, *27*(20), 5326–5337. http://doi.org/10.1523/JNEUROSCI.0991-07.2007

Talsma, D., & Woldorff, M. G. (2005). Selective attention and multisensory integration: multiple phases of effects on the evoked brain activity. *Journal of Cognitive Neuroscience*, *17*(7), 1098–1114. http://doi.org/10.1162/0898929054475172

Teder-Sälejärvi, W. A., Di Russo, F., McDonald, J. J., & Hillyard, S. A. (2005). Effects of spatial congruity on audio-visual multimodal integration. *Journal of Cognitive Neuroscience*, *17*(9), 1396–1409. http://doi.org/10.1162/0898929054985383

Thut, G., Nietzel, A., & Pascual-Leone, A. (2005). Dorsal posterior parietal rTMS affects voluntary orienting of visuospatial attention. *Cerebral Cortex*, *15*(5), 628–638. http://doi.org/10.1093/cercor/bhh164

Vercillo, T., & Gori, M. (2015). Attention to sound improves auditory reliability in audio-tactile spatial optimal integration. *Frontiers in Integrative Neuroscience*, *9*, 1–8. http://doi.org/10.3389/fnint.2015.00034

Vroomen, J., & De Gelder, B. (2000). Sound enhances visual perception: Cross-modal effects of auditory organization on vision. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(5), 1583–1590. http://doi.org/10.1037/0096-1523.26.5.1583

Vroomen, J., & Gelder, B. De. (2004). Perceptual effects of cross-modal stimulation: Ventriloquism and the freezing phenomenon. *The Handbook of Multisensory …*, (1), 141–150.

Wagner, T., Valero-Cabre, A., & Pascual-Leone, A. (2007). Noninvasive human brain stimulation. *Annual Review of Biomedical Engineering*, *9*(1), 527–565. http://doi.org/10.1146/annurev.bioeng.9.061206.133100

Wallace, M. T., Roberson, G. E., Hairston, W. D., Stein, B. E., Vaughan, J. W., & Schirillo, J. A. (2004). Unifying multisensory signals across time and space. *Experimental Brain Research*, *158*(2), 252–258. http://doi.org/10.1007/s00221-004-1899-9

Wang, L., Mruczek, R. E. B., Arcaro, M. J., & Kastner, S. (2015). Probabilistic maps of visual topography in human cortex. *Cerebral Cortex*, *25*(10), 3911–3931. http://doi.org/10.1093/cercor/bhu277

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, *88*(3), 638–667. http://doi.org/10.1037/0033-2909.88.3.638

Zuanazzi, A., & Cattaneo, L. (2017). The right hemisphere is independent from the left hemisphere in allocating visuospatial attention. *Neuropsychologia*, *102*, 197–205. http://doi.org/10.1016/j.neuropsychologia.2017.06.005

Chapter 7:

General discussion

7.1 SPATIAL ATTENTION AND EXPECTATION IN A MULTISENSORY ENVIRONMENT

Our brain is constantly bombarded by an overwhelming amount of signals from different sensory modalities and needs to combine them into a coherent percept. Attention and expectation are two mechanisms that facilitate perception by respectively selecting relevant information and constraining the interpretation of such information on the basis of prior probability. Furthermore, perception is facilitated when signals in multiple sensory modalities that share spatial and temporal features are combined. The interplay between attention and expectation (Summerfield & Egner, 2009) and between these two mechanisms in relation to multisensory processes (Santangelo & Macaluso, 2012) is multifaceted. In particular, previous research has strived to address two main questions: (1) does attention (and expectation) engage multisensory processes and (2) are attention and expectation two sides of the same mechanism?.

Previous multisensory studies showed that behavioural and neural effects of spatial attention are synergistic across sensory modalities and engage frontoparietal and temporoparietal cortices (Spence & Driver, 1996; Macaluso et al., 2002). Furthermore, the posterior parietal cortex has recently been shown to play an important role in arbitrating between integration and segregation of multisensory signals (Rohe & Noppeney, 2015, 2016) However, the causal and critical involvement of the posterior parietal cortex in multisensory processes remains unclear, in light of previous neurostimulation studies reporting contrasting evidence (Bertini et al., 2010; Kamke et al., 2012).

Critically, previous multisensory research has often confounded attention with expectation, thus behavioural and neural effects of these two mechanisms could not be disentangled. Only recently have studies attempted to separate attention and expectation (mainly in the unisensory field and at the neural level), yet they showed contrasting results (i.e., additive or interactive effects of attention and expectation, Kok et al., 2012; Lange, 2013).

The main aim of this thesis was to contribute to the understanding of multisensory processes in relation to spatial attention and expectation. By implementing a novel audiovisual paradigm which relies on the multisensory generalization effects originally showed only for attention (Spence & Driver, 1996), chapter 3 aimed to evaluate whether spatial attention and expectation, when orthogonally manipulated, have additive or interactive effects at the behavioural level. Chapters 4 and 5 investigated (1) whether interactive effects of spatial attention and expectation observed at the behavioural level could be observed at the neural level and (2) whether spatial attention and expectation (or their interaction) engaged multisensory processes at the behavioural and neural level. Finally, chapter 6 examined the causal and critical role of right anterior IPS in modifying signal reliability and in multisensory integration and segregation.

7.2 FINDINGS

The operational definition of attention as 'task-relevance' and expectation as 'signal probability' sets up a dichotomy between these two processes. Contrasting behavioural and neural evidence has been provided in support of such dichotomy (Nobre & van Ede, 2017). Chapter 3 indicated that a rigid dichotomy between attention and expectation does not explain perceptual decision processes. In fact, seemingly

opposite behavioural effects of attention and expectation can be parsimoniously explained by general and spatially selective response probability. Importantly, response probability is determined by the probability that the signal is task-relevant, suggesting that attention and expectation are intimately linked during perceptual-decision processes. Consequently, chapter 3 shed new light on the central role of task-relevance in attention and expectation processes. Behavioural results of chapter 3 were further replicated in chapter 5.

Using the same multisensory approach, chapter 4 enriched the results of chapter 3 by providing more information about how attention and expectation processes are shared across sensory modalities. Importantly, while chapter 3 aimed to characterise the interplay between spatial attention and expectation by capitalising on crossmodal links of attention (and expectation) previously described in multisensory research (Spence & Driver, 1996), chapter 4 provided evidence in support of such links. In particular, chapter 4 demonstrated that response probability (i.e., the probability to respond in function of spatial attention and expectation) exhibits similar crossmodal qualitative patterns for vision and audition. This expands the conclusions of chapter 3 which were limited to a specific multisensory setup (i.e., crossmodal effects were evaluated only in vision). Furthermore, chapter 4 showed that crossmodal effects of response probability were quantitatively smaller than modality-specific effects, hence demonstrating that response probability is not fully supramodal but generalises crossmodally. Importantly, results of chapters 3 and 5 could not address the question of whether modality-specific and crossmodal effects of response probability differed because of intrinsic properties of attention and expectation mechanisms or because of sensory features of the signals. Findings of chapter 4 supported a combination of the former and the latter, by showing that response probability effects emerge slower over

time for audition than vision, possibly because of lower reliability of auditory signals compared to visual signals (Knudsen & Brainard, 1995; Jabar & Anderson, 2015; Rohe & Noppeney, 2015).

Crucially, chapters 3 and 4 drew conclusions about how attention and expectation processes interact during perceptual decisions. Hence, they investigated a stage in perceptual decision-making processes where decision-making and motor preparation play a decisive role (Summerfield & Egner, 2009, Summerfield & de Lange, 2014; Banerjee et al., 2017). Moreover, given our operational definition of attention as requirement to respond, in chapters 3 and 4 data for the task-irrelevant hemifields were not collected for the primary modality. fMRI allows tracking neural responses of brain areas involved in different stages of perceptual decision-making and for signals that do not require a response. Chapter 5 further complemented results of chapters 3 and 4 by investigating whether spatial attention and expectation engaged similar systems at the neural level and whether such systems were modality-specific or were shared across sensory modalities. First, chapter 5 showed that spatial attention and expectation indeed share neural systems. Second, with regard to whether such systems have multisensory properties, chapter 5 replicated previous results showing that attentional resources are interactively controlled by a frontoparietal system across sensory modalities (Macaluso et al., 2000, 2002; Banerjee et al., 2011). Furthermore, it demonstrated that spatial expectations are formed in primary and frontoparietal cortices in a modality-specific fashion, i.e., only for the modality where expectation is directly manipulated. This novel result suggests that neural correlates of spatial attention and expectation differ in their modality-specificity.

Finally, chapter 6 showed that vision exerts a stronger auditory capture and dominates audition (Alais et al., 2010). Moreover, it reiterated the pivotal role of

spatial proximity in multisensory integration (Wallace et al., 2004). Nevertheless, application of TMS over right anterior IPS did not alter unisensory reliability (possibly controlled by attentional top-down modulation, Martinez-Trujillo & Treue, 2004) or the magnitude of the ventriloquist effect. Considering the complexity of multisensory interactions in relation to top-down attention and in light of recent neuroimaging results that rigorously characterised the neural correlates of multisensory processes in IPS (Rohe & Noppeney, 2015, 2016), we concluded that further research is needed to verify whether the involvement of the posterior parietal cortex is critical for multisensory integration and segregation.

7.3 IMPLICATIONS, LIMITATIONS AND FUTURE DIRECTIONS

Past unisensory literature underlined the importance of dissociating attention and expectation behavioural and neural effects (Summerfield & Egner, 2009). The predictive coding framework postulates that attention and expectation function synergistically, with attention increasing the precision of prediction errors (Friston, 2009). Furthermore, past fMRI research showed that, while it is possible to set up an operational dichotomy of attention as 'task-relevance' and expectation as 'signal probability', the effects of these two mechanisms are not completely separate (Kok et al., 2012, Jiang et al., 2013). Our results contribute to this ongoing debate by showing that, at the level of perceptual-decisions, attention and expectation co-determine the probability to respond and completely separate behavioural effects cannot be pinpointed. Furthermore, we showed that a clear dichotomy in terms of neural correlates cannot be defined either, as attention and expectations engage overlapping neural systems, despite exhibiting independent neural activations. A limitation of our studies (chapters 3, 4 and 5) is that attention and task-relevance were not dissociated,

thus we could not entirely separate endogenous mechanisms of attention orienting from perceptual-decision making processes. It would certainly be possible to employ instructions alone to manipulate attention whilst holding the response requirements constant (although to our knowledge, only few studies already used this attention manipulation, e.g., Heinze et al., 1994; Yi et al., 2006). Yet, with this experimental design it would be challenging to evaluate how observers orient their attention in function of the instructions alone.

Similar conclusions can be drawn for the differences that we found between attention and expectation mechanisms with regard to multisensory processes. Previous research showed that a stimulus that is task-relevant can trigger attention reorienting to the unattended hemifield, as indicated by frontoparietal activations (Indovina & Macaluso, 2007). Our neural results (chapter 5) can be reconciled with this finding by postulating that in our task attention reorienting initially took place regardless of stimulus modality, as visual stimuli were always task-relevant (i.e. they required a response whether they were presented in the attended or unattended hemifields). Only at this stage perceptual decisions could be formed, with task-relevance (i.e., whether a response was required depending on the stimulus modality) and signal probability co-determining response probability. Consequently, we may conclude that while reorienting of attention was triggered by any stimulus presented in the attended hemifield as in both hemifields a proportion of stimuli always required a response, expectations were formed independently in each sensory modality to be combined only at decisional stages. This would suggest that in our experiments task-relevance influenced attention reorienting even before decisional stages. In conclusion, whether it is possible to investigate attention (and expectation) only at the perceptual level by

fully factoring out task and decision-making processes remains an open question (Nobre et al., 2018).

In chapters 4 and 5 we showed that spatial attention and expectation behavioural effects extended crossmodally and that attention engaged multisensory systems. However we also found that expectations were formed in a modality-specific fashion. Importantly, while at the behavioural level we were able to demonstrate that the way expectations and response probability are shaped depend on sensory properties of the signal, in our fMRI study we only tested for modality-specific effects of auditory expectations and the respective crossmodal effects in vision but not vice versa. Thus, in this specific study we could not determine whether modality-specific expectation effects crucially depended on sensory-specific properties of the auditory signal.

Overall, our behavioural and neural results emphasized the pivotal importance of sensory properties of the signal when investigating multisensory processes. As demonstrated by previous studies, visual information is space-based and continuously available, whereas auditory information is frequency-based, intermittent and depends on the occurrence of an event (Lauter et al., 1985; Neumann et al., 1986; Sereno et al., 1995). In chapter 6 we reported critical differences in terms of reliability and spatial properties between auditory and visual signals. Nevertheless, we could not support the causal involvement of IPS in modulating signal reliability or audiovisual binding in function of signal location. Future neuroimaging studies may provide further evidence about how IPS arbitrates between multisensory integration and segregation in function of sensory-specific features, taking into account the fundamental different functions of audition and vision in perception and action (Neumann et al., 1986).

REFERENCES

Alais, D., Newell, F. N., & Mamassian, P. (2010). Multisensory processing in review: From physiology to behaviour. *Seeing and Perceiving*, *23*(1), 3–38. http://doi.org/10.1163/187847510X488603

Banerjee, S., Grover, S., Ganesh, S., & Sridharan, D. (2017). Sensory and decisional components of endogenous attention are dissociable ., 1–48. http://doi.org/dx.doi.org/10.1101/231613

Banerjee, S., Snyder, A. C., Molholm, S., & Foxe, J. J. (2011). Oscillatory alpha-band mechanisms and the deployment of spatial attention to anticipated auditory and visual target locations: supramodal or sensory-specific control mechanisms? *Journal of Neuroscience*, *31*(27), 9923–9932. http://doi.org/10.1523/JNEUROSCI.4660-10.2011

Bertini, C., Leo, F., Avenanti, A., & Làdavas, E. (2010). Independent mechanisms for ventriloquism and multisensory integration as revealed by theta-burst stimulation. *European Journal of Neuroscience*, *31*(10), 1791–1799. http://doi.org/10.1111/j.1460-9568.2010.07200.x

Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends in Cognitive Sciences*, *13*(7), 293–301. http://doi.org/10.1016/j.tics.2009.04.005

Heinze, H. J., Mangun, G. R., Burchert, W., Hinrichs, H., Scholz, M., Münte, T. F., … Hillyard, S. A. (1994). Combined spatial and temporal imaging of brain activity during visual selective attention in humans. *Nature*. http://doi.org/10.1038/372543a0

Indovina, I., & Macaluso, E. (2007). Dissociation of stimulus relevance and saliency factors during shifts of visuospatial attention. *Cerebral Cortex*, *17*(7), 1701–1711. http://doi.org/10.1093/cercor/bhl081

Jabar, S. B., & Anderson, B. (2015). Probability shapes perceptual precision: A study in orientation estimation. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(6), 1666–1679. http://doi.org/10.1037/xhp0000121

Jiang, J., Summerfield, C., & Egner, T. (2013). Attention sharpens the distinction between expected and unexpected percepts in the visual brain. *The Journal of Neuroscience*, *33*(47), 18438–47. http://doi.org/10.1523/JNEUROSCI.3308-13.2013

Kamke, M. R., Vieth, H. E., Cottrell, D., & Mattingley, J. B. (2012). Parietal disruption alters audiovisual binding in the sound-induced flash illusion. *NeuroImage*, *62*(3), 1334–1341. http://doi.org/10.1016/j.neuroimage.2012.05.063

Knudsen, E., & Brainard, M. S. (1995). Creating a unified representation of visual and auditory space in the brain. *Annual Review of Neuroscience*, *18*(1), 19–43. http://doi.org/10.1146/annurev.neuro.18.1.19

Kok, P., Rahnev, D., Jehee, J. F. M., Lau, H. C., & de Lange, F. P. (2012). Attention reverses the effect of prediction in silencing sensory signals. *Cerebral Cortex*, *22*(9), 2197–2206. http://doi.org/10.1093/cercor/bhr310

Lange, K. (2013). The ups and downs of temporal orienting: a review of auditory temporal orienting studies and a model associating the heterogeneous findings on the auditory N1 with opposite effects of attention and prediction. *Frontiers in Human Neuroscience*, *7*(June), 1–14. http://doi.org/10.3389/fnhum.2013.00263

Lauter, J. L., Herscovitch, P., Formby, C., & Raichle, M. E. (1985). Tonotopic organization in human auditory cortex revealed by positron emission tomography. *Hearing Res*, *20*, 199–205.

Macaluso, E., Frith, C. D., & Driver, J. (2002). Supramodal effects of covert spatial orienting triggered by visual or tactile events. *Journal of Cognitive Neuroscience*, *14*(3), 389–401. http://doi.org/10.1162/089892902317361912

Macaluso, E., Frith, C., & Driver, J. (2000). Modulation of human visual cortex by crossmodal spatial attention. *Science*, *289*(2000), 1206–1208. http://doi.org/10.1126/science.289.5482.1206

Martinez-Trujillo, J. C., & Treue, S. (2004). Feature-based attention increases the selectivity of population responses in primate visual cortex. *Current Biology*, *14*, 744–751. http://doi.org/10.1016/j

Neumann, O., Van der Heijden, A. H., & Allport, D. A. (1986). Visual selective attention: Introductory remarks. *Psychological Research*, *48*(4), 185–188. http://doi.org/10.1007/bf00309082

Nobre, A. C., & van Ede, F. (2018). Anticipated moments: temporal structure in attention. *Nature Reviews Neuroscience*, *19*(1), 34–48. http://doi.org/10.1038/nrn.2017.141

Rohe, T., & Noppeney, U. (2015). Cortical hierarchies perform Bayesian Causal Inference in multisensory perception. *PLOS Biology*, *13*(2), e1002073. http://doi.org/10.1371/journal.pbio.1002073

Rohe, T., & Noppeney, U. (2016). Distinct computational principles govern multisensory integration in primary sensory and association cortices. *Current Biology*, *26*(4), 509–514. http://doi.org/10.1016/j.cub.2015.12.056

Santangelo, V., & Macaluso, E. (2012). Spatial attention and audiovisual processing. *The New Handbook of Multisensory Processing*, 359–370.

Sereno, M. I., Dale, a M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., … Tootell, R. B. H. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, *268*(May), 889–893. http://doi.org/10.1126/science.7754376

Spence, C., & Driver, J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(4), 1005–1030. http://doi.org/10.1037/0096-1523.22.4.1005

Summerfield, C., & de Lange, F. P. (2014). Expectation in perceptual decision making: neural and computational mechanisms. *Nature Reviews Neuroscience*, *15*(October), 745–756. http://doi.org/10.1038/nrn3838

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*, *13*(9), 403–409. http://doi.org/10.1016/j.tics.2009.06.003

Wallace, M. T., Roberson, G. E., Hairston, W. D., Stein, B. E., Vaughan, J. W., & Schirillo, J. A. (2004). Unifying multisensory signals across time and space. *Experimental Brain Research*, *158*(2), 252–258. http://doi.org/10.1007/s00221-004-1899-9

Yi, D. J., Kelley, T. A., Marois, R., & Chun, M. M. (2006). Attentional modulation of repetition attenuation is anatomically dissociable for scenes and faces. *Brain Research*, *1080*(1), 53–62. http://doi.org/10.1016/j.brainres.2006.01.090

Chapter 8:

Supplementary Materials

CHAPTER 3

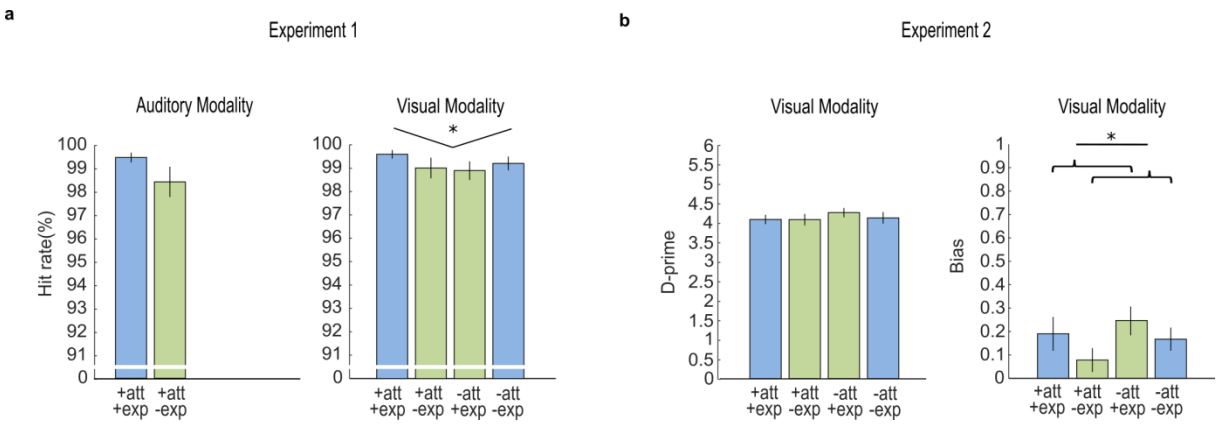


Figure 8.1: Experiment 1: hit rates; Experiment 2: d-prime and bias.

a. Experiment 1: hit rates for each of the six conditions with response requirements. The brackets and stars indicate significance of main effects and interactions. * $p < 0.05$.

b. Experiment 2: d-prime and bias values for each of the four conditions for the visual modality. The brackets and stars indicate significance of main effects and interactions. * $p < 0.05$.

Table 8.1: Spatial signal probability, general response probability and spatially selective response probability for each condition in experiment 1 and experiment 2.
S = signal location, $H_{..att..exp}$ = hemifield e.g. +att +exp, R = response.
Please note that these probabilities are computed pooled over auditory and visual modalities.

	Conditions			
	+att +exp	+att -exp	-att +exp	-att -exp
Experiment 1				
Spatial signal probability $P(S = H_{..att..exp})$	0.65	0.35	0.65	0.35
General response probability $P(R = +)$	0.9	0.6	0.6	0.9
Spatially selective response probability $P(R = + \mid S = H_{..att..exp})$	1	1	0.384	0.714
Experiment 2				
Spatial signal probability $P(S = H_{..att..exp})$	0.625	0.375	0.625	0.375
General response probability $P(R = +)$	0.123	0.123	0.123	0.123
Spatially selective response probability $P(R = + \mid S = H_{..att..exp})$	0.131	0.218	0.065	0.109

The spatial signal probability is the probability that a signal is presented in a particular hemifield, e.g. in experiment 1: P = 0.65 for the hemifield that is attended and expected vs. P = 0.35 for the hemifield that is unattended and unexpected (n.b. the spatial signal probabilities sum to one for the two conditions presented in the same run).
The general response probability is the probability that a trial requires a response in a particular run irrespective of where the signal is presented, e.g. in experiment 1: P = 0.9 for runs where attention and expectation are congruent vs. P = 0.6 for runs where expectation and attention are incongruent.
The spatially selective response probability is the probability that a response is required conditioned on the signal being presented in a particular hemifield, e.g. in

experiment 1: the probability of a response is P = 0.384, if the signal is presented in a hemifield that is not attended, but expected.

As a result, we can compute the general response probability by multiplying the spatially selective response probability with the spatial signal probability and marginalizing over the spatial signal probability; e.g.

$P(R = +) = P(R = +|S = H_{+att+exp})\ P(S = H_{+att+exp}) + P(R = +|S = H_{-att-exp})\ P(S = H_{-att-exp})$

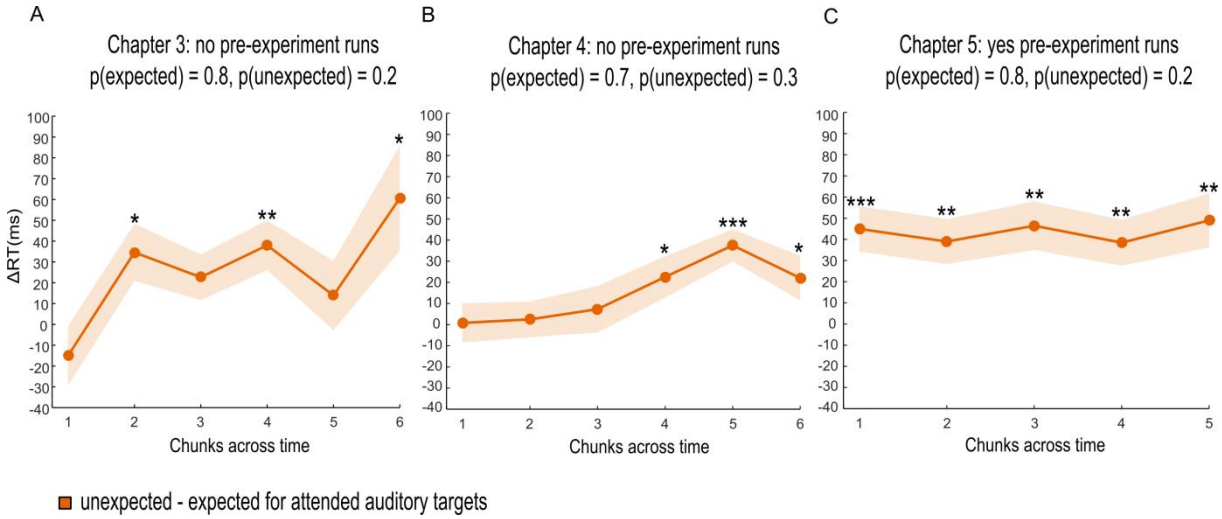


Figure 8.2: Comparison of modality specific effects of response probability (across subjects' mean ΔRT) over time (i.e., across chunks) for audition in chapter 3, 4 and 5. ΔRT for each chunk was tested as > 0 using one-sample t-tests. Shading represents standard error of the mean. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. A. Chapter 3: Auditory stimuli were presented with a ratio of 4/1 in the expected/unexpected hemifields. B. Chapter 4: Auditory stimuli were presented with a ratio of 2.33/1 in the expected/unexpected hemifields. C. Chapter 5 (fMRI experiment): pre-experiment runs, i.e., runs prior to the experiments, were used to boost the implicit learning of auditory spatial signal probability. Auditory stimuli were presented with a ratio of 4/1 in the expected/unexpected hemifields. The difference between expected and unexpected auditory targets (when attended, i.e., the effect of response probability) was absent during the first temporal chunk and increased over time (i.e., significantly > 0) when participants were not exposed to expectation manipulation before the main experiment (i.e., no pre-experiment runs, chapters 3 and 4). Conversely, when participants' implicit learning of auditory spatial signal probability was previously boosted via pre-experiment runs, the effect of response probability was significantly > 0 already during the first temporal chunk and it remained constant.

CHAPTER 5

Table 8.2: Group mean reaction times (RT) for each stimulus modality in each condition for psychophysics and fMRI experiment.

Note. Standard errors of the mean (SEM) are given in parentheses.

Experiment	Auditory modality		Visual modality			
	+att +exp	+att -exp	+att +exp	+att -exp	-att +exp	-att -exp
Psychophysics						
RT (ms) (SEM)	530.7 (17.1)	566.8 (15.3)	446.9 (10.2)	458.3 (9.4)	487 (11.4)	472.6 (11.8)
fMRI						
RT (ms) (SEM)	508.4 (24.5)	552.9 (27.4)	432.3 (13.9)	441.2 (12.9)	467.1 (14.9)	454.2 (14.7)

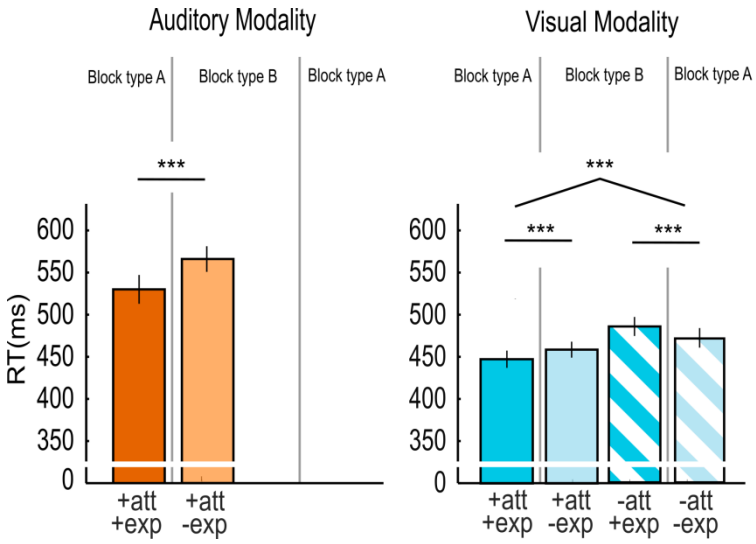


Figure 8.3: Behavioural results for the psychophysics experiment.

Bar plots representing across subjects' mean (±SEM) response times for each of the six conditions with response requirements for the experimental runs of the psychophysics experiment. The brackets and stars indicate significance of main effects and interactions. ** $p < 0.01$; *** $p < 0.001$. audition: orange; vision: blue; attended: full pattern; unattended: striped pattern; expected: dark shade; unexpected: light shade.

Table 8.3: Main effect of auditory spatial attention for auditory and visual stimuli. *p*-values are FWE corrected at the cluster level for multiple comparisons within the entire brain. Auxiliary uncorrected voxel threshold of $p < 0.001$. L, Left; R, right.

Brain regions	**MNI coordinates**			**z-score (peak)**	**Cluster size**	***p*-value (cluster)**
	x	***y***	***z***			
Attended vs unattended hemifield Auditory stimuli						
R cerebellum	18	-54	-20	>8	1786	0.000
L thalamus/ R thalamus	0	-6	6	5.16		
L central sulcus	-40	-20	54	>8	1386	0.000
L frontoparietal operculum	-54	-18	20	4.22	294	0.003
	-44	-6	14	4.89	177	0.032
R caudate	18	12	18	4.68	1781	0.000
L caudate	-16	8	20	4.60		
L putamen	-20	8	-4	4.47	717	0.000
L hippocampus	-20	-20	-16	4.17		
R hippocampus	22	-20	-18	5.08	339	0.001

[continues from previous page]: Main effect of auditory spatial attention for auditory and visual stimuli.

p-values are FWE corrected at the cluster level for multiple comparisons within the entire brain. Auxiliary uncorrected voxel threshold of $p < 0.001$. L, Left; R, right.

Unattended vs attended hemifield Auditory stimuli						
R superior frontal gyrus	24	-8	60	6.37	5494	0.000
R superior frontal sulcus/ R precentral sulcus	28	-10	58	6.02		
R postcentral sulcus/ R intraparietal sulcus	42	-26	38	5.22		
R inferior parietal lobule	60	-38	30	3.85		
R anterior cingulate gyrus	10	24	26	4.40		
L superior frontal gyrus	-16	-10	62	4.15	668	0.000
L superior frontal sulcus/ L precentral sulcus	-32	-4	54	4.90		
R middle frontal gyrus	30	46	22	5.73	524	0.000
R anterior insula	36	16	2	5.26	1239	0.000
R posterior insula	36	-12	8	4.54		
L anterior insula	-30	22	4	4.72	418	0.000
Unattended vs attended hemifield Visual stimuli						
R superior frontal gyrus	18	-4	64	4.09	2335	0.000
R superior frontal sulcus/ R precentral sulcus	26	-4	46	4.45		
L superior frontal gyrus	-24	-4	60	4.70		
L superior frontal sulcus/ L precentral sulcus	-26	-8	54	4.19		
R anterior cingulate gyrus	10	14	36	3.86		
L anterior cingulate gyrus	-4	8	50	4.54		
R postcentral sulcus/ R intraparietal sulcus	42	-34	42	4.02	472	0.000
L postcentral sulcus/ L intraparietal sulcus	-44	-34	38	4.49	488	0.000
R anterior insula	30	20	8	4.41	227	0.012

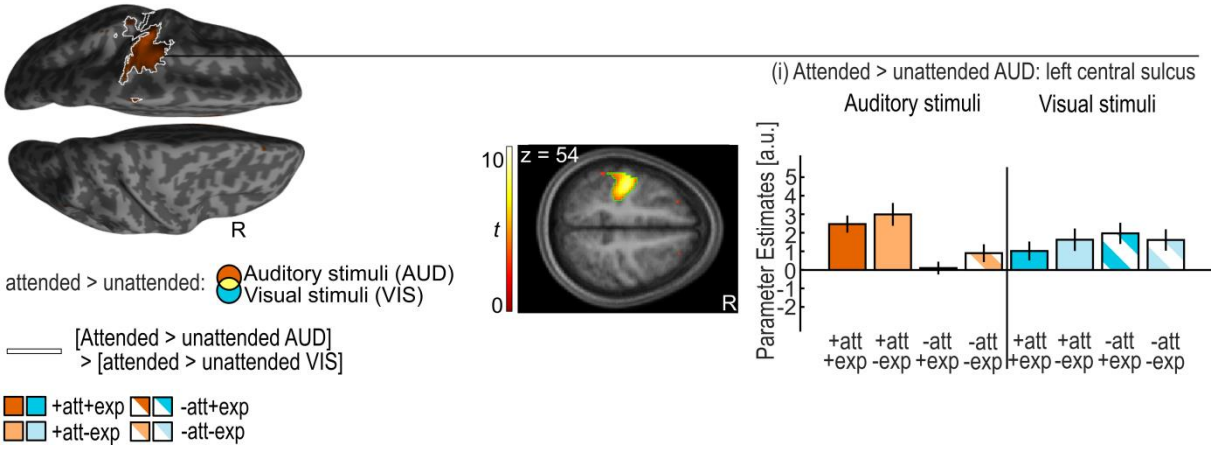


Figure 8.4: Attended > unattended for auditory and visual stimuli.

Activation increases for attended > unattended stimuli for auditory stimuli (orange) are rendered on an inflated canonical brain; they are encircled in white, if they are significantly greater for auditory than visual stimuli (i.e. interaction). Height threshold of $p < 0.001$, uncorrected; extent threshold k > 0 voxels.

Bar plots show the parameter estimates (across participants mean ± SEM, averaged across all voxels in the green encircled cluster) in the left central sulcus that are displayed on axial slices of a mean image created by averaging the subjects' normalized structural images; the bar graphs represent the size of the effect in nondimensional unit (corresponding to percentage whole-brain mean). Audition: orange; vision: blue; attended: full pattern; unattended: striped pattern; expected: dark shade; unexpected: light shade.