THE EFFECT OF VISUAL PERCEPTUAL LOAD ON AUDITORY PROCESSING

KATHARINE MOLLOY

Institute of Cognitive Neuroscience

University College London

Submitted in partial fulfilment for the degree of Doctor of Philosophy

June 2019

DECLARATION

This dissertation is the result of my own work and includes nothing that is derived from other sources or is the outcome of work done in collaboration except where specifically indicated in the text. It has not been previously submitted, in part or whole, to any university of institution for any degree, diploma, or other qualification.

In accordance with the University College London guidelines, this thesis does not exceed 100,000 words, and it contains fewer than 150 figures.

Signed: ………………………………………………………………………………

Date: 29th December 2019

Katharine Molloy

London

ABSTRACT

Many fundamental aspects of auditory processing occur even when we are not attending to the auditory environment. This has led to a popular belief that auditory signals are analysed in a largely pre-attentive manner, allowing hearing to serve as an early warning system. However, models of attention highlight that even processes that occur by default may rely on access to perceptual resources, and so can fail in situations when demand on sensory systems is particularly high. If this is the case for auditory processing, the classic paradigms employed in auditory attention research are not sufficient to distinguish between a process that is truly automatic (i.e., will occur regardless of any competing demands on sensory processing) and one that occurs passively (i.e., without explicit intent) but is dependent on resource-availability. An approach that addresses explicitly whether an aspect of auditory analysis is contingent on access to capacity-limited resources is to control the resources available to the process; this can be achieved by actively engaging attention in a different task that depletes perceptual capacity to a greater or lesser extent. If the critical auditory process is affected by manipulating the perceptual demands of the attended task this suggests that it is subject to the availability of processing resources; in contrast a process that is automatic should not be affected by the level of load in the attended task.

This approach has been firmly established within vision, but has been used relatively little to explore auditory processing. In the experiments presented in this thesis, I use MEG, pupillometry and behavioural dual-task designs to explore how auditory processing is impacted by visual perceptual load. The MEG data presented illustrate that both the overall amplitude of auditory responses, and the computational capacity of the auditory system are affected by the degree of perceptual load in a concurrent visual task. These effects are mirrored by the pupillometry data in which pupil dilation is found to reflect both the degree of load in the attended visual task (with larger pupil dilation to the high compared to the low load visual load task), and the sensory processing of irrelevant auditory signals (with reduced dilation to sounds under high versus low visual load). The data highlight that previous assumptions that auditory processing can occur automatically may be too simplistic; in fact, though many aspects of auditory processing occur passively and benefit from the allocation of spare capacity, they are not strictly automatic. Moreover, the data indicate that the impact of visual load can be seen even on the early sensory cortical responses to sound, suggesting not only that cortical processing of auditory signals is dependent on the availability of resources, but also that these resources are part of a global pool shared between vision and audition.

IMPACT STATEMENT

In everyday environments, our auditory and visual perceptual systems are bombarded by a large amount of information – there is a perpetual stream of incoming sights and sounds that we cannot process all at once. In order to focus on a task, we must filter out data that is irrelevant and give our attention only to information that relates to the task at hand. This process, called selective attention, is fundamental to our ability to avoid distraction. However, filtering out all information that is not related to the current task leaves us open to missing unexpected but nonetheless important events, such as someone calling our name, an alarm, or even someone walking suddenly into the road while we are driving. It is therefore vital that the brain maintains some degree of monitoring of the environment outside our current task so that any potentially relevant information can be brought to our attention. Since we receive auditory information from all directions regardless of the way we are facing, and we cannot shut our ears as we do our eyes, hearing has the potential to act as an early warning system that constantly monitors the environment. However, processing all the information that reaches our ears is a complex task, and it is currently unclear whether auditory processing requires access to (potentially limited) resources. If this is the case, in situations where there is a lot of information to process, the auditory system may become overwhelmed and fail to perform its role as an early warning system. Understanding whether this is the case is critical so that we know whether we can rely on our hearing to warn us of potential dangers. Another question that remains unclear is whether any resource pool for processing is exclusive to hearing, or if perhaps it is shared with other sensory systems. Can incidents such as failing to hear your name when called be exacerbated when you are focused on a demanding visual task such as playing video games or driving? With modern technology and the advent of smart devices, we are more and more often subject to multiple streams of information coming in both visually (e.g., the news feed on a phone, the street in front of us as we walk) and via audition (e.g., alert tones, music, and people around us talking), making this issue a timely one. The aim of the work in this thesis is to understand the degree to which visual information processing can interfere with hearing. The results show that high demands on the visual system can result in reduced sensory processing of sound to the point where we can become ‘deaf’ to sounds that should be audible. Thus, in order for hearing to serve its purpose as an early warning system, we must be careful not to overload either the auditory or visual systems with information.

ACKNOWLEDGEMENTS

> *"My largest P300 waves have occurred on days when I perceived the target with a vividness and intensity more appropriate for a visiting divinity than a simple tone. My smallest P300 waves occurred when I was tired and preoccupied with other matters. At those times, I felt that I had been given only so many P300 waves and that they should not be so foolishly squandered on insignificant tasks."*
>
> *Terence W. Picton. The P300 wave of the human event-related potential.*

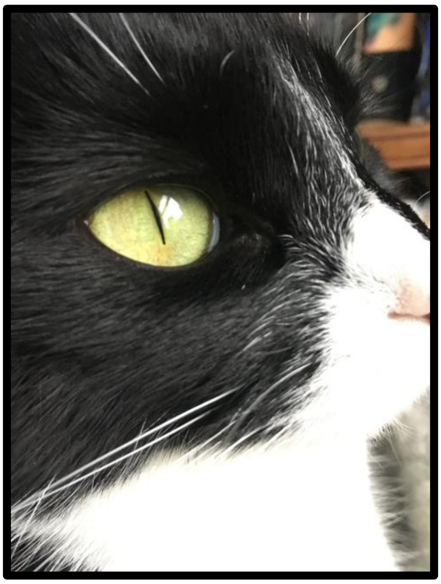
This quote from Terence Picton captures the invigorating highs and exhausting lows I have experienced during my time as a PhD student. Though it was ultimately rewarding, I would not have made it through those difficult times without the support I received from the people around me.

I have had the good fortune to be supervised by two incredible women, Nilli Lavie and Maria Chait, who are such strong leaders of their labs because they are both exceptional scientists and supportive mentors. The confidence and enthusiasm you both embody is something I hope I have also achieved.

To all the members past and present of Lavie and Chait labs thank you for making my scientific life so much richer. In particular, Anoo, Nicolas and Sijia from (the much extended!) Chait lab, Dana Luke and Josh from Lavie Lab, and Warren, Joe and Nick from the wider Ear Institute.

Huge thanks also go to my friends and the extended Molloy family, who have put up with me cancelling plans, not responding to any form of communication, and when I eventually do see them, clumsily trying to remember how to function in social situations. In particular, Amy and Laura, whose determined stream of WhatsApp messages covering our cats, medieval art and grammar memes, fluency in a wide variety of languages, and most importantly their witty and kind selves, may have been the only thing that kept me sane. The award for most forgiving goes to Tom, whose wedding was just before my submission deadline and who, though I missed said wedding, still takes his role as the instigator of communication very seriously. Massive thanks as well to Pete, who after finishing his own PhD took me on as a flatmate in spite of this extending the anxiety for a further three years!

To my darling cat Auri, who inspired the pupil diameter analysis after I frequently watched her eyes transition between the two states below. She has provided excellent company, laughter and (reluctantly) hugs, and made sure that I got out of bed to feed her, even on mornings I wasn't sure I could.



Finally, to my Mum and Dad, Val and Declan, and siblings Dom and Jess, who have been with me throughout offering practical, emotional and even financial (!) support. You can't imagine what it's meant that you've shown me such care and encouragement. You are some of the best people I know and I quite literally couldn't have done it without you.

… and I promise I'll pay you back!

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

GENERAL TERMINOLOGY

ACh	Acetylcholine
ANOVA	Analysis of variance
ARN	Attention-related negativity
BOLD	Blood oxygen level dependent
dB	Decibel
DSS	Denoising source separation
EEG	Electroencephalography
ERF	Event-related field
ERP	Event-related potential
FFR	Frequency following response
fMRI	Functional magnetic resonance imaging
FRN	Figure-related negativity
Hz	Hertz
IPS	Intra-parietal sulcus
LC	Locus coeruleus
LGN	Lateral geniculate nucleus
MEG	Magnetoencephalography
MMN	Mismatch negativity
MTG	Middle temporal gyrus
NE	Norepinephrine
ORN	Object-related negativity
PAR	Post-auricular reflex
PPA	Parahippocampal place area
PT	Planum temporale
RMS	Root mean square
RSVP	Rapid serial visual presentation
SFG	Stochastic figure ground
SPL	Sound pressure level
SSR	Steady state response
STG	Superior temporal gyrus

STS	Superior temporal sulcus

EXPERIMENTAL CONDITIONS

LL	Low load
HL	High load
A	Auditory stimulus only
V	Visual stimulus only
AV	Auditory + visual stimuli
SA	Sound absent
SP	Sound present
FA	Figure absent
FP	Figure present
CS	Colour-shape search task
LS	Letter search task
T-LS	Tone + letter search dataset
SFG-LS	SFG + letter search dataset
SFG-CS	SFG + colour-shape search task

CHAPTER 1: INTRODUCTION

1.1 AUDITORY PROCESSING AND ATTENTION

The auditory system is constantly receiving information and cannot be switched off as we do in vision when we shut our eyes. We also receive auditory information from all directions at once, unlike vision which is restricted to the field of view. Because of this, a popular view is that hearing serves as an early warning system that automatically monitors the environment and alerts us to potentially important sounds (Eramudugolla et al., 2005; Demany et al., 2010; Näätänen et al., 2010; Masutomi et al., 2016; Murphy et al., 2017). However, in order for the auditory system to perform this role effectively, analysis of auditory scenes must occur even when our attention is engaged elsewhere. It is therefore vital that we understand the extent to which auditory processing occurs for sounds that fall outside the explicit focus of attention, in order to establish whether it is safe to rely on our hearing to alert us when necessary.

The question of how attention affects auditory processing is one which has received a great deal of interest. In the first section of this chapter, I will review previous approaches and the conclusions that have been drawn.

1.1.1 Effect of active attention

1.1.1.1 Experimental paradigms

The vast majority of previous research has explored this question by manipulating the *explicit* focus of attention. This approach was facilitated by the advent of electrophysiology, since neural activity could be used as an objective measure of the processing of unattended auditory stimuli in cases where attention was engaged in an alternate task. One popular paradigm in electrophysiological studies has been to contrast a condition where attention is directed towards the critical auditory stimuli (‘active’) with a condition of ‘passive’ listening, where participants either perform no task or engage in an on-going (typically visual) task such as reading or watching a silent film. One drawback of this approach is that over the timeframes used in these passive conditions (in the order of minutes but often up to an hour), attention naturally shifts focus within the environment. The propensity for these shifts will vary from subject to subject and also throughout the experiment, so it is possible (perhaps even likely) that there are times when participants’ attention is directly focused on the auditory signals.

A second approach that attempts to account for this confound is the ‘attend vs ignore’ paradigm, in which neural activity is contrasted between active attention and a condition where participants perform an alternative task based on separate stimuli. In this paradigm both critical and ‘decoy’ stimuli are presented in all conditions – only the focus of attention is manipulated – giving control of low-level stimulus characteristics. Additionally, since the critical and decoy task stimuli are usually presented concurrently, performance on the decoy task can be used to gauge how well attention was controlled during the ignore condition. For these paradigms, the decoy tasks are typically visual, or an auditory task presented in the opposite ear to the critical stimuli, to avoid low-level stimulus interference.

A third approach to investigate the effect of attention is to contrast responses to sounds that were cued versus uncued. This paradigm relies on the fact that attention can be pre-emptively focused on certain aspects of the auditory scene, for instance a certain frequency or spatial location. During uncued listening, attention must be spread widely to accommodate the large variety of potential stimuli, resulting in relatively diffuse coverage, whereas a cue allows the listener to explicitly focus their attention in expectation of the stimulus.

Across this variety of approaches, active attention has reliably been shown to improve auditory processing. In the rest of this section, I will briefly summarise the ways in which active attention has been shown to affect various aspects of auditory processing.

1.1.1.2 Enhancement of evoked responses

The most straightforward influence of active attention is manifest as an enhancement of evoked responses to auditory stimuli. This effect has been seen as early as the brainstem. When auditory signals are modulated in amplitude at a set rate (usually at relatively low frequencies to avoid inducing a pitch percept) they will elicit activity called the frequency following response (FFR), where neural firing in the brainstem timelocks to the envelope of the auditory stimulus. FFR responses have been shown to be stronger to stimuli presented to the attended ear relative to those presented to the unattended ear (Hoormann et al., 2000). Similarly, early cortical sound-evoked responses such as the P1, N1 and P2 onset responses (and their MEG counterparts, the M50, M100 and M200) are larger for sounds that are attended versus ignored, and for cued versus uncued sounds (Hillyard et al., 1973; Parasuraman, 1978; Hansen and Hillyard, 1983; Woldorff et al., 1987; Woldorff et al., 1993).

1.1.1.3 Change detection

In addition to the basic signal enhancement that can occur when attention is engaged, there is also evidence that attention can affect various vital aspects of auditory processing. One of the most fundamental of these is the task of detecting changes in the auditory soundscape, since any change from the current, known state may provide new and potentially relevant information about the environment. There has been considerable work invested in mapping the neural activity associated with changes in the auditory environment. This work has characteristically focused on an event-related potential (ERP) component termed the mismatch negativity (MMN) that is elicited in response to a deviation within an established auditory signal (Näätänen et al., 2007 provides an extensive review).

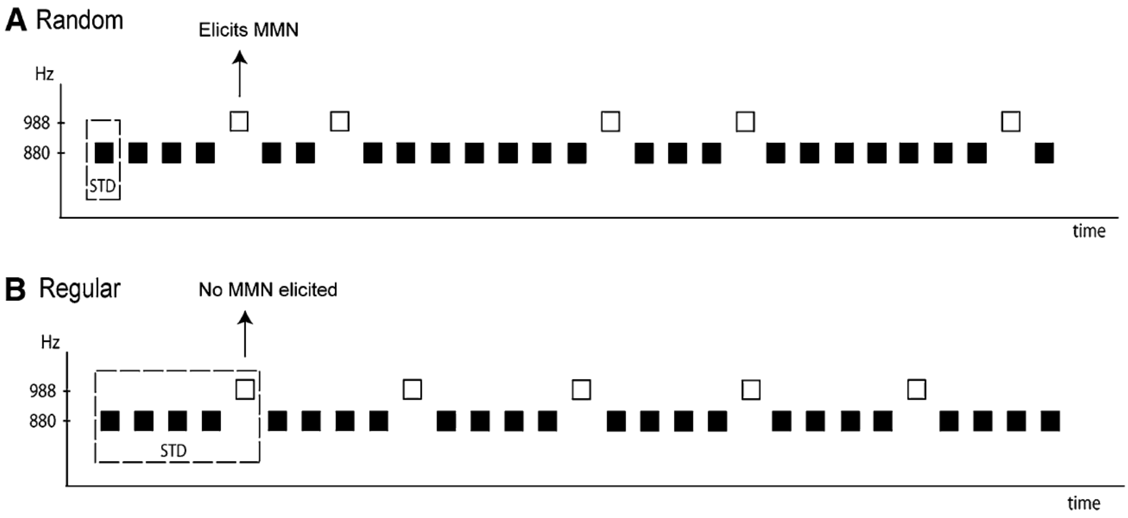


Figure 1.1: Schematic of the MMN 'Oddball' Stimulus, from Sussman et al. (2014)

A. *Schematic for the classic oddball paradigm, in which frequent 'standard' tones are presented with occasional 'deviant' tones occurring at random, infrequent positions. This stimulus design elicits an MMN response.* ***B.*** *Schematic for a modified oddball task where the deviant tone is equally rare, but occurs at predictable intervals. This stimulus design typically does not elicit an MMN.*

The stimuli classically used to explore the MMN consist of a regular sequence of identical ‘standard’ tones, with rare deviants that occur along a low-level feature such as pitch, intensity, or duration – referred to as ‘oddball’ paradigms (Figure 1.1). However, MMNs have also been found to occur in response to violation of more complex patterns (e.g. Saarinen et al., 1992; Alain et al., 1994; Alain and Woods, 1997; Sussman et al., 1998b; Alain et al., 1999a; Alain et al., 1999b; Paavilainen et al., 2001), suggesting that the MMN is not simply a low-level deviance-detection mechanism. A particularly good illustration of this is a paradigm where a local pitch deviant (B) occurs at a predictable point in a sequence (e.g., AAAAB-AAAAB), with rare pattern deviants also occurring (see Figure 1.1B). For slow presentation rates the B tone reliably elicits an MMN (Scherg et al., 1989), but at faster rates that encourage temporal grouping, B tones in the common sequence (AAAAB) did not elicit an MMN, only those within pattern deviants triggered the response (Sussman et al., 1998b). The same effect can be produced at slower rates by asking participants to explicitly attend to the pattern (Sussman et al., 2002), illustrating that focused attention can guide the MMN process. A similar effect of attentional focus was illustrated by Sussman et al. (1998a), who showed that guiding attention towards a subset of tones within a sequence (high tones) allowed for elicitation of an MMN to deviant patterns within that subsequence that did not occur when both low and high tones were attended together. One model proposed by Sussman (2007) to explain the effect of attention focus is that it guides the formation of the standard: during stimulus competition, selective attention isolates stimuli of interest on which the standard should be based, and also allows for the formation of more sophisticated rules.

1.1.1.4 Segregating concurrent, static sounds

In order to perform many of its core functions (including some of the more sophisticated versions of change detection described in the previous section), the auditory system must first separate information that arrives at the ear as one compound signal back into separate auditory objects. This process, broadly referred to as scene analysis, must occur even though signals from separate objects typically overlap spectrally and temporally, and also fluctuate over time.

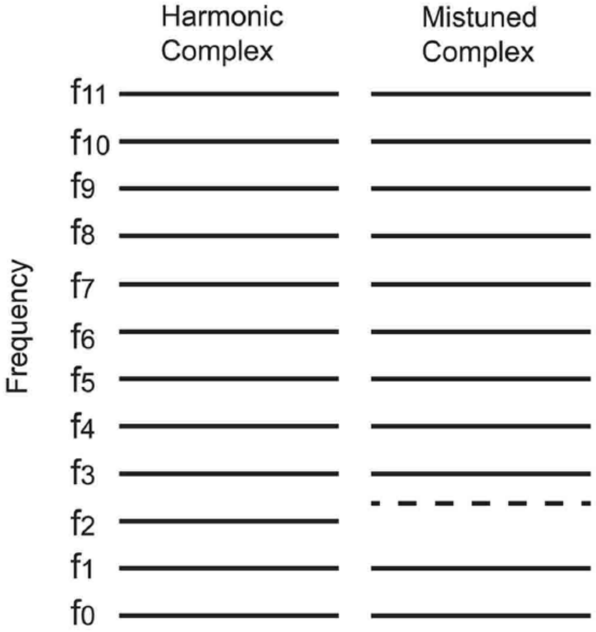


Figure 1.2: Schematic of the 'Mistuned Harmonic' Stimulus from Snyder et al. (2012)

Typically, natural sounds consist of harmonic complexes (left), which are formed of frequencies of integer multiples of the fundamental, f0. In the mistuned stimulus (right) the frequency of one of the components is shifted so that it is no longer an integer multiple of the fundamental.

One of the low-level cues underlying scene analysis is harmonicity: concurrent sound components that form a harmonic complex (i.e., consist of integer multiples of a common frequency, see Figure 1.2) are likely to be grouped together and perceived as a single auditory object (Lipp et al., 2010). Similarly, on- and offset synchrony also act as a strong cue that sound components have a shared source. The effect of mistuning one of the sound components, or giving it an onset asynchrony relative to the main complex, is that the single component is likely to be segregated perceptually from the rest of the sound, and heard as a distinct auditory object. The percept of a separate object occurs more readily for components that show a large deviation from the expected frequency or time, i.e., these factors provide an inherent, or endogenous cue for segregation. However, components that are cued exogenously (by, for example, a cue tone that precedes the sound complex and occurs at the frequency of the segregable component), are also more likely to be perceived as segregated (Darwin et al., 1995). ERP studies support these behavioural findings, showing that there is a greater distinction between the ERPs to tuned versus mistuned sound complexes during active listening than passive listening (Alain et al., 2001). There is also evidence that top-down attention can be used to enhance stimulus processing of particular features of interest; Dyson and Alain (2008) changed the task instructions so that participants attended to either the harmonicity (tuned vs mistuned) or the overall pitch (high vs low) of harmonic complexes, and illustrated that an enhanced P2 occurred in response to mistuned signals only when the current task was based on harmonicity cues. These findings suggest that while attention can improve segregation whether it is en- or exogenously cued, endogenous cues alone may not always be sufficient for segregation.

1.1.1.5 Segregating simple sequences of tones – the streaming paradigm

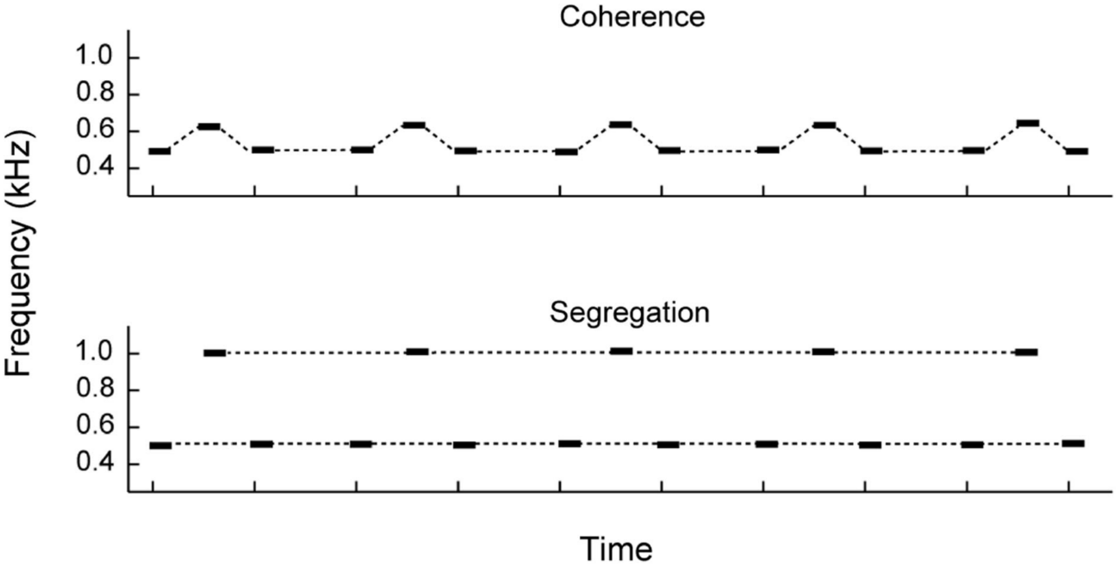


Figure 1.3: Schematic of the 'Streaming' Stimulus, from Snyder et al. (2012).

The streaming stimulus consists of sequences of regularly presented tones of two different frequencies (A and B) in a pattern: ABA-ABA-. Two percepts are possible; the first is illustrated in the top panel, in which tones at both frequencies are perceived as part of a single stream with an alternating pitch pattern – referred to as 'coherent' or 'fused' percept. The second possibility (illustrated in the bottom panel) is that the tones at the two frequencies are perceived as separate streams, each of which consists of tones at only one frequency – referred to as 'segregated' or 'streamed'.

In contrast to segregating concurrent sounds – a process during which overlapping signals must be resolved into independent auditory representations – another process that contributes to resolving an auditory scene is the binding of components that occur at different times if they are likely to represent on-going activity from a single source. This process allows an evolving auditory scene to be perceptually segregated into multiple 'streams' that represent different objects. The paradigm that has been most widely used to explore this stream segregation uses regular sequences of pure tones, typically of just two different frequencies (A and B), which are presented in a regular pattern ABA-ABA-ABA-, where the dash indicates a silent gap (van Noorden, 1975). With the right parameters for tone frequencies and presentation rate, this creates an ambiguous signal that can be perceived as fused (i.e., one stream with alternating frequencies) or segregated into two on-going streams (A - A - A - A and B - - - B - - - B). Typically, sounds are initially perceived as a single fused stream, but over time the percept of two separate streams can build up.

Stimulus parameters that affect the build-up of streaming include the frequency difference between tones in the stream, and the rate of presentation, with larger frequency differences and faster rates both leading to a faster build-up of segregated streams (van Noorden, 1975; Anstis and Saida, 1985). However, there is also considerable evidence that top-down attention can impact the build-up of streaming. Carlyon et al., (2001) presented the ABA segregation stimuli to one ear and a competing auditory task to the other ear – participants performed one condition where they attended to the streaming signals and reported when they perceived two streams, and a second condition when they performed the competing task for 10 seconds before switching to the streaming task. The results showed that when participants had been attending to the alternative task stimuli, their percept of the ABA signals was considerably less likely to have built up to 'segregated' relative to when they had been attending to the streaming signals throughout. Using a similar paradigm, Cusack et al., (2004) illustrated that any time attention was switched away from the streaming signals, the build-up of streaming was abruptly reset to the default of one fused stream. These findings suggest that not only does attention provide substantial facilitation of streaming, but in some cases, it may even be necessary for streaming to occur.

1.1.2 What is the fate of ignored sound?

The cases where auditory processes appear to fail if attention is not engaged highlight an important question: though there is broad consensus that auditory processing can be improved with explicit, directed attention, the extent to which effective auditory processing occurs when our attention is directed elsewhere remains much more unclear. There is considerable evidence that some degree of auditory processing can occur without explicit intent; for instance, it is common for an unexpected sound to capture our attention even when we are focused on a task (e.g. Näätänen and Michie, 1979). This phenomenon of attention capture serves a vital evolutionary function, since unexpected or salient auditory events may indicate a (potentially dangerous) change in the environment, but requires the auditory system to be perpetually monitoring the auditory environment. Fully processing the auditory scene at all times would be computationally intensive, and in many situations (such as when we need to maintain focus on a visual task) may in fact prove disruptive if our attention is easily captured by irrelevant sounds. It is therefore vital to understand how a balance is established between these two factors, and in particular to determine which aspects of auditory processing occur when sound is ignored.

1.1.2.1 Change detection

One of the key aspects likely to be part of any early warning system that allows unexpected sounds to trigger attention capture is a change detection mechanism (Näätänen and Michie, 1979). There is considerable evidence that changes in low-level features of an auditory signal do indeed lead to attention capture, even in situations when these changes are expected and known to be irrelevant (Dalton and Lavie, 2004).

During attention capture paradigms, MMN responses typically occur to all deviant stimuli, regardless of whether participants are actively engaged in a task (Escera et al., 1998; Schröger and Wolff, 1998; Escera and Corral, 2007), though they are larger for deviants that result in attention capture (Sabri et al., 2013). This is consistent with the classic view that the MMN reflects an automatic process, making it a plausible candidate for an early warning system.

However, there have also been studies which call the role of the MMN as automatic into question. Even for simple deviants such as a change in pitch or intensity in a series of otherwise uniform pure tones, directing attention away from signals to a competing set of stimuli can abolish the MMN responses to the unattended ear. Woldorff, Hackley and Hillyard (1991) presented independent sequences of tones to each ear, with participants attending to one ear and ignoring the other. They found that MMNs elicited by intensity deviants in the unattended ear were smaller than those to the attended ear, and that when the presentation speed was increased, the unattended ear MMNs decreased to the point where they were not detectable. Subsequent studies have shown similar effects of attention on MMNs to deviants along various stimulus features (e.g. Trejo et al., 1995; Woldorff et al., 1998; Sussman et al., 2003), suggesting that many facets of the process may be more reliant on attention than was commonly understood.

1.1.2.2 Sound segregation

Another aspect of auditory processing where the role of attention remains unclear is sound segregation. As reviewed in section 1.1.1.5, some studies have suggested that in order to achieve perceptual segregation in the streaming paradigm, attention must be maintained on the stimuli since the percept of two separate streams appears to build up, and is rapidly disrupted when attention is disengaged (e.g. Carlyon et al., 2001; Cusack et al., 2004; Thompson et al., 2011; Billig and Carlyon, 2016).

However, there is also some evidence that this type of streaming can occur outside the focus of attention. For instance, Macken et al. (2003) probed streaming using the irrelevant sound effect, in which presentation of distracting sounds during a delay period can impair short-term (internally rehearsed) recall of letter sequences. This paradigm relies on the fact that irrelevant sounds are more disruptive to serial recall when they are perceived as an alternating single source compared to when they are heard as separate, ‘static’ sources. Macken et al. used the ABA- streaming signals as the distractor sounds while participants actively engaged in a serial recall task. They showed that increasing the speed of presentation of the streaming signals (a manipulation known to improve the build-up of streaming when sounds are attended) affected the degree to which the irrelevant sounds interfered with the recall task. This suggests that the degree of streaming of the distractor sounds had been manipulated, in spite of the fact that explicit attention was not focused on the streaming signals in any of the conditions. Evidence that some streaming can occur even for ignored signals was also shown using MMN paradigms, where, for example Sussman et al. (1998a) presented MMN-style deviance signals while participants read a book. The MMN signals contained two independent sequences, presented within distinct frequency regions, for which pattern deviants were only apparent once the signal was separated into the two sub-streams. They showed that for slow presentation rates (where streaming is less likely) there were no MMNs to the deviants, but at a faster rate the

MMNs were elicited, suggesting that for the fast presentation, signals were streamed even though attention was not (apparently) involved. This study and others with similar results (e.g. Sussman et al., 1999; Sussman et al., 2001; Nager et al., 2003; Müller et al., 2005) led to claims that streaming must occur preattentively, and even that it was part of an 'automatic auditory intelligence' that performs processes such as sound organisation, stream segregation and pattern encoding, and which monitors the resulting objects for change (Näätänen et al., 2010).

The conclusions drawn based on different experimental paradigms differ profoundly, with one view stating that auditory streaming is part of an automatic scene analysis that is engaged in passively monitoring the environment, and the other suggesting that attention must be actively maintained on the streaming signals for segregation to occur.

1.2 ATTENTION THEORY

The examples given in the previous section illustrate that there remains a deep-seated confusion about the degree to which unattended auditory stimuli are processed. Within vision this question was also the basis for a long-standing debate; 'early selection' accounts (e.g. Broadbent, 1958; Neisser and Becklen, 1975) proposed that irrelevant sensory input is filtered out at a low level before it is processed, whereas 'late selection' accounts (e.g. Deutsch and Deutsch, 1963; Tipper, 1985) claimed that unattended information is processed to a much higher level before being discarded by the attentional filter. There was a wealth of support for both arguments, with irrelevant information being processed to different levels in different situations (see Driver, 2001 for a review). One theory that proposes a resolution to these apparently disparate accounts is Load Theory of Attention and Cognitive Control (Lavie, 1995, 2005, 2010). The model has received substantial support within vision, providing an account that can explain instances of both early and late selection. However, Load Theory has yet to be applied with any consensus to auditory processing.

1.2.1 Load Theory

Load Theory proposes that there is a limited processing capacity available during perception. Attention guides perception by giving priorities, so that task-relevant stimuli are processed preferentially. However, crucially, sensory processing continues automatically until capacity is reached, even if this means processing stimuli that are outside the explicitly attended task set. Thus, the extent to which irrelevant stimuli are processed depends both on how much processing capacity is used up by the attended task (*perceptual load*), and how well the cognitive system maintains the processing hierarchy (an ability which is affected by the degree of *cognitive load*).

1.2.1.1 Perceptual Load

The degree to which a task depletes processing resources is referred to as its perceptual load. If a task involves a high level of perceptual load, it may exhaust perceptual capacity so that irrelevant stimuli are not processed (leading to, in effect, early selection). In contrast, if an attended task is of low perceptual load, the excess perceptual capacity will ‘spill over’ and process irrelevant stimuli (i.e., late selection).

Perceptual load is quantified by the overall amount of perceptual information that must be processed in order to perform the task, and the time available for that processing. The overall amount of perceptual information is affected by a number of factors. The first is the number of objects to be processed: assuming all other factors are equal, tasks where a large number of objects must be processed generally induce a higher perceptual load than those with fewer objects. The second factor is how complex the objects are: simple objects such as a uniformly coloured circle (which are comprised of relatively few low level features) can be considered to contain inherently less perceptual information than a more complex objects formed of multiple features (for instance a series of different coloured circles embedded within one another), and thus induce a lower level of perceptual load. Coupled closely with these first two factors is the level to which the objects must be processed in order to complete the task (Craik and Lockhart, 1972). For a display of different coloured shapes, for example, one might be required to identify whether there is any blue shape present ('pop out' or feature search; relatively low load) or whether there is a blue triangle present (conjunction search; higher load) – the former requires only low level features to be processed, while the latter requires the separate features of every object to be bound together (Treisman and Gelade, 1980), and is thus impacted both by the number of relevant features and the overall number of objects. Finally, the speed at which processing must occur can affect the perceptual load, with higher rates of information processing (for example a rapid serial presentation of objects, or a single display with a very short presentation time) inducing higher perceptual load.

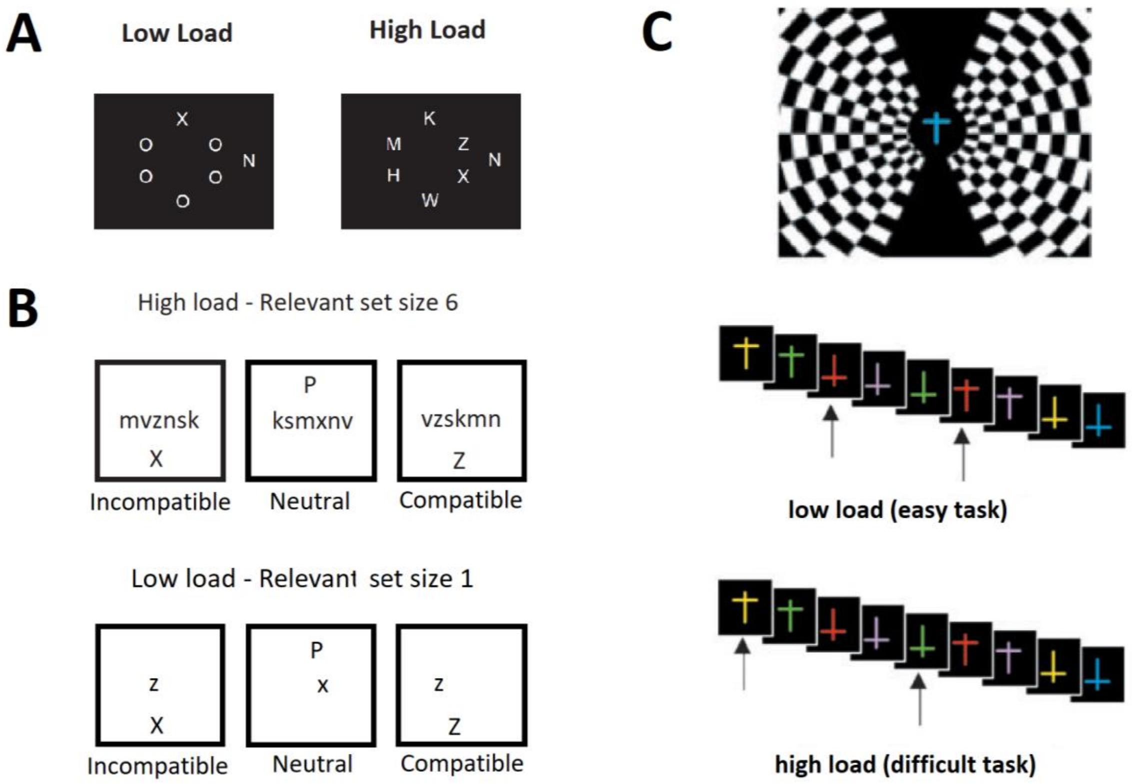


Figure 1.4: Examples of Perceptual Load Manipulations.

A. *Visual search task adapted from Lavie (2005). The task is to detect a target letter (X or N) within the circular array presented at fixation. Perceptual load is manipulated by varying the similarity between target and non-target objects. An irrelevant visual stimulus (N), which may be congruent or incongruent with the current target, is displayed peripherally. Both array examples contain a target of X, making the peripheral N stimulus incongruent in both cases.* ***B.*** *Visual search task adapted from Lavie (1995). The task is to detect a target letter (X or Z) within the string of letters presented horizontally across the centre of the display. Perceptual load is manipulated by the number of letters in the search array. An additional irrelevant stimulus that is compatible, incompatible, or neutral relative to the target is presented peripherally.* ***C.*** *RSVP task adapted from Schwartz et al. (2005). The task involves rapid foveal presentation of a series of coloured crosses which have an upright or inverted orientation. The low load task is to detect any red cross (feature search), while in high load targets are defined by a combination of features (yellow upright or green inverted) – conjunction search. Radial checkerboard stimuli are presented around the periphery, with black and white segments that are inverted at a regular rate (e.g., 8.5Hz).*

In experimental situations, perceptual load is often manipulated within a visual search task, where a number of objects are displayed simultaneously on screen for a brief duration (Lavie, 1995; Lavie and Cox, 1997; Lavie and Fox, 2000), or in a rapid serial visual presentation (RSVP), during which single objects are presented in series in the same location with a fixed presentation rate (Schwartz et al., 2005). For these tasks perceptual load can be manipulated based on the factors described above, for example increasing the number of items in the search display (e.g. Lavie, 1995; Figure 1.4B) contrasting a feature search (low load) with a conjunction search (high load) (e.g. Lavie and Cox, 1997; Schwartz et al., 2005; Figure 1.4C), changing the degree to which non-target objects resemble the target (few features in common – low load; many shared features – high load; Figure 1.4A) (Lavie and Cox, 1997; Macdonald and Lavie, 2008; Lavie et al., 2009), and is higher when stimuli are presented more briefly or at faster rates. Another task which has been shown to manipulate perceptual load is the multiple object tracking task, where participants must monitor the movement of a subset of target objects within a display of independently moving, visually identical objects (Jovicich et al., 2001; Eayrs and Lavie, 2018). For this task, both the number of target objects and the speed at which the objects move affect the perceptual load.

1.2.1.2 Cognitive Load

Load Theory proposes that another important factor during selective attention is cognitive control. It posits that during perception the cognitive system enforces a hierarchy of priorities on stimulus processing based on the top-down task set, so that relevant information is processed preferentially. Cognitive control may also play a role later in processing: if perceptual capacity exceeds the demands of the attended task and irrelevant stimuli are processed in addition to task-relevant stimuli, cognitive resources may also be involved in a late selection process to filter out the extraneous information so that it does not disrupt the conscious decision-making process.

In contrast to high perceptual load, which typically reduces distractor interference, if there is high load on cognitive resources, selective attention functions may be disrupted, and so irrelevant information is in fact more likely to be processed. This occurs in two ways: firstly, the hierarchy for allocation of perceptual resources may not be enforced appropriately, so that irrelevant information may be given the same weight as task-relevant information. In this situation, task-relevant stimuli may not be processed sufficiently to perform the task. Moreover, the irrelevant information is likely to have been more highly processed than desired. This compounds with the second factor, which is that the reduced availability of cognitive resources means the system is less able to filter out irrelevant information at a late stage. Together, these effects of high cognitive load mean that irrelevant information is more likely to capture attention, interfering with performance or even distracting from the attended task.

Cognitive load has typically been realised in experimental settings by placing demand on short term echoic memory, for instance requiring participants to hold a sequence of digits in mind while they perform another task (Lavie et al., 2004; Lavie and De Fockert, 2005), with a longer digit string inducing higher cognitive load. Another classic task uses a series of sequentially presented items, for which participants have to identify matches between the current item and one that was presented a set number of places earlier (n-back matching), with higher cognitive load occurring with larger n-back (Schulz et al., 2014). Interestingly, visual short-term memory tasks (e.g. Todd et al., 2005), lead to effects on distractor processing that align with perceptual, rather than cognitive load. This suggests that though maintaining information in auditory/echoic short-term memory acts as cognitive load, the process of visual short-term recall, which may rely on maintaining a visual image in mind, may engage perceptual processing mechanisms.

Though in experimental settings cognitive load is operationalised in ways that offer controlled manipulation of resources, in reality factors that place load on the cognitive system are likely to extend broadly to other aspects of executive function such as task-switching, remembering a complex rule set or high level decision making. Similarly, while it is important to be able to separately manipulate load on cognitive and perceptual resources experimentally in order to explore their effects independently, most 'real-world' tasks place load on both perceptual and cognitive systems, and thus are likely to induce a complex effect on selective attention.

1.2.1.3 Load as a tool to probe sensory processing during resource scarcity

One consequence of the automatic use of leftover resources is that it is not trivial to distinguish experimentally between processes which are truly automatic (i.e., do not require access to processing resources, and will occur without fail regardless of attentional state) and those which will occur by default when resources are available, but may still be disrupted if resources are sufficiently depleted. The approach that has been developed to examine whether a process is susceptible to capacity limits is a perceptual load paradigm. The aim is to control the perceptual resources available to the process of interest (the critical process) by focusing explicit attention on a concurrent competing task that depletes these resources to a greater or lesser degree depending on its perceptual load. A process that is truly automatic should proceed equally well regardless of the load of the attended task, while a process that occurs by default but is nonetheless dependent on resource availability will suffer as load is increased.

There are several considerations to be borne in mind when designing a perceptual load manipulation. The first is that perceptual and cognitive load have strikingly different effects on the processing of irrelevant stimuli: high *perceptual* load decreases distractor processing (as there is less surplus processing capacity) whereas high *cognitive* load increases both the processing of irrelevant information and the likelihood that it captures attention. It is therefore important that tasks manipulate perceptual demands (in order to control the resources available to the critical task) while keeping cognitive demands relatively low, so that processing priorities are maintained. It is also important that the perceptual load of the attended tasks is calibrated to levels of demand where there is likely to be a measurable effect on the processing of the critical stimuli – ideally the low load task should leave plenty of surplus capacity while the high load task should approach or exceed perceptual capacity.

Another aspect of the design that must be considered is how to obtain a measure of processing of critical stimuli that are (ideally) irrelevant from the participants' perspective. For behavioural experiments this leaves relatively few options; one is that participants should be engaged for several trials solely in the load task, with the critical stimulus presented unexpectedly during the final trial. Participants' percept is then probed (e.g., 'did you see anything unexpected in the last trial?), giving a measure of how well the irrelevant stimulus was processed. While this provides a strong control of attention, the paradigm is effectively limited to one critical trial per participant, and different levels of load must be compared between groups. Another option is that participants are asked to perform two tasks simultaneously ('dual task' paradigms), but are encouraged to prioritise the load task, while the critical task is secondary. Various strategies can be employed to help ensure that task priorities are maintained, such as requiring a speeded response to the primary task, or giving feedback or reward based on primary task, but the resource manipulation will necessarily be weaker than in a single task setting, since the critical stimuli are always part of the explicit task set rather than being irrelevant, and also because the coordination required to perform the dual task induces a level of cognitive load. Other paradigms use an indirect, objective measure of processing of the irrelevant stimuli, such as measuring the impact the irrelevant stimuli have on performance on the attended task (contrasting performance on trials when a critical stimulus is present versus absent), or by probing the degree to which irrelevant stimuli result in priming, for example.

For neural imaging studies a much more straightforward approach can be employed, since they provide a direct measure of processing. However, it is necessary to ensure that neural responses to the critical stimuli are separable from those to the load task stimuli, and that the neural response associated with the critical stimuli is sensitive enough to act as a quantitative index of processing. There are a number of ways in which responses to the load and critical stimuli can be separated. With fMRI data there is sufficient spatial accuracy that responses in two distinct modalities, or even stimuli presented to different locations in the visual field (such as left versus right hemisphere or foveal versus peripheral locations), can be distinguished based on the cortical region activated. Similarly, some brain regions respond to specific visual features (such as V5/MT which responds to movement) or objects (such as the parahippocampal place area, PPA, or the fusiform face gyrus, FFG, which respond to locations and faces respectively), so using critical stimuli that will activate these areas (with load stimuli that do not) allows for a direct index of critical stimulus processing. For EEG and MEG, it is possible to select different subsets of sensors to reflect different presentation modalities (for instance temporal sensors may reflect responses to auditory stimuli while occipital sensors reflect visual stimuli), but the fact that current spread occurs across the scalp means that there is some cross-contamination of activity. A more elegant approach relies on the fact that the high temporal resolution provided by M/EEG allows the recorded response to be deconvolved into component frequencies. It is therefore possible to 'tag' each stimulus by modulating it at a set frequency (e.g., flicker rates for visual stimuli, or amplitude modulation for auditory stimuli) so that the power at the corresponding frequency in the neural signal will give an index of processing. Another strategy is to present the critical stimuli on only half of trials, so that the signal that is specific to the critical stimuli can be identified by comparing neural responses in trials where it was present versus absent.

A final option for objective indices of perceptual processing is to use a physiological measure such as skin conductance or pupillometry. However, the same issue of ensuring the measure is sufficiently sensitive and specific to the critical stimuli applies, and there are fewer options for dissociating responses to multiple stimuli within one-dimensional data such as pupil diameter compared to the more rich, multi-channel datasets provided by neural imaging. Nonetheless, pupil diameter in particular has been found to be a very useful index of perceptual processing, and is one I will introduce in more detail in Chapter 5.

1.2.2 Evidence for Load Theory within vision

Load Theory has proven to be well supported by research within vision. The basic premise that processing of irrelevant stimuli (often termed 'distractors') depends on the perceptual load of the attended task has been corroborated by a wealth of behavioural studies. For example, in a task where participants must respond based on the identity of a target object, a distractor object whose identity is incongruent with the target may interfere with the response (response competition), leading to slower reaction times and higher error rates compared to when the distracter is congruent or neutral. This effect has been shown to be significantly reduced when the attended stimuli involve a higher level of perceptual load (Lavie, 1995; Lavie and Cox, 1997), supporting the prediction from Load Theory that processing of the irrelevant stimuli should be reduced under these conditions. Similar effects of reduced processing of irrelevant stimuli under high perceptual load have been shown using other objective measures such as priming (Lavie and Fox, 2000) and implicit learning of irrelevant stimuli (Jiang and Chun, 2001). These effects hold even for highly salient distractors such as faces (Jenkins et al., 2005).

This reduction in processing of irrelevant stimuli under high perceptual load can lead to substantial deficits in perception. A striking example is the phenomenon of inattentional blindness, where observers fail to notice stimuli that are readily perceptible under normal conditions because they are focused on a concurrent, demanding task (e.g. the 'Invisible Gorilla’ from Simons and Chabris, 1999). The same effect can be induced by manipulating the perceptual load of the attended task - a ‘load-induced blindness’ as illustrated by Cartwright Finch and Lavie (2007). This study used the ‘critical trial’ paradigm described in the previous section, where participants were engaged in either a low or high load perceptual decision task, with the irrelevant stimulus appearing only on a final trial. The results showed a strong effect of load on participants’ awareness of the critical stimulus, with over half of participants in the low load group reporting it compared to 10% in the high load group. The effect has since been replicated using a dual task paradigm, where participants simultaneously performed a central visual search task and reported the presence or absence of a peripheral stimulus that was not part of the search (Macdonald and Lavie, 2008). The data showed that even though participants were actively trying to detect the peripheral stimulus, high perceptual load in the visual search task led to a significant drop in detection sensitivity (d’).

There is also a substantial body of support for Load Theory from neural imaging studies, which illustrate that the level of perceptual load in an attended task is an important determinant of the magnitude of neural responses to task-irrelevant stimuli. The earliest demonstration was an fMRI study (Rees et al., 1997) in which a load task was presented centrally on screen along with irrelevant moving dots that were located peripherally. During high load blocks, activity in the region of the visual cortex which responds to motion (V5/MT) was significantly reduced relative to the low load blocks. A similar effect was shown for response to colour: activity in visual cortical area V4 elicited by irrelevant coloured images presented in the opposite visual hemisphere to the attended (monochrome) task was reduced as the perceptual load of the attended task increased (Pinsk et al., 2004). Yi and colleagues (2004) illustrated that high perceptual load can also reduce repetition suppression, an effect where neural responses to stimuli that are presented multiple times are reduced in amplitude on repeated presentation relative to the first. They presented scenes that typically elicit activity in a region of the cortex termed the parahippocampal place area (PPA), as irrelevant backgrounds surrounding a central attended task. They showed the expected effect that high perceptual load in the attended task reduced PPA activation compared to low load conditions, but also found that for background scenes that were repeated within the experiment, the typical reduction in activation on second presentation of an image (repetition suppression) was seen only under low load conditions.

More recent work has focused on determining how early in the visual pathway the effects of load can be seen. Schwartz et al. (2005) showed that load effects can be seen as early as V1 – the first cortical region to respond to visual stimuli. They recorded responses to flashing checkerboard patterns (a stimulus which is known to elicit activity from the earliest stages of the visual processing hierarchy) presented around the periphery of the visual field, with a centrally positioned series of coloured crosses presented in RSVP, where the task was either a feature search within the RSVP stimuli (colour; low load) or conjunction search (colour and orientation; high load. See Figure 1.4C). They showed significant effects of perceptual load on responses to the irrelevant stimuli throughout V1-V4, with the effect becoming more pronounced at each successive stage. A study that used a similar experimental paradigm but with a focus on recording sub-cortical activity provided data that suggests that the effect of load may be apparent even as early as the Lateral Geniculate Nucleus (LGN) within the thalamus (O'Connor et al., 2002).

The effect of load has also been explored using electrophysiological recordings. Handy et al. (2001) engaged participants in a response competition paradigm, with a central discrimination task of high or low load, and peripheral distracters that were congruent or incongruent with the target. On occasional ‘probe trials’ they presented only a neutral distractor, with no foveal task. Behaviourally, the results within the task trials were consistent with previous reports of response competition under load. ERPs from the probe trials were analysed to determine whether the load of the task in the current block had an impact on the neural responses to peripheral stimuli – though note that in order to ensure that the ERPs measured corresponded only to the distracter stimuli, there was no active task during the probe trials, merely the expectation of one. The results showed that in high load blocks the P1 response to the irrelevant stimuli was significantly reduced relative to low load blocks. This suggests that there may be a pre-emptive dampening of responses to visual stimuli at unattended locations, that is stronger if the current task is of high perceptual demand. There is also ERP evidence of load affecting the processing of concurrent, irrelevant stimuli. Parks et al. (2011; 2013) presented the cross colour-orientation RSVP load task (as used by Schwartz et al., 2005; Figure 1.4C) centrally, with an irrelevant checkerboard pattern that inverted (white areas turning black and vice versa) at a set rate of 8.5Hz, presented around the periphery. The recorded EEG signal was transformed into the frequency domain, where the power at 8.5Hz was used as an index of the degree of processing the irrelevant checkerboard stimuli received. The results indicated a significant reduction in the processing of the irrelevant stimulus under high compared to low perceptual load. There is also evidence from ERP studies that the modulation by perceptual load of neural responses to irrelevant visual stimuli can occur relatively early in the processing hierarchy; Rauss et al. (2009) presented an RSVP task similar to those described above, with occasional peripheral distractors presented at jittered delays (250-500 ms) relative to the RSVP stream. Analysis of the evoked responses to the irrelevant stimuli focused on the early C1 component (Kelly et al., 2008) that occurs at 70-80 ms in response to non-foveal stimuli, and is understood to reflect the afferent projections into the earliest cortical area, V1. The data indicated that the amplitude of this response was modulated according to the load of the attended, foveal task, corroborating the evidence from fMRI studies that suggests the locus of the load effect includes early cortical or even sub-cortical visual responses.

1.3 PERCEPTUAL LOAD AND AUDITORY PROCESSING

While Load Theory has received substantial empirical support from studies of visual processing, the question of whether it applies to audition is less clear. If perceptual resources are limited, it makes sense from an evolutionary perspective to allow any excess capacity that is not required for active tasks to be used to monitor the environment in a manner similar to the proposed early warning system provided by the auditory system. However, if resources are indeed limited, the system would be susceptible to disruption during situations of high perceptual demand, when potential dangers could be missed.

As I established in section 1.1, the degree to which unattended auditory stimuli are processed has been a matter of some debate. Viewing the previous body of research with Load Theory in mind, it becomes clear why the classic attention paradigms used in these studies may not have provided a consensus: paradigms such as ‘attend versus ignore’ and ‘active versus passive’ are not sufficient to disentangle separate effects of active attention and passive use of spare capacity without explicit intent. If resources allocation within auditory processing follows a system akin to that established for vision with Load Theory, the lack of control of the perceptual demand of the attended ‘distractor’ tasks could lead to a large disparity in the degree to which unattended stimuli are processed, resulting in sophisticated scene analysis in some cases, with only basic auditory features processed in other cases.

While applying Load Theory to auditory processing has exciting potential, the efforts thus far have been somewhat curtailed due to a lack of consensus on how perceptual load should be operationally defined for auditory stimuli. Many studies have taken the approach of trying to use a direct analogue of tasks used to modulate visual perceptual load, in spite of some fundamental differences between the visual and auditory processing pathways – in particular the retinotopic mapping that is characteristic of the ascending visual pathway, compared to the tonotopically organised auditory system. For instance, Rees, Frith and Lavie (2001) designed an auditory load experiment that paralleled the seminal 1997 fMRI study in which they recorded responses to the motion of irrelevant background objects while participants engaged in a task based on words presented on screen. In the visual version of the task, the written words had to be categorised either by the letter case (lower or upper

– low load) or the number of syllables (high load). For auditory task in the follow-up study, participants heard spoken words which they had to categorise based either on the overall intensity (low load) or the number of syllables (high load). However, while in vision the syllable task involved binding the low-level features of each letter together in order to identify it, reading the word as a whole, and determining how many syllables it contains when pronounced, the auditory version of the syllable task is perceptually trivial, with a simple assessment of the volume envelope of the stimulus providing the response cue. Perhaps unsurprisingly, then, the results did not show an effect of the auditory task manipulation on the degree to which the motion of the irrelevant visual stimuli was encoded. Though many other studies have since made efforts to manipulate auditory perceptual load, the large disparity in methods and results means that no consensus has been reached (see Murphy et al., 2017 for a comprehensive review).

In the cases where there was no significant impact of the task manipulation on processing of irrelevant stimuli it is hard to know whether this is a true null effect or a consequence of a poor load manipulation. However, some of these studies have shown effects of the task manipulation that may provide some insight into whether Load Theory applies in audition, and so I will review the results in the following section.

1.3.1 Auditory processing under auditory load

Alain and Izenberg (2003) used a load manipulation to explore whether cues used for sound segregation were impacted as perceptual resources were depleted by a concurrent task. They manipulated auditory perceptual load using a task based on complex tones (i.e., tones that are formed of multiple harmonic components, rather than a single frequency). These tones were presented in sequence to one ear and participants monitored the stream for occasional deviants that were shorter in duration. For the low load task, they simply had to respond to indicate the deviant duration, while in the high load task they also had to categorise the tone as tuned or mistuned, depending on whether it contained a harmonic component that was shifted in frequency (see section 1.1.1, Figure 1.2). An irrelevant sequence of tones containing the same occasional deviant stimuli was presented concurrently to the unattended ear. The neural responses to the unattended ear illustrated a reduction in the amplitude of the MMN response to the duration deviants during the high relative to low load task. However, an ERP component referred to as the object-related negativity (ORN, discussed in further detail in Chapters 3 and 4) that occurs in response to the mistuning that was present in some of the deviant stimuli, was unaffected by task load. The authors interpret the results as suggesting that segregation of sounds presented concurrently based on harmonicity cues is a process that is truly attention-independent, while segregation of sound sequences based on cues like the MMN may be susceptible to resource limits.

Sabri and colleagues (2006) designed a task to see whether auditory attention capture was affected by load. Participants were engaged in a duration discrimination task that was either demanding (50 ms vs 60 ms) or easy (50 ms vs 100 ms), and which contained occasional task-irrelevant frequency deviants. Behaviourally, these frequency deviants resulted in attention capture, with performance on the task being poorer for trials where the irrelevant deviant had occurred. The MMN response to the frequency deviants was found to be reduced in the high load condition, suggesting that the processing of this irrelevant stimulus feature was reduced when the demands of the task were higher. However, there was a larger P3 response to the deviants during performance of the more demanding task than the easier task, suggesting that the irrelevant feature was more likely to result in attention capture in those conditions. These results may reflect the fact that the task used did not purely load perceptual resources – within vision it has been shown that degrading perceptual discriminability may lead to performance decrements due to the fact that there is limited data available for the perceptual decision rather than because more resources are required to process it (Norman and Bobrow, 1975). Thus, while difficult perceptual discriminations may lead to some increase in perceptual resource allocation in order to ensure the information gathered is as precise as possible, it is cognitive resources that are more heavily loaded when a decision must be made based on ambiguous data (Lavie and De Fockert, 2003). If both perceptual and cognitive resources were loaded more in the demanding task used by Sabri et al., there may have been both a reduction in the degree to which the irrelevant feature was encoded (leading to a smaller, but still existent, MMN), and a lapse in cognitive control such that the irrelevant deviant was more likely to capture attention.

Chait et al. (2012) investigated the effect of perceptual load on auditory change detection. The auditory load task, presented to one ear, was based on a rapid serial presentation of sounds, which participants had to monitor for targets that were either perceptually highly dissimilar to non-targets (wide-band noise bursts; low load), or one of a set of 'auditory shapes' which were only distinguishable from non-targets after combining features such as pitch and intensity profiles (conjunction search; high load).

The change detection stimuli presented to the unattended ear consisted of rapid sequences of tone pips that contained transitions between a state where all the tones were of a fixed frequency (‘constant’ state) and a state where the tone frequencies changed randomly (‘random’ state). The transition from a constant to random state, where the first tone of a different frequency represents the clear violation of the simple established rule, is inherently different from the transition from a random to constant state, where the evidence for the transition must build up to the point where it seems unlikely that the frequency has remained unchanged for this period by chance. The experiment showed that the neural response to transitions from constant to random states were unaffected by the load in the attended task, while shifts from random to constant states showed reduced neural responses under high load. These findings suggest that neural responses to a simple rule violation was less susceptible to resource depletion than the process by which evidence for a new rule built up. This is in keeping with the theory by Sussman (2007) based on MMN evidence that when attention is available it allows more sophisticated rule-formation within the model of the standard.

Interestingly, Alain and Izenberg’s (2003) data suggest that MMNs to a feature deviant (in their case duration) can be affected by the depletion of resources, while there was no effect of the load manipulation in the Chait et al. (2012) study on the response to the constant to random transition that was essentially a frequency deviant occurring after a series of standards. This suggests that the high load task used by Chait et al. depleted resources sufficiently that the more complex process of detecting the emergence of a regular pattern was impacted, but not enough to impair the more basic deviance detection process. Thus, it is possible that with a more demanding load task the constant to random transition response might also have been affected.

1.3.2 Limitations of auditory load manipulation

The studies described above begin to form a picture of auditory processing where many processes such as scene analysis, rule-forming and prediction occur by default, but with varying degrees of complexity depending on the processing resources available. There is also some evidence that in cases of extreme resource depletion, even the basic levels of these processes may fail to execute. However, due to the paucity of data, lack of consensus on how to manipulate auditory load, and a significant number of null results, the studies based on auditory load manipulations have not provided strong evidence as to the degree to which auditory processing is dependent on the availability of perceptual resources.

Another important factor limiting the conclusions that can be drawn from this line of research is that when the attended task and irrelevant stimuli are presented within the same modality, it is difficult to firmly rule out accounts based on low-level stimulus competition. For visual processing, Tsal and Benoni (2010) proposed that the spatial proximity of stimuli in a visual display is the critical factor in determining how much the critical stimulus is 'diluted' by the other stimuli, rather than perceptual load, per se. There is also some indication from retinotopic mapping studies that the degree to which visual distractor processing is reduced under high perceptual load is dependent on how close it is to the foveal task (Parks et al., 2013). While there are plausible explanations for these findings within a Load Theory framework (see, for example, Lavie and Torralbo, 2010), it is not possible to dismiss that many low-level interactions between visual stimuli do occur.

In audition there are also a number of well-known low-level stimulus interactions whereby a signal of interest can be masked by energy from a different source at the level of the cochlea if they occur within similar frequency bands and temporal windows (energetic masking), and these effects are compounded by the fact that low-level auditory signal processing is highly non-linear. There may also be interference at a higher level based on which auditory features are relevant to the attended task-set. For example, in the cases where the MMN to feature deviants have been abolished (e.g. Woldorff et al., 1991), the effect apparently occurs only for the feature on which the active task is based: Sussman et al. (2003) showed that for an active task where deviants within a certain feature (e.g. frequency) were to be detected, responses to feature deviants in the unattended ear were abolished for deviants along the task-relevant feature, but not for other low-level features (e.g. intensity).

A simple way to negate any potential confounds of stimulus interference is to use a cross-modal design, where the attended load task is in a different modality to the critical stimuli. For exploring the degree to which *auditory* processes may be susceptible to resource availability, the natural option for the other modality is vision, in which a number of tasks have been firmly established to control the depletion of perceptual resources. This approach also offers the opportunity to explore whether any limits on perceptual resources are modality-specific (i.e., separate resource pools for visual and auditory processing) or whether sensory processes may be subject to limits on resources that are shared across modalities.

1.3.3 Applying visual perceptual load to auditory processing

To date there have been a small number of studies that have explored the effect that placing perceptual load on the visual system has on auditory processing.

Based on the phenomenon within vision of inattentional blindness (e.g. Cartwright-Finch and Lavie, 2007), Macdonald and Lavie (2011) designed a study to explore whether high levels of visual load might result in failure to perceive auditory stimuli. They used a critical trial design similar to that used by Cartwright-Finch and Lavie, where participants engaged in a visual search task for multiple trials and the irrelevant stimulus was presented unexpectedly on the final trial. The effect on perception of irrelevant auditory stimuli was even more striking than that for irrelevant visual stimuli, with detection of the unexpected sound being reduced from 80% to 20% under low and high visual load, respectively. The effect has since been replicated using dual-task conditions (Raveh and Lavie, 2015): participants simultaneously performed a primary visual search task of high or low perceptual load, and a secondary auditory detection task based on tones that were quiet (30 dB SPL) but easily audible to participants during single-task conditions. The results showed that high perceptual load in the visual task led to a significant reduction in detection sensitivity (d’) for the auditory target. Analysis of participants’ response criterion (β) showed that the effect was not simply due to participants responding more cautiously when under high load, but reflected a difference in perception. However, it is unclear from these behavioural studies which stage of the auditory processing hierarchy is affected by visual load – it may be that sensory input was reduced at early levels of the cortex, but could equally have occurred at a later stage, such as when the auditory signal reaches conscious perception.

A pair of studies by Parks et al. (2009, 2011) explored this question using electrophysiology, presenting irrelevant auditory signals alongside a visual load task. The perceptual load manipulation was effected using the well-established feature versus conjunction search within an RSVP of coloured crosses (Figure 1.4C). In the first paper (Parks et al., 2009) the measure for auditory processing was a microreflex called the post-auricular reflex (PAR). The PAR occurs in response to abrupt sound onsets and is mediated within the brainstem and had previously not been found to be modulated by attentional influences. The auditory stimuli used to evoke the PAR were extremely brief (5 ms) wide-band noise bursts that create a sharp auditory 'edge', which were presented at regular intervals throughout each RSVP trial. The second experiment (Parks et al., 2011) used the same visual task but instead used EEG to record steady state responses (SSR) to a frequency-tagged irrelevant pure tone that was presented throughout each trial. The response to the tone within the composite EEG signal was identified by measuring the signal power at the frequency which matched the rate of the amplitude modulation that had been applied to the tone. The SSR recorded in this way represents cortical sensory processing and is therefore higher in the processing hierarchy than the presumed generator of the PAR measured in the 2009 study. Somewhat surprisingly, the results of the studies showed that high perceptual load in the visual task had a significant effect on the amplitude of the PAR, but not on the SSR response. The authors propose that this may indicate that the PAR, while generated within the brainstem, may be modulated by backward projections from higher cortical areas that do show modulation under visual load, where the early cortical responses to sound do not. However, it is worth noting that for the SSR study, the auditory signals were present constantly throughout the RSVP trials (~5 seconds each), while the stimuli for the PAR study were extremely brief. Thus, it is possible that there was more opportunity to process the auditory SSR signals during periods when the load in the RSVP task happened to be momentarily lower (e.g., in the periods towards the end of each object's presentation). Regardless of the explanation, the fact that these two studies have divergent results means that it remains unclear whether visual load impacts the amplitude of evoked activity during early sensory processing or at a later stage.

Visual load tasks have also been applied to study the degree to which higher level auditory processes may rely on resource availability. (Dyson et al., 2005) ran an experiment to parallel the auditory load experiment by Alain and Izenberg (2003) described previously, in which the process of auditory segregation based on harmonicity and pitch cues was explored within the irrelevant auditory stimuli. The critical auditory stimuli used in both experiments were sequences of tones that contained occasional features that act as cues for segregation – a mistuned component within a harmonic complex, which elicits the ORN response, and a deviant based on the pitch within sequentially presented sounds, which elicits an MMN. The visual load task introduced by Dyson et al. was based on a serial presentation of coloured digits. The low load task was to classify each digit as greater than or less than five regardless of colour, while in the high load task, the response of > or < 5 was based on the current digit if it was displayed in blue, or the previous digit if the currently displayed digit was red. In contrast to the effect of auditory task load found by Alain and Izenberg, where MMN but not ORN responses were modulated by load, there was no effect of visual task load on either response to the auditory signals. However, the visual task used was not a classic manipulation of perceptual load; in fact, that requirement in the high load task to maintain the previous digit in mind is more in line with n-back tasks that are typically used to manipulate cognitive load, and regardless of which system is loaded 1-back tasks are not considered to represent high demand. Thus, it seems likely that the visual task did not function as a strong manipulation of the perceptual resources left over for auditory processing.

A visual load paradigm has also been applied in a study of auditory change detection. The study by Chait and colleagues (2012) described in Section 1.3.1 , where an auditory search task presented to one ear was used to explore load effects on unattended signals in the other ear which transitioned from constant to regular pitch patterns or vice-versa. The visual task condition involved participants performing a version of an RSVP feature-conjunction search, with coloured shapes being presented in sequence and targets being determined by colour for the low load task and a colour-shape conjunction search for high load. Though the study illustrated a significant effect of auditory perceptual load on responses to the irrelevant auditory change signals (as described previously) there was not a significant effect of visual perceptual load. The authors interpret this as signifying that load within the auditory system can affect processing of irrelevant sounds, but that loading the visual system does not place a drain on the resources used for auditory processing – i.e., that both visual and auditory perception are subject to resource limitations of the type described by Load Theory, but that the resource pools are not shared between sensory modalities. However, it is also possible that the visual RSVP task was a weaker manipulation of perceptual load than the auditory task – even in the high load version of the visual task, 67% of the stimuli could be discounted as targets based on their colour, meaning that the further operation of orientation detection was not necessary, making the high load task functionally equivalent to the low load task for a significant portion of the trials.

Based on these studies, it is currently unclear whether visual perceptual load can modulate the amplitude of responses within the auditory system, or impact the efficacy of auditory processing. One of the aspects of the visual load tasks used in these studies which may have contributed to null findings is that most used a visual task where load is presumed to be induced over an extended time period, but may in reality fluctuate as processing demands wax and wane, as one might expect during serial presentation. If resource allocation occurs in a dynamic manner, for example due to the automatic use of spare capacity as proposed by the Load model, these brief periods when load is reduced may allow processing of irrelevant auditory signals. While this has not been considered an issue in studies which used visual distracters, the auditory system is notoriously adept at processing information based on brief temporal snapshots or ‘glimpses’, such as the process at play when an auditory signal of interest surfaces briefly above a dynamic noise background – so called ‘dip listening’ (Cooke, 2006). In contrast to the RSVP paradigms, visual search tasks, which consist of multi-element arrays presented only for a brief duration (100-200 ms), may provide a stronger control of perceptual resources. In these paradigms the timing of the irrelevant auditory signals is matched precisely with the load task, which maximises the potential for competition for resources. Indeed, the two studies which used visual search (Macdonald and Lavie, 2011; Raveh and Lavie, 2015) both showed significant effects of load on auditory perception. However, this was possible only because the auditory signals used were very short; one reason that other studies have favoured longer trials is that many auditory processes rely on the integration of sensory information over time, so restricting the duration of a load task may rule out its use to study the process of interest. Another factor, particularly for electrophysiological studies, is that allowing temporal asynchrony between task stimuli and irrelevant stimuli provides a straightforward method by which to separate evoked responses to each stimulus, by time-locking the recorded signal relative to onset of each stimulus. For stimuli that are presented concurrently, separating response components that were elicited by different stimuli must rely on more complicated strategies.

1.4 SUMMARY

In summary, many fundamental aspects of auditory processing occur without explicit intent, leading to the assumption that they are ‘automatic’ or ‘pre-attentive’. However, models of attention and perceptual processing highlight that even processes that occur by default may require access to perceptual resources and can therefore be susceptible to failure when a task that is the explicit focus of attention exhausts the limited perceptual capacity. Thus, the classic paradigms employed in auditory attention research are not sufficient to distinguish between a process which is truly automatic (i.e., will occur regardless of any competing demands on the sensory system) and one that occurs passively but is dependent on resource-availability. An approach that has proved useful within vision is to control the level of resources that are available to the critical process of interest by studying it in a situation where perceptual capacity is being depleted by a separate, attended task. If the critical process is affected by manipulating the perceptual load of the attended task, this suggests that the critical task is subject to the availability of processing resources; in contrast a process that is automatic should not be affected by the level of load in the attended task.

This ‘load task’ approach has been firmly established within vision, but applying the paradigm to explore which auditory processes may be susceptible to capacity limits has had mixed results. For studies which aimed to manipulate load within an auditory task, a lack of consensus on how load should be defined for auditory stimuli has led to a range of approaches from which it is difficult to draw firm conclusions. Within-modality manipulations of load also present a potential confound in that there may be low-level stimulus interactions that affects results. To avoid these confounds, and to establish whether limits on resources may apply across sensory modalities, some studies have used visual load tasks to explore auditory processing under load. However, these have favoured the use of RSVP load tasks in which stimuli are presented over extended timeperiods. It is likely that within these tasks there are fluctuations in the real-time demands and resource availability, which may allow processing of irrelevant stimuli to occur even during tasks where the overall load is high. A more controlled approach may be to use brief load tasks where the timing of the auditory signals is matched directly to the load task, ensuring that the potential for resource depletion during auditory perception is maximised. However, limiting the duration of the stimuli places restrictions on the auditory processes that can be explored, and matching the timing of two stimuli makes it difficult to distinguish their separate neural responses from the aggregate recorded signal.

1.5 THESIS OUTLINE

In the experiments presented in this thesis, I use MEG, pupillometry a behavioural dual-task designs to explore how auditory processing is impacted by visual perceptual load.

Chapter 2 explores how neural responses to simple auditory stimuli are affected during load-induced inattentional deafness, with a view to establishing which stages in the auditory processing hierarchy show evidence of modulation by visual load. For the experiments in this chapter, I used the visual search task that has been shown previously to induce inattentional deafness (Macdonald and Lavie, 2011; Raveh and Lavie, 2015), presented concurrently with irrelevant, near-threshold pure-tones. The first experiment is an MEG study in which participants engaged in the load task and auditory signals were irrelevant. MEG data illustrate an effect of load on the evoked response to sound during both early cortical sensory processing (~100ms post stimulus onset) and during later awareness-related responses. The second experiment is a complementary dual-task behavioural study in which participants detected the auditory signals in a concurrent but deprioritised task alongside the visual search task, to establish that the task parameters used for the MEG study did indeed induce an inattentional deafness effect behaviourally.

Over the next two chapters I focus on testing whether high visual load can affect the computational capacity of the auditory system. In the first set of experiments, presented in **Chapter 3**, I review models of auditory segregation and select temporal coherence-based segregation as a process that is representative of computation within the auditory system. I then establish signals that can be used to probe this segregation mechanism within very brief stimuli, so that they are compatible with brief visual load tasks such as the one established to modulate auditory sensory responses in Chapter 1. I present MEG signals recorded while participants listened passively to these stimuli in order to characterise a neural response that gives a sensitive measure for the degree of segregation that has occurred.

In **Chapter 4** I couple these auditory signals with brief visual search tasks in order to determine whether auditory segregation mechanisms are impacted by limiting perceptual resource availability. The first experiment presents MEG data for responses to irrelevant segregation stimuli recorded while participants performed the visual search task used in Chapter 1. The second experiment replicates the findings of the first experiment while using a novel visual search task designed to manipulate perceptual load without any change between the physical stimuli used for low and high load tasks. Finally, a behavioural dual-task experiment that parallels the task design of the second MEG experiment is presented to establish any parallel effects on perception.

In **Chapter 5** I present pupillometry data collected during the MEG load experiments from Chapters 1-3. For each experiment pupillary responses are separated into responses to the visual load tasks and to the auditory signals, and I quantify the effects of load on the pupillary measures of both auditory and visual processing for each of the three datasets. Within the three study designs there are two manipulations of visual perceptual load, and two auditory processes, so comparing the effects between the three experiments provides a replication and a contrast for both the visual task and the auditory processes.

In **Chapter 6** I discuss how the data I have presented further our understanding of ‘passive’ auditory processing and that it is in fact affected by the availability of attention resources. I argue that the classic paradigms used to explore auditory attention are not sufficiently well controlled, and therefore that many aspects of auditory processing which have been broadly accepted to occur ‘automatically’ may need to be re-examined under more stringent experimental conditions. I discuss how the work I have presented highlights the need for a concrete method to quantify perceptual load, particularly within audition, and that along with neural imaging methods, pupillometry may provide an efficient way to index perceptual processing under load. Finally, I reiterate how the cross-modal design used in this thesis provides important insight into the mechanisms of selective attention, and the process by which resources are allocated between sensory systems.

CHAPTER 2: LOAD-INDUCED INATTENTIONAL DEAFNESS[1]

2.0 MOTIVATION

The phenomenon of inattentional deafness has been illustrated using behavioural measures which show reduced awareness of quiet auditory stimuli that are presented concurrently with a visual task of high relative to low perceptual load. However, it is not clear whether these effects on awareness reflect an impact of visual load on sensory auditory processing or on later processes related to conscious perception. To address this question, I used MEG to record evoked responses to irrelevant sounds that were presented while participants performed a concurrent task of low or high visual load. The timing of any load effect on the ERFs, along with source localisation of the effect of load on auditory processing should provide an insight into the stage(s) in the auditory processing hierarchy where visual perceptual load can impact activity.

2.1 INTRODUCTION

Much research has shown that the level of perceptual load in an attended task is an important determinant of the magnitude of neural responses to task-irrelevant stimuli. Tasks of high perceptual load (e.g. visual search in multi-element arrays) are associated with significantly reduced visual cortex response to task-irrelevant stimuli that can elicit a robust response in conditions of low load (e.g. search for a feature

---

[1] Data from this chapter have been published in: Molloy, K., Griffiths, T. D., Chait, M., & Lavie, N. (2015). Inattentional deafness: visual load leads to time-specific suppression of auditory evoked responses. *Journal of Neuroscience*, *35*(49), 16046-16054. Although I have rewritten some of the text so that it works in concert with the other chapters in this thesis, there may be some repetition of information.

“pop-out”) (Rees et al., 1997; Handy et al., 2001; Yi et al., 2004; Schwartz et al., 2005). These findings are explained with the limited capacity model offered in the Load theory of attention (Lavie, 1995, 2005, 2010). Load modulations have been found throughout visual cortex and in some subcortical structures, from the lateral geniculate nucleus and superior colliculus (Rees et al., 1997; O'Connor et al., 2002), through striate and extrastriate cortex (V1-V4 and MT: Rees et al., 1997; Schwartz et al., 2005) to category-selective regions that respond to meaningful stimuli (TEO: Pinsk et al., 2004; PPA; Yi et al., 2004). ERP studies confirm that these effects are apparent in early cortical evoked responses 70-80 ms post stimulus onset (Rauss et al., 2009), and throughout later peaks and frequency locked responses (Handy et al., 2001; Parks et al., 2011; Parks et al., 2013).

Load-induced modulation of visual responses leads to the phenomenon of 'inattentional blindness', where observers fail to notice unattended stimuli if they are presented during task conditions of high perceptual load (e.g. Cartwright-Finch and Lavie, 2007). This occurs even when subjects are instructed to detect any additional stimuli beyond the task set (Macdonald and Lavie, 2008; Carmel et al., 2011). The inherent limits in perceptual capacity underlying these behavioural effects are believed to relate to a central limited attention resource (Rees et al., 1997; Dehaene and Changeux, 2011), and this raises the possibility of a cross-modal perceptual load effect, whereby increased perceptual load in a visual task may result in reduced auditory cortex responses to task-unrelated auditory stimuli. Consistent with this premise, recent behavioural findings demonstrated reduced detection sensitivity for auditory tones presented during visual tasks of high perceptual load (load-induced inattentional deafness, Macdonald and Lavie, 2011; Raveh and Lavie, 2015). However, the neural underpinnings of these effects remain as yet unknown. Specifically, it is not understood whether inattentional deafness results from a modulation of early sensory (auditory cortical) activity, or of processing at a later stage such as when information either reaches subjective awareness or is filtered out because it is not task relevant.

In the first experiment presented in this chapter, I used MEG recordings to explore the temporal and cortical locus of load-induced inattentional deafness. Participants performed a visual search task that involved different levels of perceptual load, with irrelevant pure tones presented at near-threshold intensities on half of trials. In order to maximise the potential impact of the drain on perceptual resources by the visual task (see comments in section 1.3.3) auditory stimuli occurred at precisely the same time as the visual task stimuli, and auditory and visual neural responses were separated during data processing based on a PCA-like analysis that is described in detail in the methods. The second experiment is a behavioural study which used an identical stimulus design to the MEG experiment, but where participants concurrently performed both the visual load task and attempted to detect the auditory signals.

2.2 EXPERIMENT 1: LOAD INDUCED DEAFNESS: MEG

2.2.1 Methods

2.2.1.1 Participants

Fourteen paid participants (8 female; mean age of 28.3 years, SD = 4.5 years) took part in the main MEG experiment. Nine further participants (4 female, a mean age of 28.7 years, SD = 4.1) participated in a passive viewing control condition (see sections '2.2.1.3 Stimuli' and '2.2.1.5 Passive Group' for details). One of the participants from the main experiment was excluded from the analysis because they showed a poor neural response to the auditory stimuli when they were presented in isolation. All participants were right handed, had normal or corrected to normal vision and reported normal hearing and no history of neurological disorders. The experimental protocol for this and all subsequent experiments presented in this thesis were approved by the University College London research ethics committee.

2.2.1.2 Apparatus

The experiment was run using Matlab 7.12 and Cogent 2000 (http://www.vislab.ucl.ac.uk/cogent.php). The magnetic signals were recorded continuously (600Hz sampling rate; 100Hz hardware low-pass filter) using a CTF-275 MEG system (axial gradiometers, 274 channels, 30 reference channels, VSM MedTech, Canada) in a magnetically shielded room. Subjects were seated in an upright position, with the visual stimuli projected onto a screen placed ~52cm from the participants' eyes. Sounds were presented via tubephones (E-A-RTONE 3A 10 Ω, Etymotic Research, Inc.) inserted into the ear-canal. Presentation latencies for auditory (~15 ms) and visual (~17 ms) stimuli were measured using a microphone and photodiode, and data were adjusted so that the evoked responses reported are appropriately aligned to the onset of stimulus presentation.

2.2.1.3 Stimuli

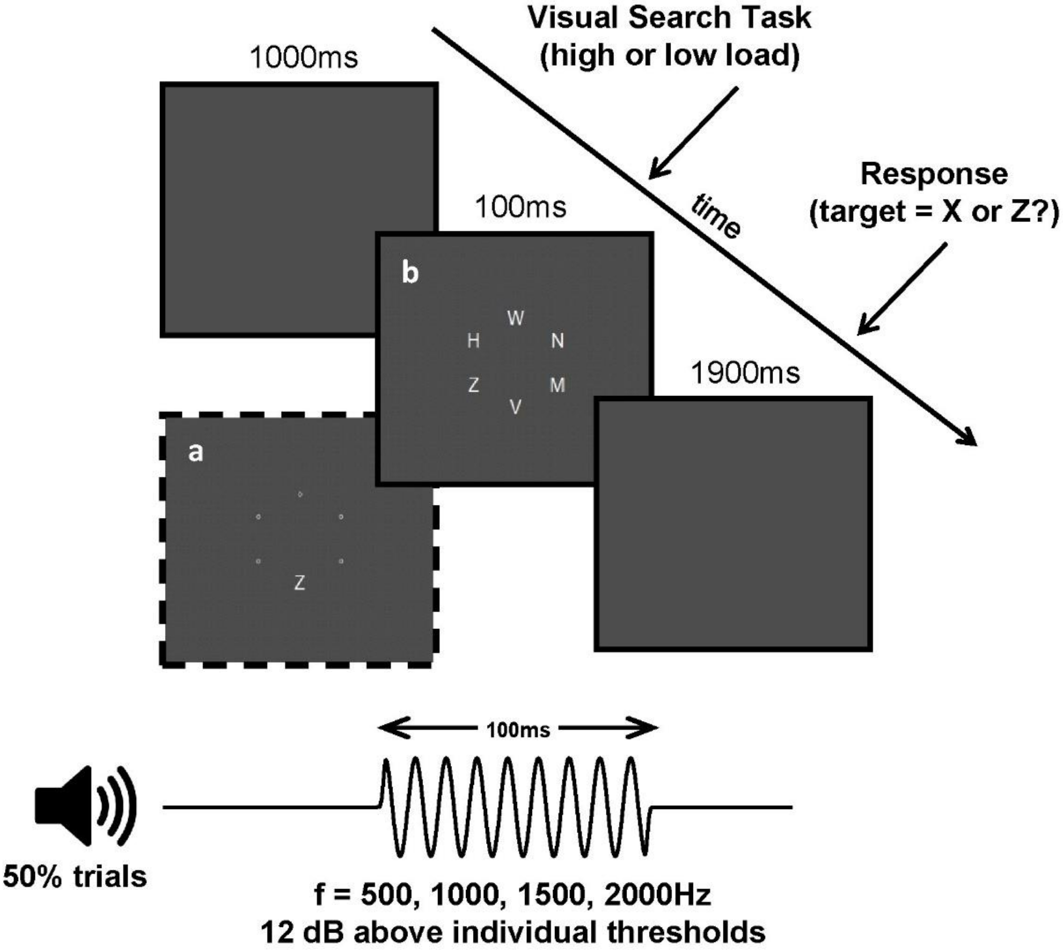


Figure 2.1: Experiment 1 - Load Task Paradigm

Letter search task of low load (a; target dissimilar from distractors) or high load (b; target similar to distractors). Subjects were instructed to identify which of two target letters (X or Z) had been present in the display. Search arrays were presented for 100 ms. Auditory stimuli (which occurred on 50% of trials; no auditory task) were presented at precisely the same time as the visual stimuli, covering the 100 ms during which the search array was displayed on screen.

The visual task consisted of six items spaced equally around a circle centred at fixation and subtending a 1.9° viewing angle (Figure 2.1). The background of the display was dark grey (RGB: 77 77 77); the letters and fixation cross appeared in white. The target letters were X or Z, in equal proportion, both measuring 0.6° x 0.6°. The non-target letters in the high load condition were the letters K, W, V, N and M (all the same size as the target letters), and those in the low load condition were smaller Os (0.2° x 0.2°) as used in previous load research (Lavie and Cox, 1997; Forster and Lavie, 2008; Raveh and Lavie, 2015) to maximise the difference in search load. The positions of the letters were randomised on each trial so that the target had an equal probability of occurring in each position. The low load non-target letters were easily distinguishable from the targets based on low-level visual features such as line curvature and orientation, resulting in target pop-out. For the high load condition, to distinguish non-targets from targets requires binding together the low-level features of line orientation with spatial location. This conjunction search involves more visual processing than the basic feature search (Treisman and Gelade, 1980) and therefore induces a higher visual perceptual load. A passive viewing control condition (see section 2.2.1.5 Passive Group below) was included to verify that the slight difference between the visual displays in the high- and low- load conditions, did not mediate the load effect. If load is the critical factor, MEG responses should not differ between displays in the passive viewing condition.

The auditory stimuli were 100 ms long, diotically presented, pure tones with frequency of 500, 1000, 1500 or 2000Hz and an envelope shaped by 10 ms raised cosine ramps. The intensity of the tones was adjusted for each participant individually (see section 2.2.1.4 Procedure). During the main experiment, tones were presented randomly on 50% of trials. Overall, equal proportions of each of the 4 possible frequencies were used, and tones were selected randomly on each trial to discourage participants from narrowing their auditory attention to a certain frequency band.

2.2.1.4 Procedure

Thresholds for each of the pure tone frequencies were determined for each participant at the beginning of the experiment using an adaptive staircase procedure. The staircase gave an estimate of the 79.4% correct point on the psychometric function (Levitt, 1971), and these thresholds were increased by 12dB to produce the tones used for the remainder of the experiment. Participants were informed that they may hear some sounds during the experiment which were part of the head localising process, and that the the threshold procedure was used to determine an appropriate sound level for this process (i.e. one that was audible but not disruptive).

The main experiment consisted of 8 blocks of 64 trials each, four low load and four high, with the order of blocks counterbalanced between participants. No feedback was given during experimental blocks, but at the end of each block participants were provided with a score of percentage correct on the visual task, to encourage engagement. Blocks lasted for ~4 minutes each, and participants were allowed to take breaks between blocks when needed.

Figure 2.1 shows a schematic diagram of the trial structure. Each trial began with a fixation cross presented at the centre of the screen for 1000 ms. Subsequently, a visual search array of either low (Figure 2.1, display a) or high load (Figure 2.1, display b) was presented for 100 ms. The very brief presentation of the visual display in combination with the heightened computational demands of the conjunction-search task results in high perceptual load. On 50% of the trials, the visual display was accompanied by a 100 ms auditory tone, which was timelocked with the visual display. A blank screen was then presented for 1900 ms, during which participants were to make a speeded response on a button box regarding the identity of the visual target (using their right hand – button 1 for X and 2 for Z).

Before the main experimental blocks, MEG data was collected from two short blocks (4 minutes each) which contained only auditory stimuli and served to characterise the response to auditory signals while there was no concurrent visual activation. Each block consisted of 200 presentations of pure tones, 50 at each of the four frequencies used in the experiment, with ISIs randomly distributed between 700 and 1500 ms in 100 ms increments. Participants fixated at the centre of the screen and did not respond to the tones. Participants were told that these blocks were run to help localise their brain in the scanner.

2.2.1.5 Passive Group

To ensure that any effects of load on evoked responses found in the main group were not trivially due to differences in the visual stimulus displays between load conditions, an additional group of naïve subjects participated in a control session in which they viewed the visual stimuli, but did not perform any task. No tones were presented to this group. Blocks consisted of 200 presentations of the visual displays (one block contained the displays from the low-load task and the other contained the displays from high load task; block order counterbalanced across subjects) and lasted 3 minutes each.

2.2.1.6 Analysis

Behavioural Data

Data from the visual search task were subjected to within-subjects comparisons of proportion correct and reaction times between low and high load. For the reaction time data, a paired t-test was performed. The proportion correct scores were close to ceiling in both load conditions, so to account for the non-normal distribution, a mixed effects logistic regression was applied. The regression modelled correct trials as a proportion of all trials, with Load (low vs high) as a fixed factor and Subject as a random factor.

Preprocessing

The MEG data were epoched into 700 ms intervals, including 200 ms pre-onset, and baseline corrected to pre-onset interval. Epochs with amplitudes above 3pT (~6% of trials) were considered to contain artefacts and discarded. A PCA-based, De-noising Source Separation (DSS, de Cheveigné and Parra, 2014) routine was applied to each stimulus condition (Auditory alone - **A**, Visual alone Low load - **VL**, Visual alone High load - **VH**, Visual and Auditory Low Load - **AVL**, Visual and Auditory High Load - **AVH**) to remove 50Hz electrical noise, and extract stimulus-locked activity (maximise reproducibility across trials). Scree plots indicated that the first 3 components were considerably more repeatable than the others, but to ensure all activity of interest was retained, a conservative selection of the 12 most repeatable components in each condition were projected back to sensor space.

Separating Auditory and Visual Responses

In the main experiment auditory stimuli were always presented concurrently with the visual search array. While primary auditory and visual cortices are spatially distinct, MEG sensors at one site may pick up activity from distant neural sources due to spread of magnetic fields. This means that data recorded from temporal sensors may be contaminated by visual activity, and any effects found at those electrodes might not strictly reflect auditory processing. To separate the auditory and visual responses, a second stage of DSS analysis was applied to each subject’s data. This analysis was designed to identify response components that differed between audio visual trials (AV, tone-present trials), and visual only trials (V, tone-absent trials). For details of the covariance matrices used to define this criterion, see de Cheveigné & Parra (2014). The first two components were found to explain the vast majority of the variance between V and AV trials and were therefore projected back into sensor space and analysed as the auditory response (Fig 2.2B). The remaining components (i.e., those which explained very little of the variance between V and AV trials) were projected back to sensor space and analysed as the visual response (Fig 2.3B). The DSS reweighting was calculated over the entire dataset (collapsed over low- and high- load blocks; all channels) and this same channel reweighting was then applied to both high and low load datasets. I also tested similar analyses where the reweighting was calculated based on only the low, or only the high load data (and then applied to both) or where I applied the low load and high load reweighting independently, and all methods showed comparable results. I used the data where the reweighting was done based on all trials together as I considered this to be the most neutral analysis.

Importantly, the DSS procedure is 'blind' in that it does not separate the data based on prior expectations of the time-course or source of auditory or visual responses, but implicitly defines the auditory response as activation which is observed in the AV but not V condition. To assess the success of the DSS procedure at separating the auditory and visual response components from the AV response, I compared them to auditory and visual responses that were recorded in isolation. These consisted of visual responses from the trials (50%) which did not contain a tone, and auditory responses recorded during the blocks at the beginning of the experiment where the tones were presented alone. The evoked responses and source estimates for these trials are pictured in Figure 2.2 (visual) and 2.3 (auditory) alongside the data from the DSS analysis. Both auditory and visual responses separated via DSS match the unimodal counterparts in terms of shape, time course and source, which indicates that the DSS analysis was successful.

All analysis of load effects was run on the data from AV trials after separation via by DSS (i.e., Panel B in Figures 2.2 and 2.3) in order to assess processing during concurrent visual and auditory stimulus presentation, when competition for resources may have occurred.

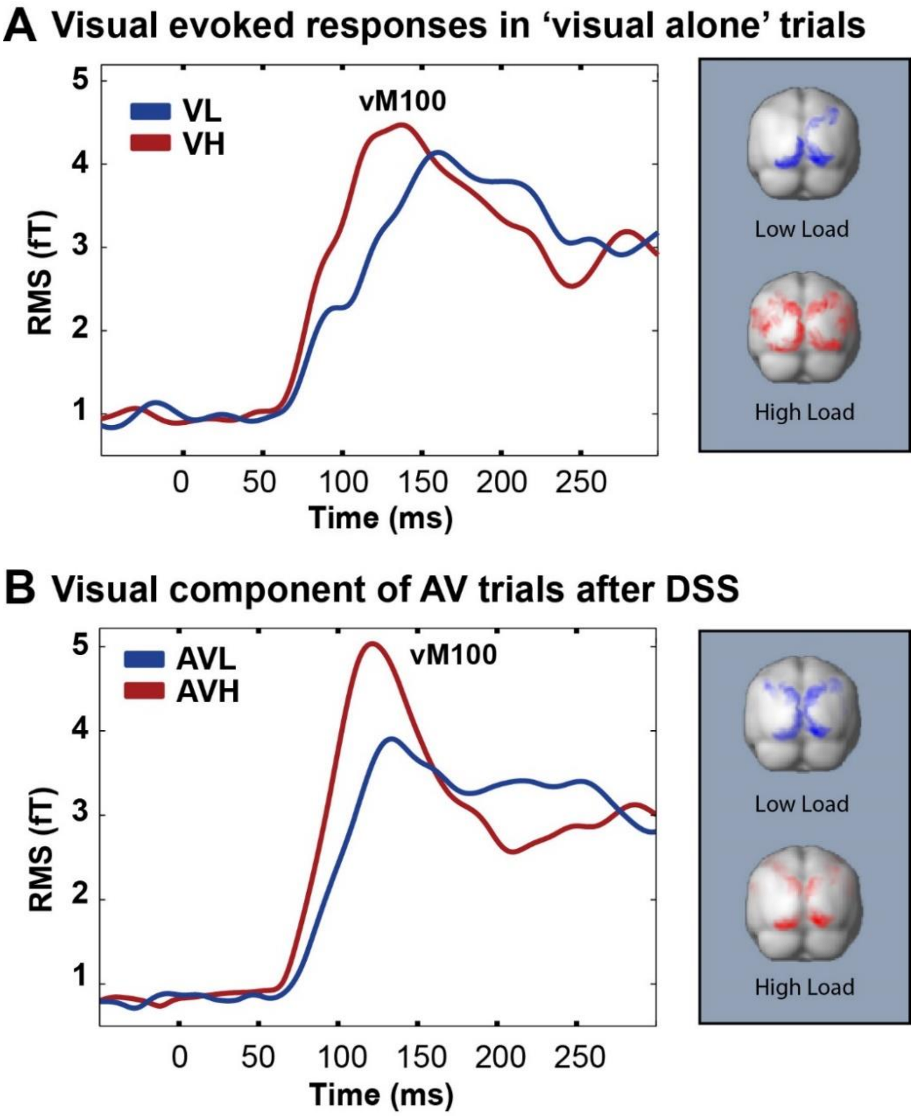
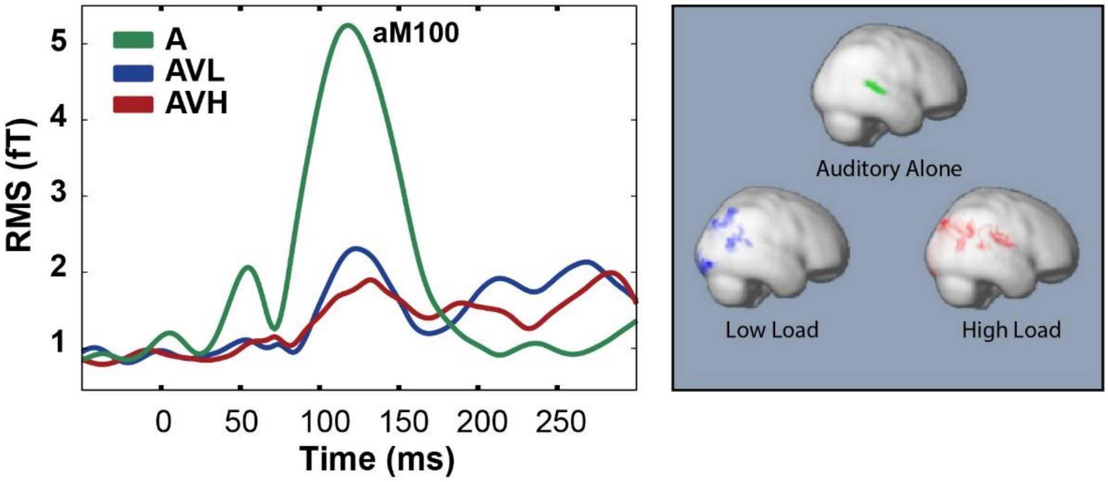


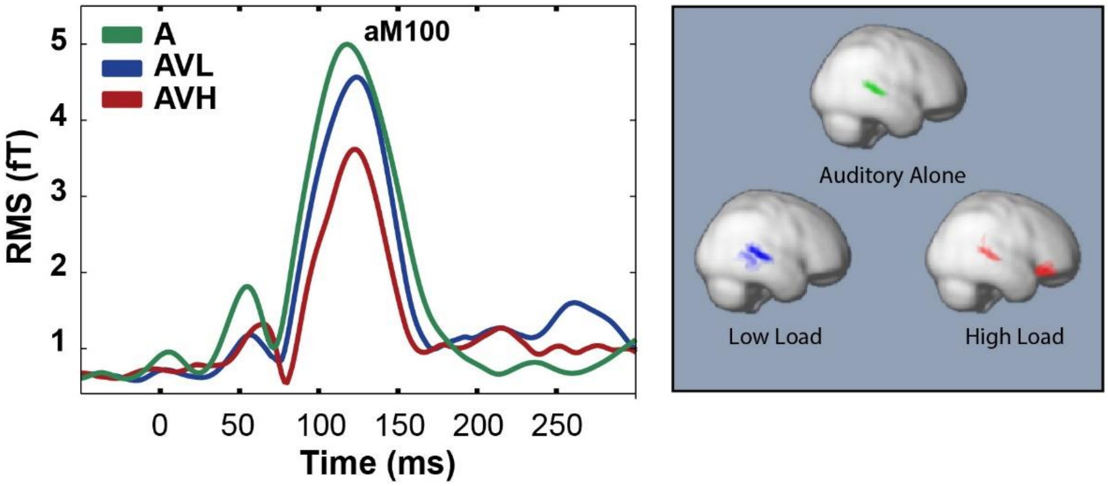
Figure 2.2: Experiment 1 - Visual evoked responses from visual only trials vs DSS analysis

A: *Data from the VL and VH conditions (i.e., when the visual stimuli were presented without an accompanying tone).* ***B:*** *DSS-derived visual responses from the AVL and AVH trials. Both the evoked responses and the source analysis results are comparable to those in Panel A.*

A Auditory Components Before DSS



B Auditory Components After DSS



<u>Figure 2.3: Experiment 1 - Auditory evoked responses before and after DSS analysis</u>

***A**: Data for the Auditory Alone condition (A), and the Low and High load 'Visual and Auditory' conditions (AVL, AVH) before DSS was applied (left: grand RMS over auditory sensors; right: source analysis). The RMS plot demonstrates a weak auditory signal in the AVL and AVH conditions, likely due to an overlap with concurrent visual activity. Source analysis confirmed dispersed activation in the AVL and AVH conditions relative to the A responses. **B:** Data from the same conditions after the DSS analysis was applied. The auditory responses across the three conditions are similar in shape (but with a clear effect of load, explored further in Figure 2.4). Source analysis confirmed clear auditory activity, with no remaining activity in the visual cortex.*

Channel Selection

From the blocks where auditory signals were presented alone, the auditory M100 (aM100) onset response was identified (c.f. Hari, 1990; Roberts et al., 2000; see Figure 2.3A). The aM100 current source is generally robustly localised to the upper banks of the superior temporal gyrus in both hemispheres (Hari, 1990; Pantev et al., 1996; Lütkenhöner and Steinsträter, 1998), and is apparent in MEG data as a source/sink pair located over the temporal region of each hemisphere on the scalp map. For each subject, the 40 most strongly activated channels at the peak of the aM100 (20 in each hemisphere, 10 positive, 10 negative) were considered to best reflect activity in the auditory cortex, and thus selected for the analysis of the data in the main experiment. Similarly, using the data for the **VL** and **VH** conditions (visual stimuli only; collapsed over load conditions), the visual M100 (vM100), which dominates the evoked response (see Fig 2.2A), was identified for each subject as a source/sink pair over the occipital lobe (c.f. Hashimoto et al., 1999). The source of the vM100 (and its EEG counterpart, the P1) is reliably located within the striate cortex (Nakamura et al., 1997; Shigeto et al., 1998; Hashimoto et al., 1999). For each subject, the 40 channels (20 in each hemisphere) which showed the strongest activity during the vM100 response were then selected for analysis of visual responses in the main experiment. Importantly, the DSS analysis detailed above was based on data from all channels; the channel selection as described here was used, after the application of DSS, for deriving the RMS activation measure (see below), time-frequency data, and subsequent statistics.

Evoked Responses

For each condition, in each hemisphere, the root mean square (RMS) of the field strength across the 20 selected channels was calculated for each sample point. The time course of the RMS, reflecting the instantaneous power of neural responses, is employed as a measure of the dynamics of brain responses. For illustrative purposes, group-RMS (RMS of individual subject RMSs) is plotted but statistical analysis was always performed across subjects. Analysis was performed separately for each hemisphere, but results were comparable in each, so reported results are collapsed over both hemispheres.

To compare responses between load conditions (**AVL** vs **AVH**, for the auditory and visual responses) the difference between the RMSs for high and low load were calculated for each participant, and subjected to bootstrap re-sampling (1000 iterations, balanced; Efron and Tibshirani, 1993). The difference was judged to be significant if the proportion of bootstrap iterations which fell above/below zero was more than 95% (i.e., $p < .05$) for 15 or more adjacent samples (25 ms). The bootstrap analysis was run over the entire epoch (200 ms pre onset to 500 ms post onset); all significant intervals identified in this way are indicated in the relevant figures.

Though the time-domain auditory and visual responses were drawn from different subsets of sensors, there was a small amount of overlap in some participants. However, the load effects reported remained the same when any overlapping sensors were removed. Moreover, when the auditory responses were derived only from the more frontal sensors (which have the opposite polarity to occipital activity) versus the posterior sensors (which have the same polarity as occipital activity), the load effect remains the same in both cases, indicating that the effects reported are not trivially due to an imperfect separation of auditory and visual responses.

Time-Frequency Analysis

A time-frequency analysis was conducted to examine potential load-induced oscillatory effects. This analysis was based on data which had been preprocessed to remove noise, but had not undergone the second stage of DSS analysis to separate auditory and visual responses, since this process is focused on enhancing evoked activity and may remove induced oscillatory activity. Data for each trial were converted to time-frequency space using a Morlet wavelet transform with 7 cycles, across frequencies from 5-40Hz. For each participant, the power spectra were then averaged over trials in each condition separately across the temporal and occipital sensors (selected for each individual as described above). This resulted in time-frequency images for each subject and condition, from which a t-statistic image was calculated giving a t value at each time-frequency ‘voxel’. A mixed design ANOVA was used to compare the t-statistic images of the time-frequency power spectra between conditions; Load (high vs low) and Trial Type (**V** vs **AV**) were entered as fixed factors, and Subject was entered as a random factor. The results reported refer to the main effect of load in this analysis and are significant at the $p < .05$ level after FWE (family wise error) correction at the cluster level for time-frequency voxels. There were no significant interactions between load and trial type.

Source Localisation

Activity within the time windows identified in the RMS analysis was localised using the multiple sparse priors method (MSP, Litvak and Friston, 2008). Inversions were based on all MEG channels and used a single shell head model and group constraints. Second-level analyses consisted of t-contrasts to compare activation between load conditions. Activity was thresholded at $p < .001$ (uncorrected) at the peak level and the resultant clusters were subjected to a $p < .05$ threshold with FWE corrections. Where comparisons between load conditions were used, the contrasts were defined in the direction indicated by the RMS data i.e., low > high or vice versa.

The inversions reported were performed on the data after the DSS analysis had been applied and therefore largely reflect sources of time-locked activity. Source analysis based on non-processed data was also performed to identify any additional (potentially non-time locked) sources. This analysis did not reveal any further significant activity.

2.2.2 Results

2.2.2.1 Behavioural

The high load task resulted in slower response times and poorer performance: mean task reaction time was increased in the high load (M = 826 ms, SD = 105) compared to low load conditions (M = 648 ms, SD = 84), t (1,12) = -13.1, $p < .001$. A logistic regression showed that the proportion of correct responses was significantly lower in the high load (M= 88%, SD = 3.5) compared to the low load (M = 98%, SD = 1.8) task, z = -9.78, p<.001. These findings confirm that the load manipulation was effective.

To ensure that the auditory stimuli did not capture attention, performance on trials where the auditory stimuli were present versus absent were contrasted. There was no effect of the presence of the auditory stimulus on reaction times (low load *t*

(1,12)= -1.3, *p* = .22, *ns*; high load *t* (1,12)= 1.0, *p* = .35, *ns*) and similarly, when sound (present vs absent) was entered into the logistic regression to model proportion correct, neither sound (*z* = 0.40, *ns*), nor the load x sound interaction (*z* = -0.39, *ns*) were found to be significant. Thus, participants did not seem to pay attention to the tones, as instructed (see the 'Procedures' section of the Methods).

2.2.2.2 MEG

Visual Search Task

Responses to the visual search task are displayed in Figure 2.4. The high load response showed a significantly greater amplitude than the low load response from 65-135 ms post onset. This difference occurred during the rising slope and peak of the vM100 onset response and is likely to reflect the higher level of perceptual processing involved in the high compared to the low load conditions. Source localisation demonstrated that the load effect was associated with increased activation in broad regions of the occipital and parietal cortices (*p* < .05 FWE corrected, see Figure 2.4 and Tables 1 and 2). These findings are in line with previous fMRI studies of visual perceptual load (Rees et al., 1997; Donner et al., 2002), which show increased activity in the visual cortex and areas of the attention network under greater visual load.

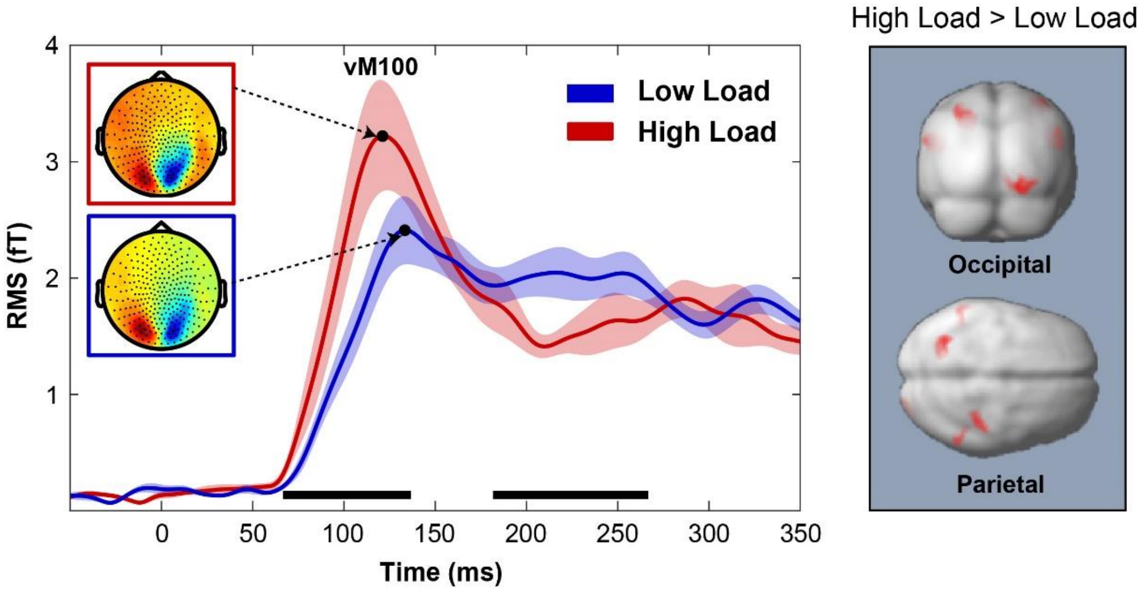


Figure 2.4: Experiment 1 - Visual Evoked Responses

Left: grand RMS responses from the DSS-separated visual responses from the AVL and AVH conditions. Horizontal black bars indicate intervals where response amplitudes differ significantly between load conditions. Inset are the scalp field maps of a representative participant at the peaks of the vM100 response. Right: areas where source localisations indicated a significant difference between activation in low and high load (HL>LL displayed no regions were found to be significant for the opposite contrast).

Table 2.1: Experiment 1 - Visual Activity Source Contrast, High Load > Low Load

Regions where contrasts between high and low load indicated significant differences in activation. Activity was thresholded at p < .001 (uncorrected) at the peak level and p < .05 (FWE corrected) at the cluster level. Each entry gives the anatomical description of the region(s) the cluster covers, the Brodmann areas in the cluster, its peak MNI co-ordinates, and the p-value associated with the cluster. For the LL>HL contrast, no regions were found to be significant.

Anatomical Description	**Peak MNI Co-ordinates**			**p value (FWE corrected)**
	x	**y**	**z**	
Right lingual gyrus, BA18	30	-94	-12	<.001
Left postcentral gyrus, BA2, BA3	-32	-68	50	.027
Right supramarginal gyrus, BA40	58	-52	30	.030
Left supramarginal gyrus, BA40	-48	-48	20	.047

Later in the trial, the low load condition showed an increased amplitude compared to the high load condition, likely reflecting the response decision and selection components of the task (e.g., greater decision certainty and faster response selection), and potentially response preparation, which are associated with greater behavioural accuracy in low compared to high load. There were no regions where the source estimates showed stronger activity in low load relative to high load, suggesting that this activity may represent distributed activity that did not correlate well between subjects.

The data from a passive viewing control group confirmed that the effects of load on visual evoked activity seen in during the main experiment were specific to performing an active task; when passively viewed, there were no significant differences between brain responses evoked by the low and high load visual stimuli (Figure 2.5).

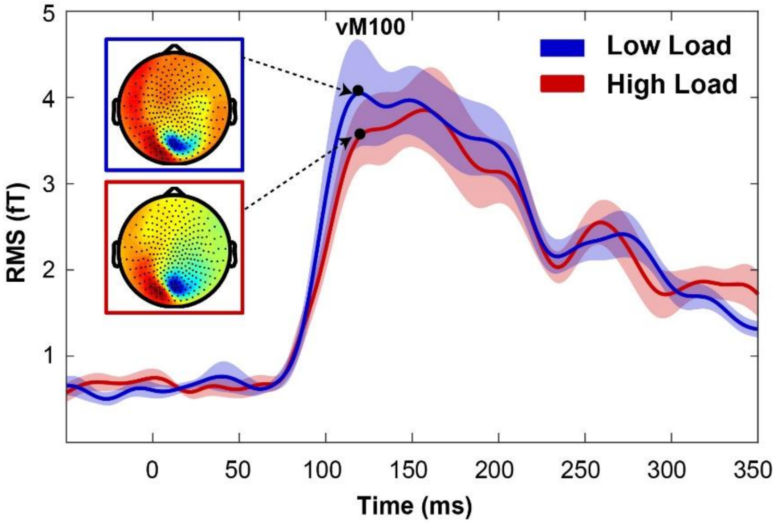


Figure 2.5: Experiment 1 - Visual Passive Group

Grand RMS responses for the passive control group, who viewed the visual stimulus arrays but did not perform a task. There were no significant differences between responses for low vs high load. Inset are the scalp field maps of a representative participant at the peaks of the vM100 response.

Auditory responses

Auditory responses are shown in Figure 2.6. Response amplitude was significantly reduced in the high- relative to low-load condition during the latter portion of the aM100 response, from 127-162 ms post stimulus onset. The sites associated with this modulation were the superior temporal sulcus (STS) and posterior middle temporal gyrus (MTG), both of which showed reduced activity under high compared to low load (p < .05, FWE corrected, see Figure 2.6 and Table 2).

An additional load effect was also observed later in the trial, between 243-272 ms post onset, where right hemisphere auditory responses showed a peak in the low but not high load condition (Figure 2.6). Response peaks with a latency in this range are identified as the P3 (Mecklinger et al., 1998; Kluge et al., 2011) known as the 'awareness positivity', with generators in Heschl's gyrus and STG (Opitz et al., 1999), and frontal regions (Comerchero and Polich, 1999; Polich and Comerchero, 2003). The P3 is observed as a positive polarity response in EEG. Its MEG counterpart is usually associated with a topography similar to the aM100, as is also found in the present data. The overall pattern is therefore consistent with a P3 response occurring under low but not high load. The difference in amplitude between load conditions in this time range is associated with increased activity in the superior temporal gyrus (STG) bilaterally in low compared to high load (*p* < .05, FWE corrected, see Figure 2.6 and Table 2).

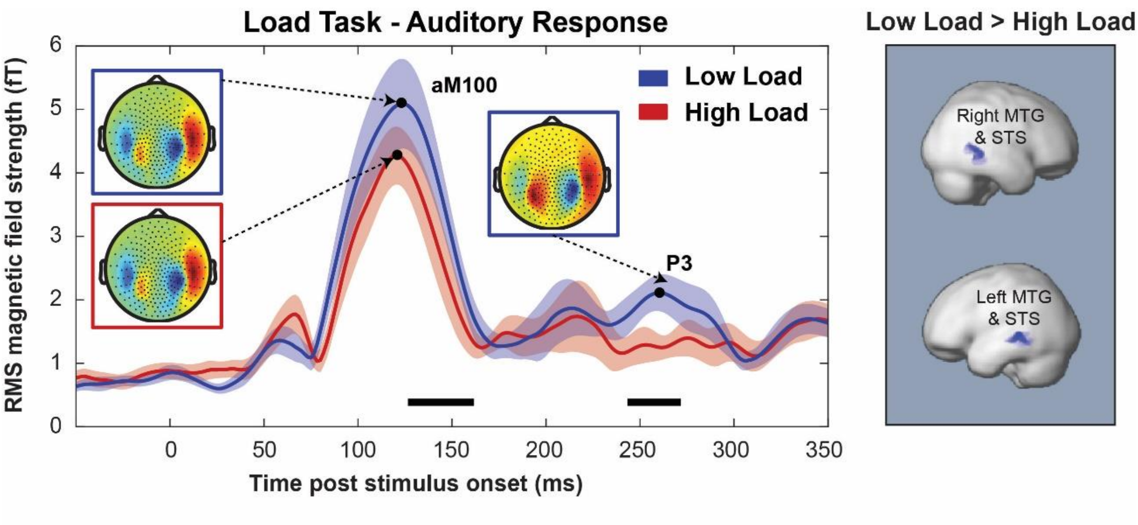


Figure 2.6: Experiment 1 - Auditory Evoked Responses

Left: grand RMS responses from the DSS-separated auditory responses from the AVL and AVH conditions. Horizontal black bars indicate intervals where response amplitudes differed significantly between load conditions. Inset are the scalp field maps of a representative participant at the peaks of the aM100 and P3 responses. Right: areas where source localisations indicated a significant difference between activation in low and high load (LL>HL displayed no regions were found to be significant for the opposite contrast).

Table 2.2: Experiment 1 - Auditory Activity Source Contrast, Low Load > High Load

Regions where contrasts between high and low load indicated significant differences in activation. Activity was thresholded at p < .001 (uncorrected) at the peak level and p < .05 (FWE corrected) at the cluster level. Each entry gives the anatomical description of the region(s) the cluster covers, the Brodmann areas in the cluster, its peak MNI co-ordinates, and the p-value associated with the cluster. For the HL>LL contrast, no regions were found to be significant.

Peak (ms)	**Anatomical Description**	**Peak MNI Co-ordinates**			**p value (FWE corrected)**
		x	**y**	**z**	
aM100 (100-170 ms)	Left MTG and STS, BA21	-66	-26	-6	<.001
	Right MTG and STS, BA21	50	-36	-6	<.001
P3 (250-300 ms)	Left STG and MTG, BA22, BA41	-48	-24	4	.012
	Right STG and MTG, BA22, BA41	46	-40	2	.004

Time-Frequency analysis

Time-frequency analysis on visual (occipital) channels revealed a significant effect of load on pre-stimulus oscillatory power: from the beginning of the epoch (-200 ms) to -93 ms there was increased power at 8-9Hz in low compared to high load trials. A similar effect was seen between 335 and 485 ms post stimulus presentation, with increased power at 8Hz in low compared to high load. These effects are consistent with numerous reports of alpha suppression during active attention (Fu et al., 2001; Kelly et al., 2006). Oscillatory power was also found to be higher in low load from 400 ms post stimulus presentation to the end of the epoch (500 ms) for frequencies between 15 and 30 Hz. This is likely to reflect response selection and preparatory motor responses (Zhang et al., 2008), which occur earlier in low load.

Crucially, the same analysis on auditory (temporal) channels revealed no significant effect of load on oscillatory power. No trends were observed even at a lower p threshold, indicating that the effects on evoked activity were not caused by pre-emptive, global, suppression of auditory cortical activity under high load.

2.3 EXPERIMENT 2: BEHAVIOURAL DUAL TASK

The findings from Experiment 1 suggest a plausible explanation for the neural effects underlying inattentional deafness: under high visual perceptual load, fewer processing resources are available to the auditory system, leading to a reduction in the sensory processing of sounds such that they do not reach awareness. To confirm that the reduced P3 measured in the MEG study was associated with reduced awareness of the tones under high visual load a further behavioural study, based on the same paradigm used in the MEG experiment, was run. Participants performed a dual task, monitoring the visual display and subsequently (see methods) reporting whether a tone was presented.

2.3.1 Methods

2.3.1.1 Participants

Eleven paid participants took part in the behavioural study. One was excluded due to an exceptionally high false alarm rate (M = 63%, for all included participants M = 8%). For the remaining ten participants (7 female), ages ranged from 18-29 years (mean = 22.2, SD = 3.3). All participants had normal or corrected to normal vision and reported normal hearing.

2.3.1.2 Apparatus, Stimuli and Procedure

The experiment was run on a Dell PC with a 13” monitor using Matlab 7.12 and Cogent 2000 (http://www.vislab.ucl.ac.uk/cogent.php). A viewing distance of 57cm was maintained throughout using a chin rest. Sounds were presented using the same tubephones as in the MEG study.

The visual and auditory stimuli were identical to those in Experiment 1 except that the number of trials in each experimental block was reduced to 56. Subjects performed the same visual task as in Experiment 1, while also performing a detection task on the simultaneous auditory stimuli. They were informed that the visual task was their priority, and made speeded responses to the visual search task (right hand: ‘0’ for X and ‘2’ for Z). After the visual response, a screen prompted participants to respond with whether or not they had heard a tone (left hand: ‘A’ for sound absent, ‘S’ for sound present). To familiarise participants with the dual task procedure, they completed a series of short demos before the main experiment.

2.3.1.3 Analysis

Data from the visual search task were subjected to within-subjects comparisons of proportion correct and reaction times between low and high load. Values for d' were not calculated since the paradigm was a two-alternative forced choice (X or Z) rather than a detection (present vs absent) task. For the reaction time data, a paired t-test was used. The proportion correct scores were close to ceiling in both load conditions, so to account for the non-normal distribution, a mixed effects logistic regression was used. The regression modelled correct trials as a proportion of all trials, with Load (low vs high) as a foxed factor and Subject as a random factor.

Data from the auditory task were analysed to compare proportion correct, detection sensitivity (d') and response criterion (β) between low and high load. Reaction times were not analysed since in the dual task paradigm the response to the auditory detection task was delayed until after the visual response had been given. For proportion correct, a mixed effects logistic regression was performed with Load as a fixed factor and Subject as a random factor. For d' and β, paired t-tests were used to compare values within subjects for low versus high load.

2.3.2 Results

Performance on the visual task paralleled that observed in Experiment 1. The visual search task showed increased reaction times from low (M = 710 ms, SD = 114) to high load (M = 883 ms, SD = 124), $t(1,9) = -7.3$, $p < .001$. As in Experiment 1 there was also a reduction in accuracy from low load (M = 94%, SD = 6.8) to high load (M = 83%, SD = 5.7), $z = -4.76$, $p < .001$. These findings confirm that visual search task performance was equivalent to that found for this task in the scanner. In keeping with the instructions to treat the auditory detection task as secondary, there was no effect of the presence of the auditory stimulus on either reaction times (low load t (1,9)= 0.6, $p = .54$; high load t (1,9)= 0.2, $p = .84$) or accuracy (effect of sound present vs absent: $z = -0.72$, *ns*; load x sound interaction: $z = 0.30$, *ns*)) or on the primary visual search task.

In the auditory detection task, participants showed significantly reduced detection accuracy rates under high load (M = 88%, SD = 5.8) compared to low load (M = 92%, SD = 5.3; $z = -4.08$, $p<.001$) conditions, as well as reduced detection sensitivity (d') under high load (M = 2.6, SD = 0.7) compared to low load (M = 3.2, SD = 0.7), $t(1,9) = 3.7$, $p < .01$. Their response criterion (β) did not differ significantly between low (M = 3.2, SD = 3.4) and high load (M = 1.3, SD = 0.8), $t(1,9) = 2.0$, $p = .07$. These data demonstrate that the load task used in Experiment 1 does indeed impact rates of awareness and conversely inattentional deafness – listeners were less likely to detect the tones when these were presented during the high load task, relative to a low load visual task (see also Raveh and Lavie, 2015).

2.4 DISCUSSION

2.4.1 Visual load induces time-specific reduction of early auditory ERFs

The present findings establish the neural underpinnings of inattentional deafness under load. Auditory cortical responses to irrelevant tones that were presented while participants performed a visual search task were reduced during the high load version of the task compared to those recorded under low load. This effect of visual perceptual load was evident at both an early stimulus-processing stage (reflected in the modulation of the aM100) and during later processing (revealed by reduced amplitude of the P3) that is associated with awareness. Results from the behavioural dual-task version of the experiment, in which detecting the tones was an active but deprioritised task, confirmed reduced detection sensitivity of the auditory stimuli under the high (vs low) load conditions. Thus, the data suggest that focusing attention on a perceptually demanding visual task leads to reduced availability of neural resources required for perception of a simple auditory stimulus, resulting in reduced detection, and leading to the experience of inattentional deafness.

Importantly, the timing and reversed direction of the load effects on visual and auditory responses is precisely what load theory predicts on the basis of shared capacity. The increased visual processing in high load was only apparent during a specific portion of the evoked response, and it was precisely at this time that auditory responses became weaker under high visual load. This is supported by time frequency analysis, which established no effects of auditory suppression in either pre-stimulus or post-stimulus stages. Thus, rather than simply dampening all auditory activity in conditions of high visual load, the cost to auditory processing occurred only when the visual system had a very high demand for perceptual resources. This trade-off between visual and auditory activation is consistent with the capacity limited “push-pull” mechanism envisaged in load theory (see also Pinsk et al., 2004), and supports the conclusion that the task involves a temporally focused draw on perceptual resources shortly after the presentation of the display. Further, the fact that the impact of increased perceptual load on auditory responses occurred during the M100 onset responses suggests that the sharing of perceptual resources has an early locus, occurring during stimulus encoding.

2.4.2 Neural locus of effect of visual load on auditory processing

The findings that the effects of load modulation were localised in associative auditory cortex (e.g., MTG/STS) offers a plausible account for the source of shared audio-visual capacity. Since associative auditory cortex is known to mediate integration of audio-visual stimuli (see Calvert, 2001 for a review), the finding that the evoked response to the tone in this area was significantly modulated by visual perceptual load suggests these regions of associative auditory cortex as a possible site for cross modal audio-visual capacity limits. However, in the evoked responses there was evidence of an effect of load during the aM100 response, which represents activity much earlier in the auditory pathway in areas such as PT and STG (Hari, 1990; Pantev et al., 1996; Lütkenhöner and Steinsträter, 1998). Thus, it is plausible that the impact of visual load on auditory processing occurred earlier in the pathway, but the MEG source analysis did not provide sufficient sensitivity to illustrate the earlier effect. In this case the differential activity seen in the associative areas may reflect integration of the auditory and visual stimuli (which is plausible since they shared on- and offset times), that was reduced under high visual load because the auditory responses were weaker.

2.4.3 P3 response reflects subjective awareness

Increased visual load abolished the P3, a response typically understood to reflect awareness of a stimulus (Comerchero and Polich, 1999; Kok, 2001). This effect is consistent with previous findings that high load can increase incidences of inattentional deafness (Macdonald and Lavie, 2011; Raveh and Lavie, 2015), and with the data from the dual task experiment here.

The response recorded here under low load conditions is likely to be a P3a, which is a relatively early component of the P3, associated with involuntary shifts of attention (Comerchero and Polich, 1999). This is consistent with the load theory proposal that attention is involuntarily allocated to task-irrelevant stimuli in conditions of low load that leave spare capacity. While the P3a is typically associated with a frontal source (McCarthy et al., 1997; Escera et al., 1998; Opitz et al., 1999), it has also been shown that P3a generated by an auditory stimulus receives a contribution from auditory cortex (Escera et al., 1998; Opitz et al., 1999). The P3 response identified here appears consistent with this temporal generator, both in terms of its timing and the source in STG. While the frontal generators of the P3a are known to be affected by attention (Comerchero and Polich, 1999; Koivisto et al., 2009), the contribution from auditory cortex was previously believed to be automatic (Escera et al., 1998). The present results, demonstrating a load effect on STG activity during the P3a, suggest that this auditory cortical contribution to the P3a may also be sensitive to the effects of perceptual load.

The amplitude of the P3 to oddball stimuli in a secondary task has been used to quantify resource sharing during dual task conditions (Isreal et al., 1980; Wickens et al., 1983; Kramer et al., 1986), with larger P3 responses to the secondary task when it receives a higher priority. The data presented here highlight that even when stimuli are entirely irrelevant to the task, the P3 may be a useful measure of the availability of 'leftover' processing resources, as determined by the level of perceptual load in the attended task.

2.4.4 Relation to previous research

The present findings shed light in the controversy over whether visual load can affect auditory processing (Otten et al., 2000; Müller et al., 2002; Dyson et al., 2005; Restuccia et al., 2005; Muller-Gass et al., 2006; Muller-Gass et al., 2007; Sculthorpe et al., 2008; Parks et al., 2011; Chait et al., 2012). The mixed findings reported previously may be due to the fact that, as demonstrated here, the load effect is time-sensitive. Depending on the specifics of the visual load task and any presentation asynchrony between the auditory and visual stimuli, the time during which the visual task places a drain on perceptual capacity may or may not coincide with a resource-dependent point in auditory processing. This issue is particularly relevant for load tasks such as RSVP, which may place a high overall load on perceptual resources, but do not necessarily guarantee that demand is consistently high in a manner that would preclude any potential for spill-over of spare capacity.

2.4.5 Auditory processing is subject to limits on global resources

These findings give a crucial insight into the mechanism of sensory processing in the brain. In situations with numerous sources of sensory information, limits on our perceptual resources can cause our system to become overloaded, leading to reduced processing of stimuli that are not directly relevant to the current task and resulting in inattentional blindness and deafness. That these limits apply across sensory systems has implications for models of attention, and also the understanding of perception and behaviour in busy, real life situations, when multisensory information competes for processing resources. They also indicate that auditory processing, which has been assumed in many cases to act as a pre-attentive early warning system, may in fact to be susceptible to substantial deficits in functional capability (e.g., being rendered temporarily ‘deaf’) not only when there are high demands on auditory processing, but even for load on other sensory systems.

CHAPTER 3: FIGURE-GROUND SEGREGATION IN SHORT STOCHASTIC SIGNALS[2]

3.0 MOTIVATION

In Chapter 2 I showed that high visual load can reduce the amplitude of early sound-evoked responses in auditory cortex. In the case of the quiet, pure-tone stimuli used in the previous experiments, this was accompanied by reduced detection sensitivity for the sounds, indicating that high visual load can impact both the neural response to, and perceptual awareness of quiet sounds. This effect appeared equivalent to a time-specific dampening of auditory evoked activity at times when visual load was highest - perhaps equivalent to turning down the volume of the stimuli, or reducing the gain of the auditory system. However, much louder sounds would presumably still elicit auditory responses strong enough to reach the detection threshold. For these sounds, it remains to be seen whether visual load might impact additional aspects of auditory processing. For instance, since natural auditory scenes are typically comprised of multiple sound sources which overlap both temporally and across the frequency spectrum, even fundamental aspects of auditory scene analysis may rely on relatively complicated computations. Might these computations also be susceptible to limits on the availability of processing resources? Or do they occur

---

[2] Data from this chapter have been published in Molloy K, Lavie N, Chait M (2019) Auditory figure-ground segregation is impaired by high visual load. Journal of Neuroscience 39:1699-1708. Although I have rewritten much of the text so that it works in concert with the other chapters in this thesis, there may be some repetition of information.

locally within an encapsulated auditory network, with auditory-visual shared capacity applied more globally to modulate the overall gain on early evoked responses?

In the next series of experiments, I aimed to address this question, examining whether high visual load could impair computational processing within the auditory system. However, before load was introduced, I needed to identify a stimulus that modelled a computational auditory process and which could be combined with a brief but intensive load task such as that described in Chapter 2, to ensure that load coincided directly with auditory processing. In this chapter I introduce stimuli which model figure-ground segregation, the process by which an auditory object is perceptually extracted from the aggregate sound input. I tested a variety of stimulus parameters to find a short stimulus that elicited a clear MEG response that indexed segregation. This is a precursor to the work presented in Chapter 4, where I combine these auditory signals with a visual load task to examine whether the auditory computations are susceptible to depletion of resources.

3.1 INTRODUCTION

Figure-ground segregation is the process by which an auditory object is perceptually extracted from the aggregate sound input. This function underlies key aspects of listeners’ ability to make sense of the acoustic environment, including recognising individual sounds within crowded scenes and understanding speech in noise. Though auditory segregation has been shown to occur without explicit intent (Teki et al., 2011; O'Sullivan et al., 2015; Teki et al., 2016), the calculations which underlie this process can be complex and rely on combining multiple features of the auditory signals. In spite of a substantial amount of previous research, it is still unclear to what extent these calculations may rely on the availability of processing resources (Shamma and Micheyl, 2010; Shamma et al., 2011; Snyder et al., 2012; Puvvada and Simon, 2017).

There are two major factors which have hindered the conclusions of previous research on the topic: the first is that much of the research has focused on stimuli that are highly simplified, and thus may not capture the complexity (and potential demand) of the analysis that must occur in order to parse natural scenes. The second is that previous work may not have used sufficiently well-controlled attention manipulations. In the present chapter I address the issue of how figure-ground segregation is modelled, and the evidence of neural markers of auditory segregation. The question of how attention is controlled will be covered in Chapter 4.

3.1.1 Streaming Paradigm

One of the most common paradigms used to study auditory segregation focuses on how an on-going sequence of tones can be segregated into more than one 'stream' that are perceived as separate auditory entities. This work has typically used the 'ABA-' signals described in Chapter 1 (Section 1.1.1.5) which with certain parameters creates an ambiguous signal that can be perceived as fused or segregated into two on-going streams. However, since the stimulus is ambiguous, typically a behavioural response is required in order to determine subjective percept at any given time. These streaming paradigms are also frustrated by the fact that the percept may not be strictly bistable, with some participants reporting alternative percepts outside the 'fused' versus 'segregated' dichotomy, and with periods of transition where the percept is not stable.

To try to circumvent the issue of behavioural reports, more recently efforts have been made to characterise a neural signal which may reflect the degree of segregation that is experienced perceptually. Gutschalk et al. (2005) recorded MEG responses to an ambiguous ABA- stimulus, asking participants to indicate each time their percept shifted between fused and segregated. When they contrasted the neural responses to the B tones within the signals at times when participants perceived one versus two streams, they found that during the segregated percept the amplitude of M50-M100 responses to the B tone was significantly larger than when the streams were perceived as fused. In a similar set of experiments, Snyder, Alain and Picton (2006) also showed that there was a positive ERP component that peaked at ~200 ms post onset of each ABA triplet, whose amplitude built up over time during active attention, which they propose may reflect the build-up of the streaming percept. Interestingly, a similar paradigm used in an fMRI study found that differences in percept were correlated with activity within the intra-parietal sulcus (IPS), with stronger activity occurring when the streams were perceived as segregated; however, activity within auditory cortex did not differ between percepts (Cusack, 2005). Though it is not clear whether IPS pays a formative role in stream segregation or simply reflects perceptual organisation that occurred earlier in the processing hierarchy (but in this instance was not captured in the fMRI analysis), this result highlights that areas outside the auditory cortex may be involved in auditory stream segregation.

However, while there has been a wealth of research using this paradigm it represents a very specific process which is reliant on stimulus parameters that were carefully balanced to result in ambiguity. Further, the signals are highly simplified, with the two objects being artificially separated temporally and in frequency space. Thus, it is unlikely to accurately represent the processes which must occur during segregation of natural sounds, where spectral and temporal overlap is highly likely.

3.1.2 Concurrent Segregation Paradigm

Another popular paradigm focuses on how cues within concurrently presented sounds can influence which elements within the signal are perceptually fused together or segregated. Typically, a harmonic complex tone is presented, which may contain a single component that is mistuned. If the degree of mistuning is sufficient, this causes the mistuned component to be perceived as a separate pure tone object alongside the main harmonic complex (Moore et al., 1986). Other cues such as inducing an onset asynchrony between the main complex and a (tuned) component also led to the same effect of a segregated percept (Lipp et al., 2010). Neurological studies using these stimuli have shown that there are reliable neural responses which occur when listeners report the percept of a component segregated from the main sound. The first of these responses is the object-related negativity (ORN), a negative EEG component which is superimposed on the N1-P2 complex and peaks between 150-300 ms post stimulus onset stimulus onset (Alain et al., 2001; Alain and Izenberg, 2003; McDonald and Alain, 2005; Alain and McDonald, 2007; Lipp et al., 2010). The ORN has been recorded even when participants are not explicitly listening to the sounds (Alain et al., 2001; Alain et al., 2002; Alain and Izenberg, 2003; Dyson et al., 2005; Lipp et al., 2010). There are also later markers of segregation which are more susceptible to attention: the P230 occurs only when attention is actively engaged on the task, and the P400 reflects detection of the separate object when it is task-relevant (Alain et al., 2001; Alain et al., 2002; McDonald and Alain, 2005).

While these paradigms have well-established neural markers that index segregation and are effective in exploring how the auditory system segregates concurrently presented objects which overlap spectrally, the signals used are temporally static and so do not capture the dynamic nature exhibited by many natural sounds.

3.1.3 Informational Masking Paradigm

To address the limitations of these classic stimulus paradigms, more recent work has focused on signals which are both spectrally and temporally complex so as to more accurately model the processes which occur during natural listening. One example of this is the informational masking paradigm, which combines a target stream of tones with a dynamic ‘tone cloud’ masker (Gutschalk et al., 2008; Elhilali et al., 2009a). The target is a regular stream of fixed frequency pure tones, while the tone clouds are formed of pure tones with scattered onsets and ISIs, whose frequencies are distributed on a log scale. The density of the tone cloud can be manipulated by increasing or decreasing the average temporal or frequency spread between the masking tones (Figure 3.1 – left versus right). However, a frequency band around the target stream is always protected, with no masker tones appearing within that area of the spectrum, leaving the target stream as the only energy within that band (see Figure 3.1). This spacing was introduced to ensure that there is no low-level ‘energetic’ interference between the target and the maskers at the stage of the cochlea, so that any failure to detect the target (which is perfectly audible when presented alone) must be attributed to 'informational' masking, which occurs further up the processing hierarchy.

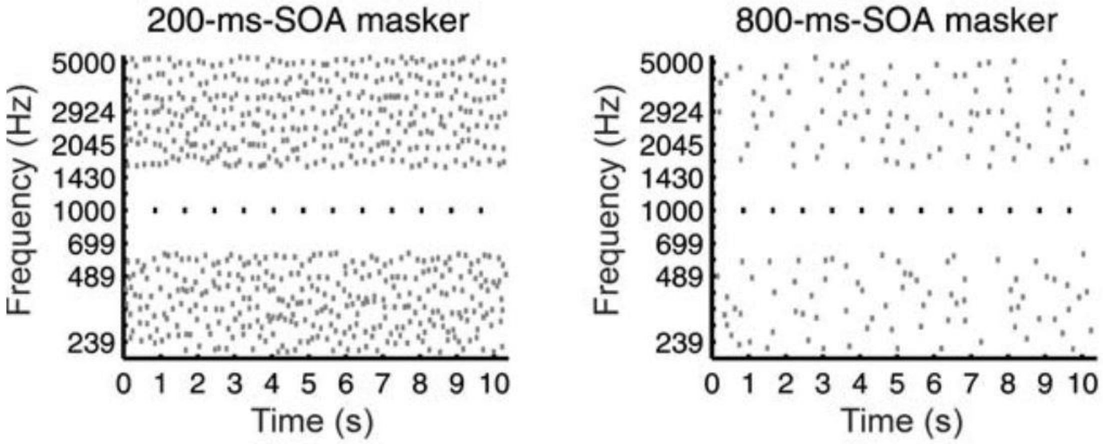


Figure 3.1: Informational Masking stimulus schematic from Gutschalk et al. (2008)

The target is a regular stream of tones of a fixed frequency within a 'tone cloud' masker. The masker is formed of pure tones log distributed over a broad frequency spectrum with scattered onsets and ISIs. The possible frequencies of the masking tones exclude a 'protected' region around the target frequency so that there is no low-level interference between masker and target. The density of the masking cloud can be manipulated by increasing or decreasing the average temporal and frequency spread between the masking tones; the left example shows a dense masker while the right example shows a sparser masker.

Studies using this paradigm have shown that there are neural markers which reflect identification of the target within the stimulus. Gutschalk et al. (2008) contrasted EEG responses between active and passive conditions for sounds that were either cued or uncued, and for targets that were detected versus undetected. They reported a negative EEG component that occurred from ~50-250 ms post stimulus onset that was only apparent in trials when participants detected a target, which the authors term the attention-related negativity (ARN). They also reported a steady state response (SSR) that phaselocked to the amplitude modulation within the target tones, which had similar strength regardless of whether participants were performing an active task, and for both detected and undetected targets. More recently, however, this SSR has been shown to be strongly modulated by the top-down focus of attention in an experiment where participants either performed a task based on the target stream or the tone cloud masker (Elhilali et al., 2009a), suggesting that the manipulation used by Gutschalk et al. may not have been sufficient to illustrate an effect of attention on the SSR.

While these studies used stimuli that are more naturalistic than the simplified stimuli used in the classic paradigms, the targets were nonetheless artificially separated in frequency space from the maskers. Since the auditory processing hierarchy is organised tonotopically (i.e., shows a spatial organisation based on frequency), this means that the target in these signals activates neurons which are spatially segregated from those which respond to the masker. This is by design, since this paradigm questions how a signal can be masked by other stimuli which, though they do not interfere physically with the target, nonetheless overwhelm the auditory system's capacity for processing. However, detection of the target may not rely on auditory computation, per se, since it produces an inherently separate signal from the maskers from even the first stage of the auditory pathway – the basilar membrane within the cochlea.

3.1.4 Stochastic Figure-Ground Stimuli

To specifically target figure-ground segregation which relies on computational processing, in the present work I used the Stochastic Figure Ground stimulus developed by Teki and colleagues (SFG: Teki et al., 2011; Teki et al., 2013; Teki et al., 2016). Signals are a regular series of chords formed of random, changing frequency components which create a stochastic background. 'Figures' are introduced by repeating a small subset of components throughout the signal, resulting in the percept of a complex, bound auditory object that is heard separate from the background (Teki et al., 2013). Importantly, components belonging to the figure and those that are part of the ground are spread across the spectrum such that at any given point in time, the figure and the ground do not provide inherently separate neural signals. Instead, the percept of a figure emerging from the background arises from processes which computationally segregate the figure from the ground via a mechanism which computes the 'temporal coherence' across frequencies (Elhilali et al., 2009b; Shamma et al., 2011; Krishnan et al., 2014). This computation is based on analysing how information in separate frequency channels changes over time, and binding together components which show similar temporal dynamics (Teki et al., 2013; Teki et al., 2016). This process allows auditory objects to be segregated despite spectro-temporal overlap, and so it is likely to play a vital role in segregating the spectrally broad, dynamic sounds we experience in natural environments. It also relies on a signal which can only be identified via computation, and so provides an index of the level of computational processing which is occurring within the auditory system.

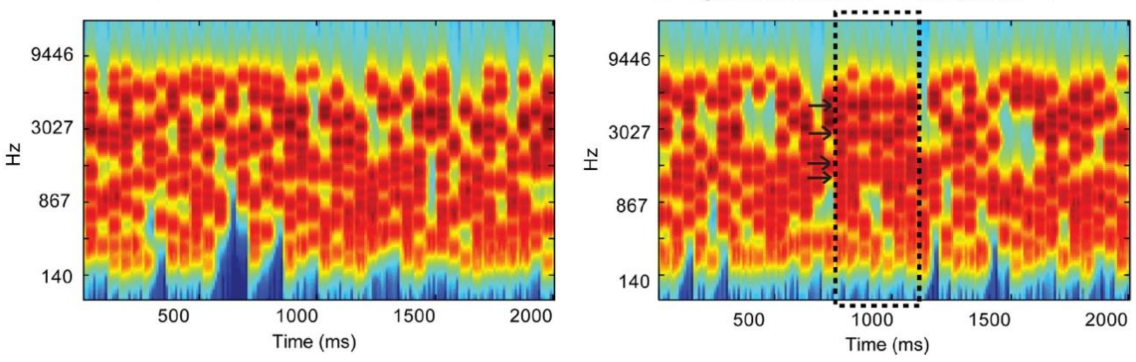


Figure 3.2: Stochastic Figure-Ground stimuli from Teki et al. (2011)

Accumulating work has shown that listeners are highly sensitive to SFG figures, and brain responses to the emergence of the figure have been consistently observed in naïve, distracted listeners performing an incidental task (Teki et al., 2011; O'Sullivan et al., 2015; Teki et al., 2016). Teki et al. (2011) recorded fMRI while presenting incidental SFG signals to participants who were engaged in a low-demand visual picture matching task. SFG signals consisted of an on-going ‘ground’ of 50 ms chords formed by random frequency components, which lasted 2000 ms in total. In half of the signals a figure, comprised of a set of repeated frequency components, appeared part way through the sequence and was present for a small number of chords. Both the number of frequency components in the figure (the ‘coherence’) and the number of chords in the figure (the ‘duration’) were varied. Figure 3.2 shows example spectrograms from a signal with no figure (A) and one with a figure of coherence 4 and duration 7 (B).

Behavioural work has shown that the percept of the figure becomes stronger as the coherence and duration are increased (Teki et al., 2011; 2013), and fMRI recordings indicate several brain regions that show a similar pattern: activation in the intraparietal sulcus and the superior temporal sulcus increased as the coherence and/or duration of the figure increased (Teki et al., 2011). This indicates that the brain responds to these spectro-temporal regularities in the signal even when they are irrelevant to the current task (and indeed, when participants are not even aware that figures might exist within the longer signals).

Similar work using EEG has shown that even more complex figures can be detected during passive listening. O’Sullivan et al. (2015) used figures that were not static, but instead comprised of sets of components that varied in frequency in the same manner over time (see red components in Figure 3.3). They found that the global field power of the EEG responses showed a significant peak ~ 200 ms after the onset of the coherent figure, even during passive listening. During active listening this response had a larger amplitude and duration, and was followed by a second significant peak around 600 ms post-stimulus onset.

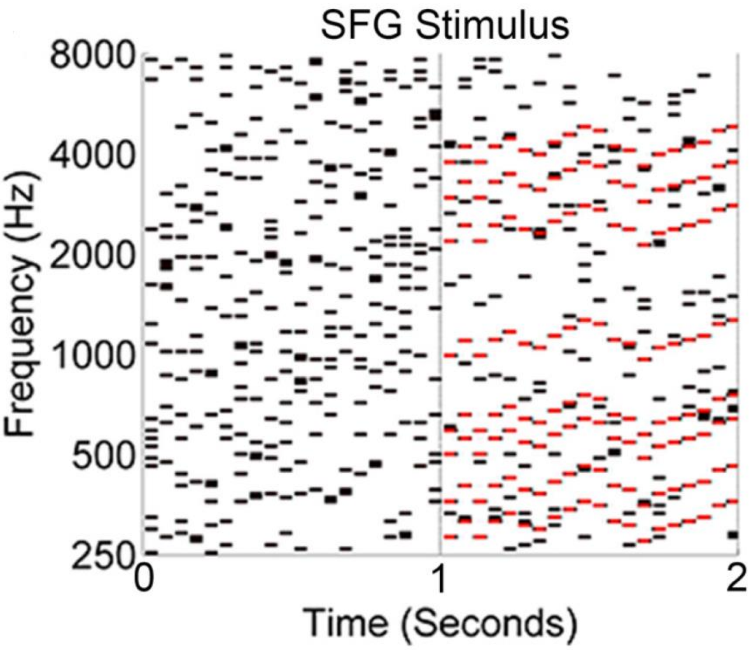


Figure 3.3: Stimulus schematic from O'Sullivan et al. (2015)

SFG signals consisted of 50 ms chords. Ground components (black) were randomised, while figure components (red, coherence 8) followed a single randomly generated pitch envelope, so that those components varied coherently over time.

A paradigm similar to the active condition reported by O'Sullivan et al. (2015) was used by Toth et al. (2016). Using a static figure, they illustrated a negative EEG component that peaked at ~300 ms post stimulus onset and a later positivity at around 600 ms. These responses appear to mirror those reported by O'Sullivan et al. (2015). Both peaks showed larger amplitude when the parameters of the stimulus were set to give a stronger figure (i.e., increased coherence or duration), but while the earlier response was not correlated to behaviour, the amplitude of the latter peak increased with detection performance (Toth et al., 2016). Together, these studies suggest that there is an early response evoked by the figure that occurs even during passive listening but can be enhanced during active listening, and that there is a later response which may reflect target detection rather than being evoked specifically by the figure-ground segregation process. In both studies a source analysis was performed on the early component and both suggested a superior temporal/inferior parietal source (O'Sullivan et al., 2015; Toth et al., 2016).

These stimuli have also been investigated using MEG recordings. Teki et al. (2016) presented ground stimuli that extended over several seconds, during which a figure of varying coherence might appear part-way through the signal and remain present until sound offset. In contrast to the 'peak' responses reported previously when shorter figures were used (O'Sullivan et al. 2015; Toth et al., 2016), this study showed an early transient after figure onset followed by a sustained response to the figure, which varied in strength according to the coherence of the figure. The data were source localised to find brain regions where activity varied parametrically with the coherence of the figure. This revealed that both the onset peak and the sustained response had sources within the planum temporale (PT) and the IPS. Activity in PT is consistent with the broad temporal sources suggested by the EEG data, though interestingly, no temporal source was indicated by the fMRI data from Teki et al. (2011). The IPS source was consistent with the fMRI results (Teki et al., 2011) and suggests that this area, which is outside the classic 'auditory cortex' is nonetheless involved in identifying the auditory figures. This is also consistent with previous evidence that the IPS may play a role in auditory segregation (e.g., Cusack, 2005).

3.1.5 Chapter Overview

The imaging studies reviewed above have shown that the brain produces reliable neural responses to figures within the SFG signals, even without participants having explicit knowledge of the signals or intending to listen. Importantly, the strength of the neural activity also appears to correspond to the strength of the figure as parameters such as coherence and duration are varied, and has been shown to be enhanced by active attention. This suggests that it may provide a sensitive index of the degree to which the figure has been processed. However, previous SFG stimuli have been formed of long, on-going scenes where the figure may appear at some point during the sound. As I have discussed previously, it is not clear whether extended load tasks induce sufficiently consistent high load to produce reliable effects experimentally, particularly on auditory processing; thus, paradigms which combine short auditory stimuli with precisely timed visual load tasks, such as that reported in Chapter 2, may have a better chance of having a measurable effect on auditory processing.

In this chapter, I describe how I adapted the SFG stimuli to produce short signals which could be used in parallel with a visual search task similar to that used in the previous experiments. I present MEG data recorded in response to these short SFG signals, and show that in spite of the substantially reduced duration there is still a clear neural signal in response to the figures.

3.2 METHODS

3.2.1 Participants

Sixteen paid participants (9 male; mean age of 24.8 years, SD = 3.0 years) took part in the experiment. All were right handed, had normal or corrected to normal vision and reported normal hearing and no history of neurological disorders.

3.2.2 Apparatus, Stimuli and Procedure

The magnetic signals were recorded using a CTF-275 MEG system (axial gradiometers, 274 channels, 30 reference channels, VSM MedTech, Canada) in a magnetically shielded room. Subjects were seated in an upright position, with the visual stimuli projected onto a screen placed ~52cm from the participants' eyes. Data were recorded continuously with a 600Hz sampling rate and a 100Hz hardware low-pass filter.

The auditory stimuli were ~200 ms long, diotically presented SFG stimuli (Teki et al., 2011; 2013; 2016, see Figure 3.1). Signals consisted of a succession of chords, each comprised of multiple frequency components. Frequencies were chosen from a log-distributed pool of 109 frequencies from 180 to 4067Hz. Each chord was comprised of between 11 and 21 (number was uniformly distributed) frequency components, which were selected from the frequency pool with equal probability. The *'figure-absent'* stimuli (FA, 50%) were formed of random frequency chords. The *'figure-present'* stimuli (FP, 50%) were constrained so that a subset of the frequencies were repeated in each chord (this parameter is referred to as the 'coherence' of the figure, see Teki et al., 2011; 2013; 2016), while the others were selected randomly for each chord. The repetition of the coherent frequencies creates the auditory 'figure' which can be heard separately from the stochastically changing background (Teki et al, 2013).

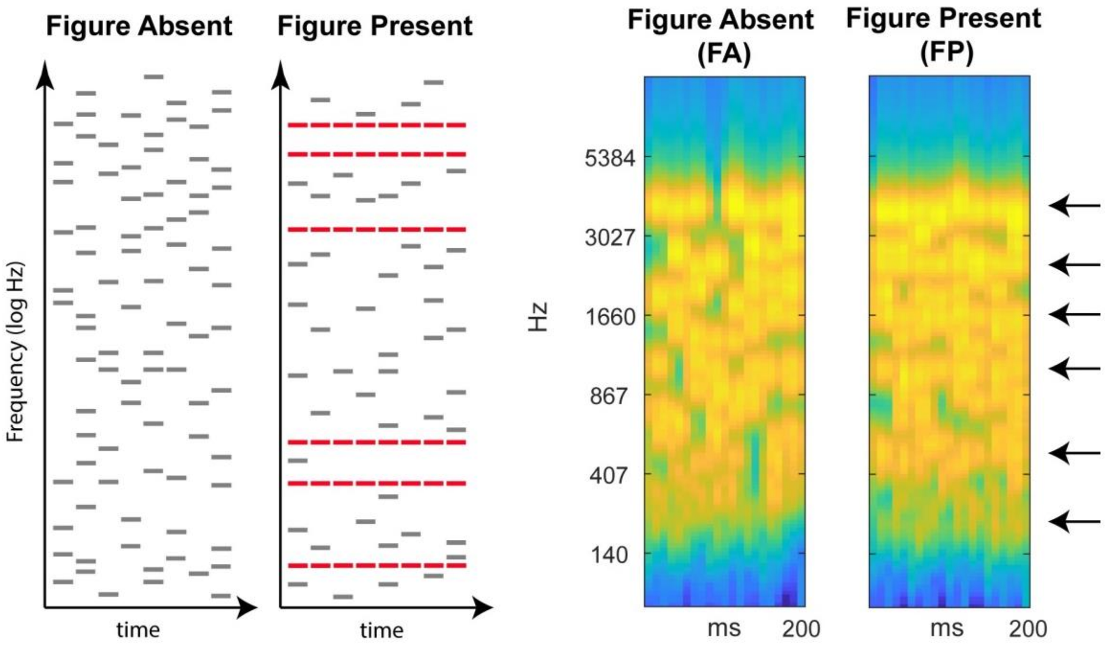


Figure 3.4: Example schematics (left) and spectrograms (right) of the Stochastic Figure-Ground (SFG) stimuli

Stimuli, adapted from Teki et al. (2011, 2013, 2016) consisted of a succession of chords, each comprised of multiple frequency components. The 'figure-absent' (FA) stimuli were chords formed of random frequencies, while the 'figure-present' (FP) stimuli were constrained so that some of the frequencies (indicated in red in the schematic representation and by black arrows in the spectrogram) were repeated in each chord of the stimulus, producing an auditory 'figure'. The example stimuli contain 8 x 25 ms chords, with coherence 6. The spectrograms were generated with a filterbank of 1/ERB wide channels (Equivalent Rectangular Bandwidth; Moore & Glasberg, 1983) equally spaced on a scale of ERB-rate. Channels are smoothed to obtain a temporal resolution similar to the Equivalent Rectangular Duration (Plack and Moore, 1990). This models processing in the auditory periphery to produce a representation of the stimulus similar to that available to the central nervous systems.

The stimuli were varied along several parameters in order to find a stimulus that elicited a strong neural response to the figure, and so could be used in the subsequent investigation of load (Chapter 4). Specifically, I varied the number of coherent frequencies used for the figure (6 or 8); the duration of the chords (25 or 30 ms); and the number of chords (6 or 8); all combinations were tested (8 possibilities). Each FP condition also had a matched FA condition where number and duration of chords, and average number of components were matched, giving 16 stimulus conditions overall. 120 exemplars of each of the conditions were randomly allocated into 4 blocks of 480 stimuli and presented with ISIs of 800 ms. Naïve participants passively listened to the signals while performing an incidental visual task. Prior to the recording, the volume of the stimuli was set to a comfortable level (~70 dB SPL) by each participant.

The visual task was designed to be very low demand. Pictures of landscapes were presented in groups of three (5 seconds per image, fade in and out over 1 second), and participants had to press a button if picture 2 or 3 was the same as picture 1. Instances of repetitions were relatively rare (~1 in 12 sets) so that button pressing was kept to a minimum. This task helped ensure that participants' eyes were open, and they were awake throughout the blocks, but did not place a high demand on processing resources.

At the beginning of the session, a short (4 minutes) 'localiser' block was recorded in order to characterise participants' neural responses to simple auditory stimuli. The measurement consisted of 200 presentations of a 100 ms long, 1 kHz pure tone with ISIs randomly distributed between 700 and 1500 ms. Participants watched a static fixation cross in the centre of the screen and were not required to perform a task.

3.2.3 Analysis

The data from the localiser block were divided into 800 ms epochs, and baseline-corrected using a 200 ms pre-stimulus interval. The auditory M100 onset response (Roberts et al., 2000) was identified for each subject as a source/sink pair in the magnetic-field contour plots distributed over the temporal region of each hemisphere. For each subject, the 40 most activated channels at the peak of the aM100 (20 in each hemisphere) were selected for subsequent sensor-level analysis of the responses evoked by the SFG stimuli.

The data from the main blocks were epoched into 1000 ms trials which covered 800 ms post-stimulus onset, and 200 ms pre-onset. All data were baseline corrected to the pre-onset interval. Epochs with amplitudes above 3pT (~6% of trials) were considered to contain artefacts and discarded. A PCA-based, Denoising Source Separation (DSS, de Cheveigné and Parra, 2014) routine was applied to the data to extract stimulus-locked activity. The 20 most repeatable components were retained and projected back to sensor space. To characterise the response at this stage, the root mean square (RMS) of the evoked field over the localiser channels was calculated for each time point to give a time-series which reflects the instantaneous power of the evoked response. For illustrative purposes, group-RMS (RMS of individual subject RMSs) is plotted (Figure 3.2) but statistical analysis was performed across subjects.

To characterise the elements of the response that are specific to FP stimuli, a further DSS was conducted, this time optimised to find components (spatial filters) which differed maximally between FA and FP trials (calculated over all channels and the whole epoch; de Cheveigné & Parra, 2014). The highest ranked DSS component was retained for the analysis and used as a spatial filter (source model) for the analysis of FA vs FP trials (Figure 3.3, inset). In all cases this spatial filter corresponded to the standard temporal dipolar pattern associated with auditory responses.

FP trials (relative to FA trials) were characterised by a sustained shift in the evoked response towards the direction associated with the M100. Though MEG data does not exhibit positive or negative peaks in the same way as EEG data (since magnetic fields take the form of dipoles), I refer to this direction as negative since the EEG counterpart of the M100 (the N1) has a negative polarity and the signals elicited by auditory figures (e.g., the ORN and ARN) also represent negative EEG components. To quantify the amplitude of the negative shift associated with a figure, the difference between the evoked responses for FP and FA trials was calculated for each participant, and subjected to bootstrap re-sampling (1000 iterations, balanced; Efron and Tibshirani, 1993). The difference was judged to be significant if the proportion of bootstrap iterations which fell above/below zero was more than 95% (i.e., $p< .05$) for 15 or more adjacent samples (25 ms). The bootstrap analysis was run over the entire epoch duration (200 ms pre onset to 800 ms post onset); all significant intervals identified in this way are indicated in the figure. Similar analyses were run to investigate the effects of different figure parameters: for each set of matched FP and FA trials (e.g., Coh 6, Num 8, Dur 25 ms) the FP-FA difference waves were calculated for each subject; conditions were then grouped according to one parameter, e.g., Coh 6 vs Coh 8 and a bootstrap analysis was run as described above.

Since all stimulus conditions showed a clear response to the figure, the data were collapsed over all parameters for the source analysis. Sources were estimated using multiple sparse priors (MSP, Litvak and Friston, 2008) analysis. Inversions were based on all MEG channels and used a single shell head model with group constraints. Second-level analyses consisted of paired t-contrasts to compare activation between FP and FA conditions (over all stimulus parameters). Results were thresholded at $p < .001$ at the peak level and then subjected to a whole brain $p < .05$ FWE correction at the cluster level. In one instance a small-volume correction (a 10mm diameter sphere centred at the peak of the cluster) was applied instead, since the cluster was small but in a location consistent with previous fMRI (Teki et al. 2011) and MEG (Teki et al. 2016) sources for similar SFG stimuli. The use of a different correction is marked in Table 3.1.

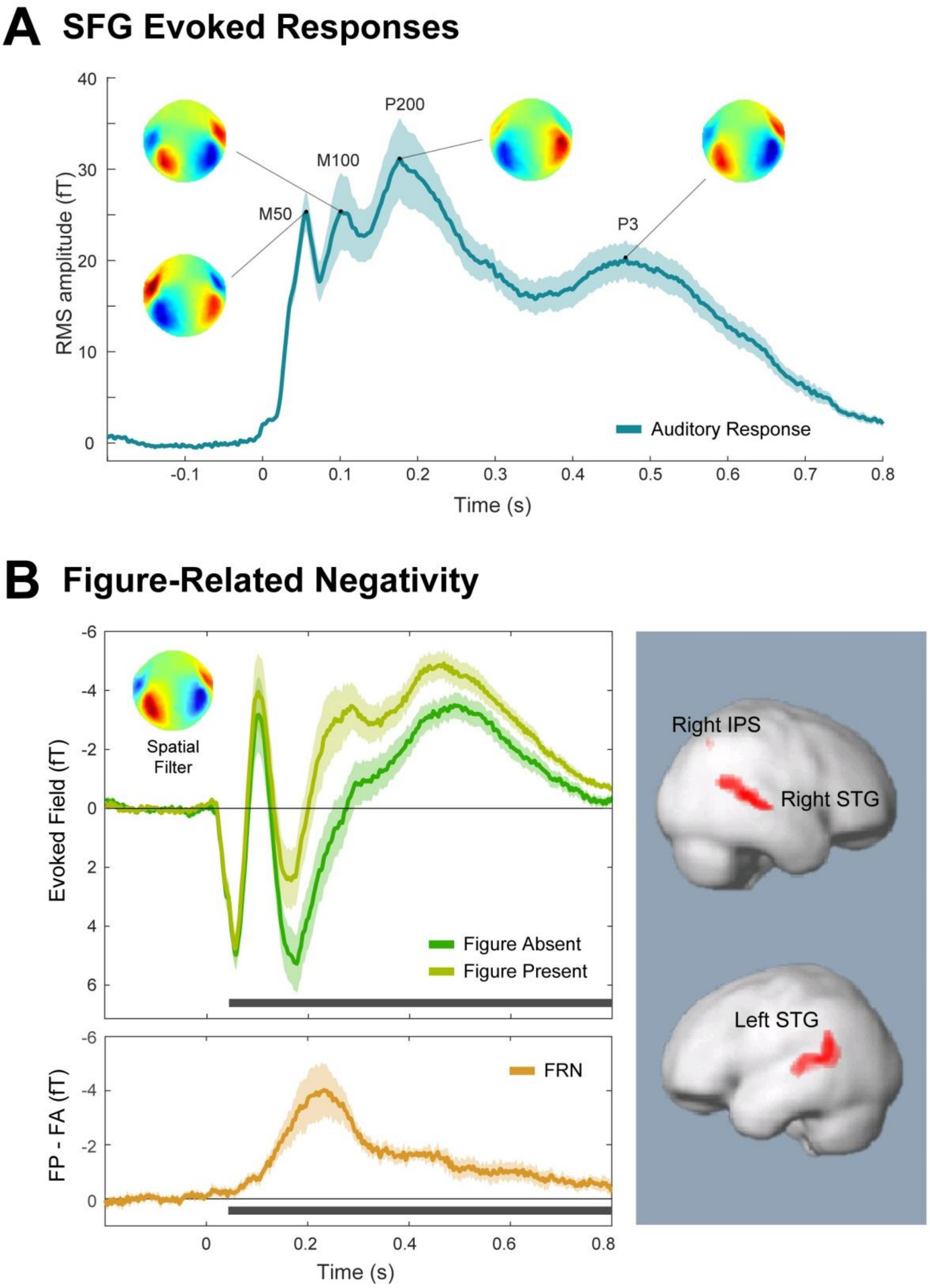
3.3 RESULTS

3.3.1 Response to Figure

Figure 3.5 A shows the average evoked response, collapsed over all stimulus parameters and over figure-present and figure-absent trials. These data are comprised of 20 components (i.e., reflect the independent activity of many processes), and so their dynamics are summarised by calculating instantaneous power (RMS; see methods) across channels. Visible is the characteristic succession of onset peaks (M50, M100, P200 at 50, 100 and 200 ms post onset), followed by a P3 response from ~400-700 ms post stimulus onset.

Figure 3.5 (Overleaf): Experiment 3 - SFG Evoked Responses.

A. Overall response to the SFG stimuli (collapsed over FA/FP conditions and all stimulus parameters). *Plotted is the mean RMS (instantaneous power) of stimulus-evoked activity collapsed over all conditions. Associated scalp topographies at major peaks are provided. Error bars reflect the standard deviation of bootstrap resamplings.* ***B. Figure-related negativity.*** *Evoked fields calculated using a spatial filter which maximised the difference between FP and FA responses (inset – see methods). Evoked fields in this and*



subsequent figures are plotted with the M100 as an 'upwards' peak, to match the convention used in EEG for its counterpart, the N1. Error bars reflect the SD over bootstrap resamplings for each condition, and significant differences between the conditions are marked by horizontal black bars at the bottom of the plot. Top panel shows the evoked responses separately for figure present (FP) and figure absent (FA) stimuli, while the bottom panel illustrates the figure-related negativity (FRN), calculated by the difference timeseries of FP-FA. Right panel shows source-level contrast: regions where FP trials showed greater activity than FA trials are indicated in red. No regions were found to be significant for the opposite contrast (FA>FP).

To identify neural activity that was specific to the figure, a source separation analysis was run to isolate activity that was most different between figure-present (FP) and figure-absent (FA) trials. The resulting spatial filter (pictured in the inset) was applied to the data to produce the time series in Figure 3.5B. The response to the FP trials relative to FA trials is characterised by a sustained ‘negativity’ (i.e., in the same direction as the M100 peak) which was significant from 60 ms post stimulus onset and persisted throughout the rest of the trial. I refer to this response as the ‘figure-related negativity’ (FRN); the difference timeseries which represents the amplitude of the FRN throughout the trial is illustrated below the main plot in Figure 3.5B.

Table 3.1: Experiment 3 – Source Contrast for Effect of Figure

*Source estimates for the difference between FP and FA trials. * indicates a small volume correction.*

Cluster		***Peaks***				
Cortical Structures	***p***(FWE-corr)	***t***	***p***(uncorr)	Co-ordinates (x,y,z)		
Left Temporal Lobe:	<.001	5.12	<.001	-58	-42	10
Superior Temporal Gyrus, BA 22		4.63	<.001	-58	-52	24
Right Temporal Lobe:	< .001	5.72	<.001	56	-42	10
Superior Temporal Gyrus, BA 22						
Right Parietal Lobe:	0.026*	3.54	<.001	34	-66	46
Superior Parietal Lobule, Inferior Parietal Lobule, BA 7		3.32	0.001	34	-62	48

Source localisation revealed several brain regions where activity differed significantly between FP and FA trials (Fig 3.5B, Table Table 3.1:3.1). FP stimuli showed greater activity in bilateral superior temporal gyri and right superior and inferior parietal lobules. This activity is consistent with the findings of Teki et al. (2011, 2016) that the SFG stimuli evoke figure-specific activity along the superior temporal planes, superior temporal sulci, and also within the intraparietal sulci.

These results confirm that there is a measurable neural response to the presence of the figure even during very brief SFG signals, consistent with previous behavioural reports (Teki et al, 2013), and despite the fact that the sounds were not explicitly attended. The fact that a response to the figure can be seen within 60 ms of scene onset (after 2 chords), suggests a very rapid figure-ground segregation process.

3.3.2 Effects of Stimulus Parameters

The parameters I varied within the stimuli created 8 conditions, each with a matched set of FP and FA trials. Here I assess how manipulating each parameter affected the FRN. I calculated the FRN (i.e., FP – FA) for each subject and condition, and then grouped the eight conditions according to each of the three parameters that were varied (coherence, duration of chords, and number of chords). The figures illustrate the average FRN (equivalent to the orange FRN plot in figure 3.5B) with data split by parameter value. Since these FRNs are based on fewer trials than the previous figure, the results are somewhat noisier.

3.3.2.1 Effect of coherence

Overall, trials where the figure contained 8 coherent components resulted in a larger FRN than those with 6 coherent components (see Figure 3.6). The effect was significant from ~150 - 300 ms post stimulus onset, during the main peak of the FRN, and later from ~400-450 ms. This is consistent with the previous MEG data from Teki et al (2016), and EEG data from Toth et al. (2016) which showed that increased coherence led to stronger evoked responses to the figure in this timeframe.

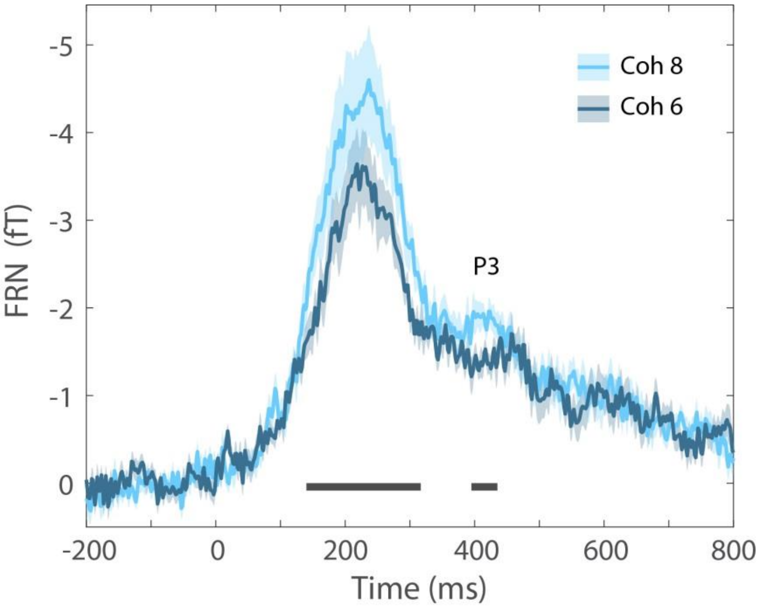


Figure 3.6: Experiment 3 - Effect of Coherence

Plotted are the FRNs (FP-FA) elicited by stimuli containing a figure of 6 vs 8 coherent components. Error bars reflect the SD over bootstrap resamplings for each condition, and significant differences between the conditions are marked by horizontal black bars at the bottom of the plot.

3.3.2.2 Effect of duration and number of chords

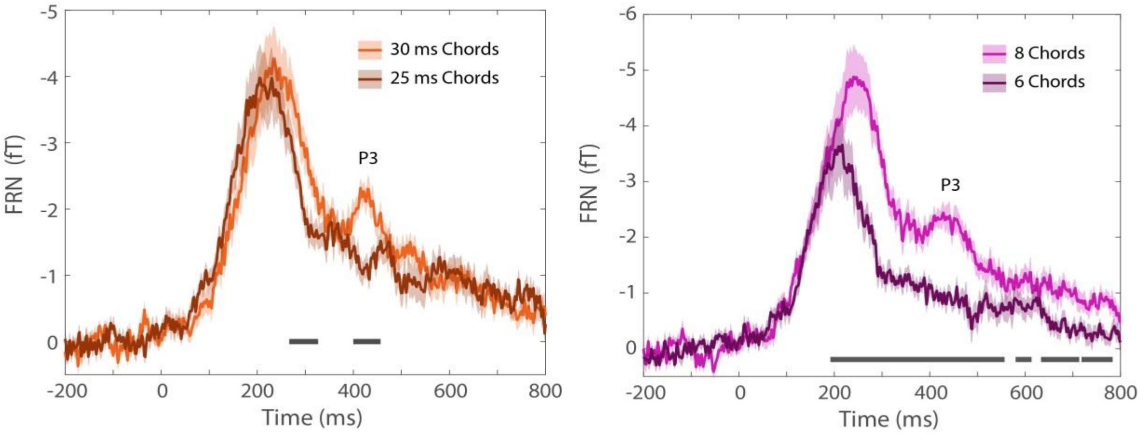


Figure 3.7: Experiment 3 - Effect of Duration and Number of Chords

Plotted are the FRNs (FP-FA) split by chord duration (left) and number of chords (right). Error bars reflect the SD over bootstrap resamplings for each condition, and significant differences between the conditions are marked by horizontal black bars at the bottom of the plot.

Figure 3.7 illustrates the effect of chord duration (left) and the number of chords (right). Increasing either factor resulted in a longer duration of the overall stimulus, and this is reflected in the fact that the FRN extended later into the epoch for both 30 vs 25 ms chords, and for stimuli with 8 vs 6 chords. When chord duration was manipulated, there was an effect on the latency of the FRN, with the shorter chords leading to a slightly earlier peak. This presumably reflects the fact that with shorter chords, information about the stimulus composition is accrued faster. In contrast, when the number of chords was manipulated, the FRNs show the same onset slope, but with a greater number of chords the FRN continued building for longer, leading to both a longer peak latency and a larger peak amplitude. This is consistent with previous work which suggests that rather than the duration of stimuli overall, the crucial factor governing both the perceptual salience and strength of the neural response to the figure is the number of chords presented (Teki et al., 2011; 2013), since the figure response builds up as information is accrued.

3.3.2.3 P3 response

Interestingly, all three parameter contrasts resulted in a significant effect from ~ 400 – 450 ms, which corresponds to the onset of the P3 response (see peak labels in figure 3.5A). In each case, the parameters that would produce the more salient figure (i.e., greater coherence, more chords) or simply a longer stimulus resulted in a larger P3 peak. This suggests that stimuli with a salient figure or those which were overall longer had a greater chance of capturing attention (which was not tightly controlled by the task used during these recordings). This P3 peak is less apparent in the grand mean FRN (when collapsed over all stimulus conditions - figure 3.5B), presumably because this average contains responses to both ‘stronger’ and ‘weaker’ figures, and perhaps also due to the mismatch in latencies of the P3 peak for different stimulus parameters, leading to a more diffuse response when averaged.

3.4 DISCUSSION

3.4.1 FRN as a marker for auditory segregation

In spite of the substantial reduction in the overall duration of the SFG stimuli relative to that used in previous work (Teki et al., 2011; O'Sullivan et al., 2015; Teki et al., 2016), there was still a clear neural response to the figure in all stimulus conditions. This is consistent with behavioural evidence (Teki et al., 2013) showing that listeners are sensitive to figures in brief SFG signals. The response to the figure was characterised by a negativity relative to the figure absent signals, which I term the figure-related negativity (FRN). The FRN emerged very early in the evoked responses, appearing after just two chords had been presented in spite of the fact that auditory signals were behaviourally irrelevant to participants. This suggests that the FRN reflects a swift segregation mechanism that operates early during processing and without explicit intent. The FRN also persisted throughout the evoked response, suggesting that it may also reflect a sustained response to the figure such as those reported in previous studies that have used protracted, temporally dynamic stimuli (Elhilali et al., 2009a; Teki et al., 2016). The FRN extended into the P3 timeframe and there was tentative evidence that there was a resurgence of the response during that period, particularly for signals that contained a strong figure, suggesting that some of the later activity captured within the FRN may relate to attentional shifts towards the auditory signals when figures were present. The amplitude of the FRN was also sensitive to stimulus parameters that are known to affect the strength of the figure percept (coherence and number of chords), indicating that it provides a suitable metric of figure-ground segregation.

3.4.2 Relating the FRN to similar evoked components

The FRN seen here is generally reminiscent of the object-related negativity (ORN) described earlier in this chapter, which occurs in response to a mistuned harmonic within concurrently presented sound components; both responses show a negative valence, have similar peak latencies, and are evoked even during passive listening. Indeed, Toth et al. (2016) refer to the EEG negativity they recorded in response to figures within SFG signals as an ORN. However, since the process underlying concurrent sound segregation based on harmonicity and segregation of on-going sequences may rely on distinct underlying mechanisms (Alain and Izenberg, 2003; Dyson et al., 2005; Micheyl and Oxenham, 2010), and since in subsequent chapters I illustrate a notable difference between the ORN and FRN, I will maintain the use of distinct terminology.

The FRN also shares characteristics with the awareness-related negativity (ARN) reported by Gutschalk et al. (2008) in their informational masking study. In contrast to the FRN, the ARN was only seen when participants successfully detected the target within the auditory scene, which the authors interpret as meaning that active awareness of the figure is required to elicit the response. However, given the similarity between the stimuli found to elicit the ARN and those used here, it is possible that these responses are driven by the same process. The signals used by Gutschalk et al. contained a target which was equivalent to a figure comprised of only one frequency component (i.e., coherence 1). Previous work on SFG stimuli has shown that the neural response to figures deteriorates as the number of coherent components is reduced (Teki et al., 2011; Teki et al., 2016; Toth et al., 2016) , and that detection of figures with only one frequency component is very poor (Teki et al., 2013). Thus, it may be that the ARN is actually an FRN to a weak signal, and thus only becomes detectable above the noise floor after amplification via active attention. This is in keeping with a combination of bottom-up and top-down influences of attention on segregation, such as those described by Elhilali et al. (2009a) and Alain et al. (2001) and Gutschalk at al. (2008).

3.4.3 Neural sources associated with auditory segregation

Source localisation of the FRN suggested bilateral superior temporal generators and an additional right parietal source. This is consistent with previous neuroimaging work which has implicated Planum Temporale (PT), the Superior Temporal Sulcus (STS) and Intra-Parietal Sulcus (IPS) in the process of detecting SFG figures (Teki et al., 2011; Teki et al., 2016). The involvement of PT and IPS in segregation also correspond to previous results that were based on simpler streaming signals (Cusack, 2005; Gutschalk et al., 2007; Wilson et al., 2007; Schadwinkel and Gutschalk, 2010). Current computational models of auditory scene analysis apply cross-channel correlation at a variety of temporal delays and are hypothesised to be supported by rapid adaptive processes in auditory cortex (Elhilali et al., 2009b; Teki et al., 2013; Krishnan et al., 2014; Lu et al., 2017). However, the involvement of sources outside typical auditory cortical regions suggests that the relevant computations are elaborated along a broad processing hierarchy; Teki et al. (2016) suggest that PT operates as a hub for the process that computes the coherence maps, while IPS is involved in encoding the signal as consisting of several sources. The FRN recorded here, which emerged very early and extended throughout the evoked response, may reflect a combination of these processes which occur throughout the hierarchy.

3.4.4 Conclusions

In summary, the FRN characterised here appears to be a robust response which reflects both early computations underlying figure segregation in auditory cortex and later activity that may be driven by a sustained response to the figure (e.g., in IPS) or awareness-related activity. Its amplitude is sensitive to the strength of the figure percept, and it is elicited even in naïve listeners who are not explicitly attending to the sounds. Since any response to the figure in the SFG signals indicates that underlying temporal coherence computations have occurred, the FRN should serve well as a neural index of the level of computation occurring in the auditory system for the next series of experiments.

CHAPTER 4: FIGURE-GROUND SEGREGATION UNDER LOAD[3]

4.0 MOTIVATION

Having established short SFG stimulus as a good model for computations within the auditory system, and the FRN as a sensitive index of processing, I turn to the question of whether this type of processing is dependent on attention and the availability of resources.

4.1 INTRODUCTION

The extent to which auditory figure-ground segregation is dependent on attention has been a longstanding issue in hearing research (Alain and Arnott, 2000; Carlyon et al., 2001; Alain and Izenberg, 2003; Macken et al., 2003; Cusack et al., 2004; Dyson and Alain, 2004; Dyson et al., 2005; Sussman, 2007; Lipp et al., 2010; Shamma and Micheyl, 2010; Thompson et al., 2011; Snyder et al., 2012; Weintraub et al., 2014; Lakatos et al., 2016; Masutomi et al., 2016; Lu et al., 2017), but despite decades of debate the answer has remained elusive.

Most previous work has focused on the effect that the top-down allocation of attention has on segregation. Typically, this is explored using a contrast between conditions where participants intentionally listen for an auditory object (‘active attention’), and conditions of either passive listening or where top-down attention is

---

[3] Data from this chapter have been published in: Molloy, K., Lavie, N., & Chait, M. (2019). Auditory figure-ground segregation is impaired by high visual load. Journal of Neuroscience, 39:1699-1708. Although I have rewritten much of the text so that it works in concert with the other chapters in this thesis, there may be some repetition of information.

allocated to a different stimulus. These studies have reliably shown that auditory segregation can be improved by active attention, with ERP measures of segregation being enhanced during an active task based on segregating the auditory stimuli. It is also clear that once auditory objects have been separated, attention can act to enhance processing of one object and supress processing of others (e.g. Woods and McDermott, 2015), a process that is usually characterised by enhanced neural responses to the attended sound relative to other sounds in the scene, and by the phase-locking of neural oscillations to patterns within the attended signal (Bidet-Caulet et al., 2007; Ding and Simon, 2012; Woods and McDermott, 2015). However, the question of whether attention is strictly necessary for segregation has been much more controversial.

4.1.1 Is attention required for segregation to occur?

4.1.1.1 Segregating simple tone sequences

Even for the relatively straightforward signals used in the ABA- streaming paradigms, there appears to be contradictory evidence as to whether streaming can occur in the absence of directed attention. As I reviewed in Chapter 1, early behavioural studies of attention and streaming, based on participants reporting their percept, suggested that there is very little build-up of streaming in situations where attention is not engaged on the ABA- signals (Carlyon et al., 2001), and that any streaming that does build up while attention is focused on the signals rapidly decays as attention is switched away (Cusack et al., 2004). Similar effects have been shown by studies that used objective measures of streaming rather than behavioural reports: Thompson et al., (2011) and Billig & Carlyon (2016) tested detection of a temporal irregularity within the ABA sequence. This task is facilitated when the signal is perceived as one stream, and so can act as an index of the degree of segregation: performance on the task was initially high while the signal was fused, but deteriorated as streaming built up. These experiments illustrated that the build-up of streaming was very slow when attention was directed away from the stimuli (Thompson et al., 2011; Billig & Carlyon, 2016). They also confirmed that during active attention, any switches away from the streaming signals caused an abrupt disruption of streaming, though they noted that if attention was brought back to the streaming signals the build-up of streaming occurred at a faster rate, suggesting that while attention switches did have a substantial effect on streaming, it was not fully reset (Billig & Carlyon, 2016).

In contrast to this set of findings, other experiments appeared to provide evidence that ignored signals can be streamed. For instance, using the 'irrelevant sound effect' paradigm Macken et al. (2003) probed the degree to which ignored ABA- signals interfered with an active short-term memory task. The task – in which participants maintain an echoic representation of a series of digits – is disrupted if irrelevant sounds that are presented during the rehearsal period are alternating sequences rather than multiple separate, static sources. ABA- signals were presented while participants performed the memory task at two different presentation rates (a factor that is known to affect the build-up of streaming when sounds are attended). Macken at al. argued that if streaming were dependent on focused attention, the irrelevant sounds should remain fused regardless of presentation rate, and so would be equally disruptive to the attended task. However, the data illustrated that a high presentation rate (which encourages streaming) led to fewer errors in task recall, suggesting that streaming had occurred even though the sounds were outside the explicit focus of attention.

Similar conclusions have been drawn from MMN studies, in which ignored sound sequences elicited MMNs in response to deviance from patterns that can only become apparent once sounds have been segregated into two on-going streams. For example, Sussman et al. (1999) presented sequences of six tones (three high and three low, but all with different exact frequencies) in an alternating high-low pattern where the tones in each band typically followed an ascending pitch pattern (L1 H1 L2 H2 L3 H3), but where occasional pattern deviants (descending pitch order) occurred within either the high or low band. At slow presentation rates (where streaming based on broad pitch bands is unlikely) there were no MMNs to the deviant sequences, but at higher presentation rates that encourage streaming, the deviant sequences elicited MMNs. These studies highlight that certain endogenous cues such as the presentation rate and the frequency separation between tones, which encourage faster build-up of streaming when active attention is engaged, may be sufficient to induce segregation even for ignored signals. However, as I highlighted in the introduction to this thesis, it is not clear from these previous studies whether this effect is purely stimulus-driven or if it is only possible only when spare resources spill over from the attended task.

4.1.1.2 Segregating specrto-temporally complex signals

The question of whether attention is necessary for segregation has also been approached using more complex segregation stimuli. For instance, Elhilali et al. (2009a) explored auditory segregation using the informational masking paradigm (described in Chapter 3 Section 3.1.3 and illustrated in Figure 3.1) in which targets comprised of a single regularly repeated frequency were presented amongst a broad-spectrum dynamic tone cloud. They used MEG to assess the strength of neural activity that was timelocked to the presentation rate of the target, contrasting conditions where participants engaged in a task that was based on the target within the stimulus with conditions where the task was based on the background. Both the power and the degree of phase-locking across MEG channels within the steady-state response (SSR) was shown to be increased when participants were attending to the target rather than the background (Elhilali et al., 2009a). This effect correlated with behaviour: the SSR built up over time in a manner that correlated to performance on the target task, and additionally, participants who showed a stronger modulation of the SSR with attention also performed better on the target task. There was also an effect of the low-level saliency of the target, with more salient targets showing superior power and phaselocking of the SSRs, regardless of the attentional state. As reviewed in Chapter 3 (Section 3.1.3) Gutschalk et al. (2008) used the same stimuli to contrast neural responses between conditions of active and passive listening, and for targets which were detected versus missed. However, unlike the results from Elhilali et al.'s study, the target-driven SSRs recorded by Gutschalk et al. did not show any modulation according to attentional state. This difference may reflect that in the passive condition used by Gutschalk et al. resources were automatically allocated to the auditory signals, while in the 'ignore' condition used by Elhilali et al. (2009a) participants were actively engaged in a task that required attention towards the background signals (which did not evoke the SSR). In addition to the SSR responses, Gutschalk et al. also reported an attention-related negativity (ARN) that occurred only for targets that were detected. As discussed in Chapter 3, this may bear some relation to the FRNs reported here that are elicited by figures in the SFG signals. Both Gutschalk et al. and Elhilali et al. concluded that there are bottom-up and top-down influences on segregation of the informational masking signals, with the implication that some level of stimulus-driven segregation is purely automatic.

For the SFG signals, there is also substantial evidence that active attention is not necessary in order for segregation to occur. For instance, Teki et al. (2011, 2016) have shown that participants who performed a visual task while irrelevant SFG stimuli were presented in the background showed clear sustained responses (recorded via MEG and fMRI) that occurred only when the figure was present, and which scaled with the coherence and duration of the figure. Similarly, O’Sullivan et al. (2015) identified an ERP component of the EEG response that occurred from ~100-200 ms post stimulus onset which correlated in amplitude with the coherence of the stimuli. This component was elicited under passive listening conditions and was enhanced (in both amplitude and temporal extent) with active attention. They draw similar conclusions to those based on the informational masking stimuli: that some segregation occurs automatically but that it is enhanced by active attention.

4.1.2 Applying Load Theory

As argued in the introduction to this thesis, the paradigms that have been used to study whether attention is required for segregation are not designed appropriately to distinguish between processes that are pre-attentive and those which occur during passive listening due to the automatic application of spare capacity. Drawing the conclusion that some segregation occurs automatically, based on evidence that segregation occurs during passive conditions, is not necessarily valid; if the auditory system benefits from spill-over of capacity, as is established to occur in vision, resources would only be fully withdrawn from task-irrelevant stimuli when a separate, explicitly attended task involves a sufficiently high level of perceptual load to exhaust all capacity. Thus, the streaming effects reported to occur to unattended signals may have occurred only because the tasks participants were engaged in did not heavily load perceptual resources. This seems plausible for the studies described above; for example, in the study based on the irrelevant sound effect (Macken et al., 2003), the active task was maintenance of a series of digits, a process known explicitly to load cognitive resources and result in enhanced processing of irrelevant stimuli. In most other studies there was no active task, merely the expectation that attentional resources are not engaged unless explicitly directed.

4.1.2.1 Impact of Load on Concurrent Sound Segregation

To examine whether auditory processes do benefit from the automatic allocation of resources when available, a load task is required to manipulate the perceptual demand of an attended task, so that the resources available to the auditory system are modulated. This approach has already been applied to measures of concurrent sound segregation. The ORN, described in detail in Chapter 3, is a response elicited when a single component is perceived as segregated from a concurrent (static) harmonic sound due to low-level segregation cues such as a mistuning or onset asynchrony. Unlike later components elicited during active performance of a segregation task (e.g., P230, P400), which appear to reflect target selection, the ORN occurs readily during passive listening (Alain et al., 2001; Alain et al., 2002; Lipp et al., 2010). Investigating whether load might affect the elicitation of the ORN, Alain and

Izenberg (2003) presented harmonic complexes to both ears, with a load task based on signals in one ear while sounds in the other ear were to be ignored. The stimuli in both ears were either tuned or mistuned (50% each), and occasional deviants (of shorter duration) occurred. The task was to detect targets that had a shorter duration. For the low load task, participants pressed one button for all duration deviants; in the high load task, they had to make a further judgement of whether the target was tuned or mistuned, pressing different buttons for each. EEG recordings revealed that duration-deviant stimuli in the unattended ear elicited a mismatch wave that was reduced as task load increased. In contrast, the ORN elicited by the mistuned stimuli was not significantly affected by task load. Although this result suggests that the ORN is preattentive, sounds in the active task ear were never presented at the same time as sounds in the 'ignore' ear – ISIs varied from 200-400 ms – so it is unclear whether the load of the active task was timed sufficiently well to critically disrupt processing of sounds in the unattended ear.

The ORN has also been investigated using a visual load paradigm. Dyson et al. (2005) presented tuned and mistuned harmonic complexes to participants who were actively engaged in a visual task that was easy or hard. The visual task was based on a serial presentation of coloured numbers onscreen. In the easy task participants categorised each digit as greater than or less than 5 (5 itself was not used in the task). For the hard task, if the number was in blue participants categorised the current number as > or < 5, while for red numbers, participants categorised the *previous* number as > or < 5. The ORN responses recorded to the auditory stimuli were not significantly different between visual task conditions (though note that ORNs were smaller under the hard task conditions, and *p* values approached the significance criterion of *p* < .05). Though this result again suggests that the ORN is not susceptible to the depletion of resources, the same criticism I described in the previous paragraph applies to this study, since the timing of the visual presentations (random ISIs of 1500-2000 ms) was not matched to the presentation of the sounds (random ISIs from 200-400 ms). Further, the visual task did not strictly load perceptual resources: it was in essence an n-back categorisation task, which would typically be classed as cognitive load, and further, a 1-back task is not usually considered to be high load. Thus, although there is some evidence that concurrent sound segregation (in particular the elicitation of the ORN) is robust to the depletion of resources, it may be that experiments have not adequately controlled load, especially in terms of ensuring resources are depleted at the time when the segregation stimuli are being processed.

4.1.2.2 Impact of Load on Segregation of Sequences

One study has employed load to explore whether segregation of sound sequences can occur when resources are depleted. Masutomi et al. (2016) created complex dynamic stimuli based on natural sounds; the background was constantly changing, while targets were formed of ten repetitions of a 320 ms segment. Segregation of this sort is driven solely by the repetition within the target and must therefore occur via some form of temporal integration (Andreou et al., 2011; McDermott et al., 2011; Bendixen, 2014). Segregation was assessed by presenting a probe sound (320 ms) at the end of each trial which either matched the segment that was repeated to form the target or was a section of the background (i.e., a plausible but incorrect probe). Masutomi et al. ran a series of experiments where this auditory segregation task was paired with a concurrent, prioritised task of low or high demand. Three load tasks were used: a multiple object tracking (MOT) task, which loads visual perceptual resources; a digit sequence memory task, which loads visual encoding and short-term auditory memory resources; and an auditory enumeration task (counting a series of tone pips), which loads auditory perceptual and short-term memory resources. For the MOT task there was no significant effect of load on auditory segregation performance. For the digit sequence and auditory enumeration tasks there was a significant effect of load on performance of the segregation task, but this was not statistically different from the effect of load on a separate auditory task that was not based on segregation. The authors concluded that, while there may have been an overall effect of load on auditory processing, it was not specific to auditory segregation. One limitation of this series of experiments is that in all cases the segregation stimuli were task-relevant, and thus any effect of load would be expected to be small. Additionally, even in the high load conditions, performance on the load tasks was very good (d’ > 2.5) suggesting that resources may would not have been fully depleted. Finally, as I have discussed previously, it is difficult to be sure that load is maintained at all times for stimuli that are presented over extended periods. Here, segregation stimuli were all several seconds long, and though the load task stimuli were presented over the same period, it is not clear how levels of load may have varied throughout stimulus presentation.

4.1.3 Chapter Overview

In the present chapter I used a visual load task to explore whether figure-ground segregation based on temporal coherence is dependent on the availability of processing resources. Importantly, unlike the load studies described previously, I used a visual load task that was precisely timed to the presentation of SFG stimuli in order to maximise the depletion of resources. I present data from two MEG studies (Experiments 4 and 5) which assess auditory processing of task-irrelevant sounds, using two different visual load tasks, and a complementary behavioural dual task study (Experiment 6) which illustrates the effect of load on deprioritised - but not ignored - auditory signals.

4.2 EXPERIMENT 4: SFG UNDER LOAD – MEG

This experiment used a slightly adapted version of the visual load task used previously in the Inattentional Deafness experiments in Chapter 2. During pilot testing this task appeared to provide a good manipulation of load based on participants' reaction times and performance. However, during the MEG experiment participants performed repeated blocks of the task and showed substantial improvements in performance between blocks. This was clear during the experiments since after each block a participant's percentage correct score was displayed on screen. I was concerned that load was not being manipulated sufficiently well during the majority of the blocks each participant completed and so aborted scanning after 16 of a planned 20 participants had been run. I designed a new visual load task which provided a stronger manipulation of load and ran a full cohort of participants in a second MEG experiment (Experiment 5) which provides the main data for this chapter. Nonetheless, I analysed the data collected as part of Experiment 4, and it illustrates some effects that corroborate the findings of Experiment 5. Thus, I present the data here with the caveat that the sample size was not as large as intended, and that the data were somewhat noisy.

4.2.1 Methods

4.2.1.1 Participants

Sixteen paid participants (9 male; mean age of 24.9 years, SD = 5.1 years) of an intended 20 took part in the MEG experiment. All were right handed, had normal or corrected to normal vision and reported normal hearing and no history of neurological disorders. None had taken part in any of the previous experiments presented in this thesis, or in experiments which used a similar load task.

4.2.1.2 Apparatus, Stimuli and Procedure

The apparatus and recording methods were identical to those described for the MEG experiment in Chapter 2.

As load was introduced as an extra factor in this and the experiments that follow, relative to the experiments presented in Chapter 3, I chose a fixed set of parameters for the SFG stimuli in order to maximise the power of the analysis. Since all stimulus parameters tested in Chapter 3 gave rise to significant FRNs, it seemed that any of these could be suitable, so I selected parameters based on what I believed might be more susceptible to a load manipulation. For the coherence value, I chose to use the smaller number of coherent components (six) to minimise the inherent signal to noise ratio of the coherent versus non-coherent components. For the chord duration, I chose shorter chords (25 ms) so that the auditory information processing would need to occur faster (and might therefore be more susceptible to disruption). Finally, I chose to present the larger number of chords (eight) so that any signal which arose from successfully integrating over multiple chords would be stronger, giving a clearer contrast to conditions where integration was un- (or less) successful. Thus, the stimuli used in this and subsequent experiments had coherence 6, with 8 x 25 ms chords, resulting in 200 ms signals.

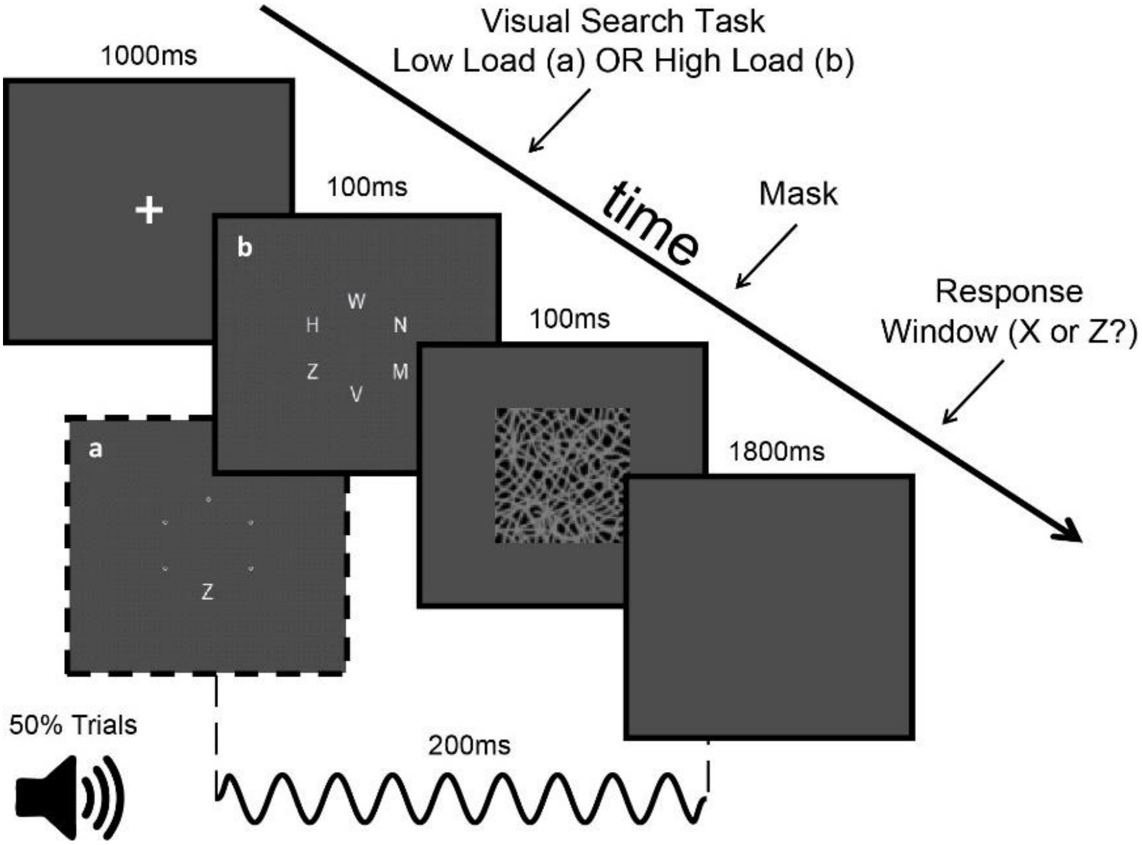


Figure 4.1: Experiment 4 - Load Task Paradigm

Letter search task of low load (a; target dissimilar from distractors) or high load (b; target similar to distractors). Subjects were instructed to identify which of two target letters (X or Z) had been present in the display. Search arrays were presented for 100 ms, followed by a 100 ms mask. Auditory stimuli (which occurred on 50% of trials; no auditory task) were presented at precisely the same time as the visual stimuli, covering the 200 ms during which the search array and mask were displayed on screen.

The visual load task used previously (Chapter 2) was a search for targets within an array of letters that were presented onscreen for 100 ms. Pilot testing indicated that increasing the duration of the visual stimulus to 200 ms (to match the duration of the auditory signals) made the search too easy, so instead the array was presented for 100 ms followed immediately by a 100 ms mask (Figure 4.1). The subsequent response period was 1800 ms (during which the screen was blank). There was no auditory task as this experiment was designed to explore processing of irrelevant sounds under load. The experiment consisted of 8 blocks of 120 trials each, and the order of blocks (low or high load) was counterbalanced between subjects. Feedback was not given within blocks, but at the end of each block participants were provided a score of percentage correct, to help maintain task engagement.

4.2.1.4 Analysis

Behavioural Data

Data from the visual search task were subjected to within-subjects comparisons of mean percentage correct score and reaction times between low and high load. For the reaction time data, a paired t-test was used. For the accuracy data a mixed logistic regression was used to model the proportion of correct responses, with load as a fixed factor and subject as a random factor. Sound was also entered into an initial model to explore whether the presence versus absence of an auditory signal had an effect on trial performance, but since neither the main effect of sound nor the interaction between load and sound were found to be significant, sound was not included in the final model.

MEG Evoked Data

The data from the localiser blocks were analysed as described in previous chapters to give a selection of channels for each participant that best represented the spatial distribution of their aM100 responses. Data from the main blocks were epoched into 1000 ms trials, including a 200 ms pre-onset interval. Pre-processing (base line correction and trial exclusion) of the trials was performed as described in previous chapters, followed by a DSS analysis (de Cheveigné & Parra, 2014) to extract stimulus-locked activity, after which 20 components were retained. RMS and source data reflecting the response to visual-only trials (LL vs HL) were calculated at this stage of the analysis.

To estimate the response to auditory stimuli, a second DSS step was performed to separate auditory responses from the measured auditory-visual combined response (as described in detail for Experiment 2, Chapter 2). Ten components were retained and projected back to channel space. RMS and source data for the overall response to the auditory SFG stimuli (under LL and HL) were calculated at this stage of the analysis.

Finally, as described for the SFG stimuli in Chapter 3, a final DSS analysis was applied to produce a spatial filter that reflected activity which was most different between FP and FA trials. The data were then separated into LL/HL and FP/FA conditions for analysis. Bootstrap analyses to contrast the evoked responses between conditions were run as described previously with significance criteria of $p < .05$ for 15 or more adjacent samples (25 ms).

MEG Source Analysis

A source analysis was attempted, but due to the small sample and the fact that the FRNs recorded here were relatively small and noisy, no sources that reached the significance criteria were identified.

4.2.2 Results

4.2.2.1 Visual Task

Participants showed a significant effect of load on performance in the visual task (Figure 4.2). Increased load led to lower accuracy (Mean: LL = 97.9%, HL = 89.3%; SD: LL = 3.7, HL = 6.7; $z = -6.87$, $p < .001$) and longer RTs (Mean: LL = 597 ms, HL = 761 ms; SD: LL = 83, HL = 79; $t(15) = -9.1$, $p < .001$) confirming that the load manipulation was successful. However, as noted in the introduction to this experiment, the effect was small.

The evoked responses from the visual only trials (Figure 4.3) show two clear onset peaks about 100 ms apart, which represent the vM100 onset response to the search array (which appeared at 0 ms), and to the mask (appearing at 100 ms). The topography plots are consistent with activity in occipital areas. High load responses were significantly stronger than those in the low load task during the vM100 onset

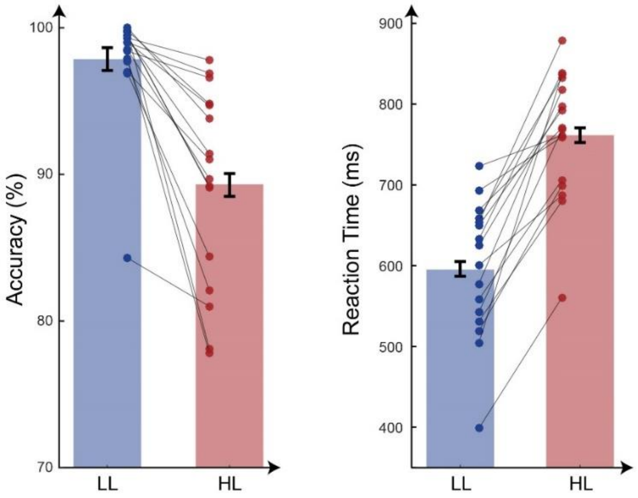


Figure 4.2: Experiment 4 - Visual Behavioural Data

Mean values for accuracy (percentage correct) and reaction times are plotted for low load (blue bars) and high load (red bars). Error bars indicate the standard error of the sample mean, corrected to reflect the within-subjects contrast. Individual data are plotted and connected by grey lines to illustrate change in performance for each participant between low and high load conditions.

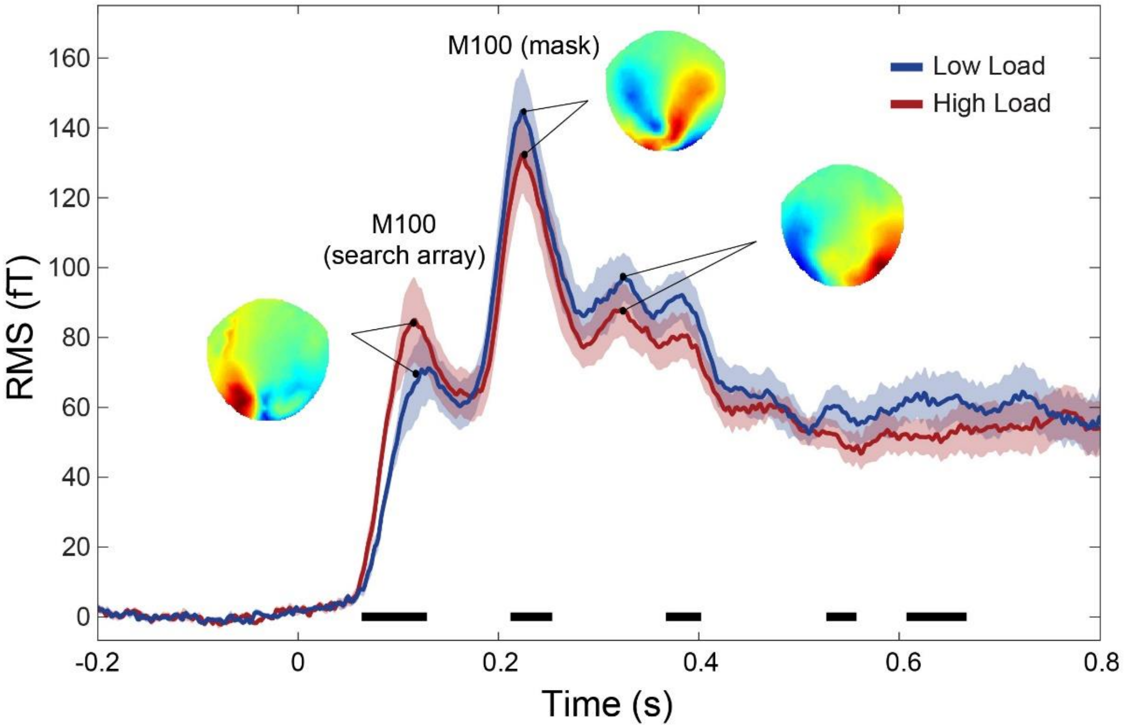


Figure 4.3: Experiment 4 - Visual Evoked Responses

Mean RMS (instantaneous power) of activity in visual-only trials, contrasting responses in LL and HL, with scalp topographies to illustrate polarity at major peaks. Black bars illustrate times when there was a significant difference between conditions.

peak to the search array (63-128 ms). This mirrors the early effect of load seen in the visual MEG data in Chapter 2 (which used the same visual search stimuli), and presumably reflects increased early visual processing under high load. In contrast, during the vM100 peak to the mask (212-253 ms) the visual response was significantly larger under low load than high. Since the mask did not provide any information relating to the visual search (in fact it obscured important information), this effect is consistent with a load theory interpretation that under high load, fewer processing resources were left available to spill over to process the (task-irrelevant) mask. During the latter part of the trial, responses remained larger under low load, with significant periods from 370 ms onwards; this later effect is also consistent with the visual ERFs reported in Chapter 2, and may reflect that responses such as the N2pc and P3 were stronger in low load since task performance was higher (Hopf et al., 2000; Molloy et al., 2015).

4.2.2.2 Auditory Processing

Effect of Load on Overall Auditory Responses

Figure 4.4 shows the auditory evoked responses, collapsed over FA and FP trials, as a function of visual task load. The activity is characterised by the standard succession of auditory response peaks, replicating the pattern observed during passive listening in Chapter 3. Field-maps associated with the major peaks are also plotted. All exhibit the standard dipolar pattern over temporal channels commonly associated with auditory activity. The fact that the shape, timing, and field pattern derived by DSS match very well with those from Chapter 3 confirms that overall, the auditory activity was isolated successfully from the response mixtures.

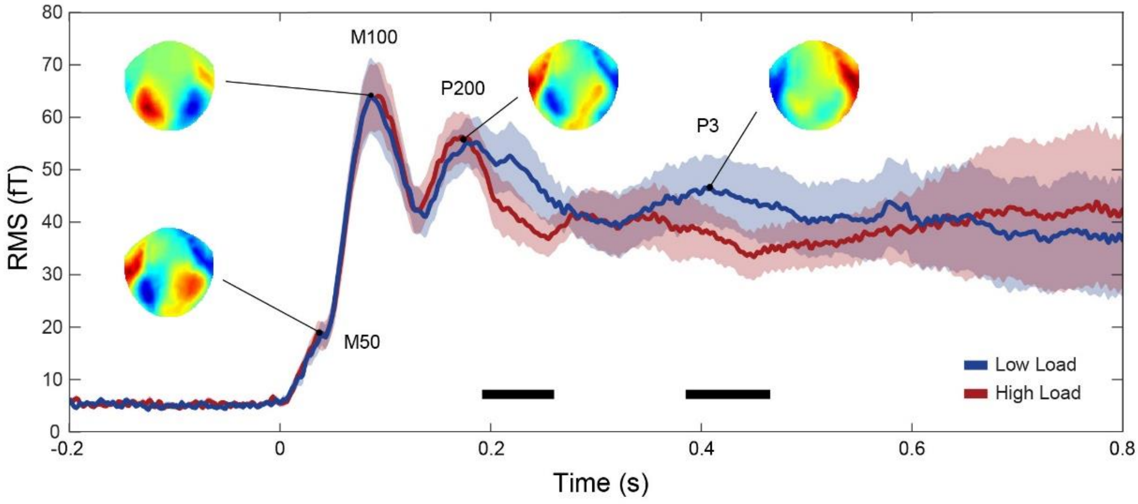


Figure 4.4: Experiment 4 - Auditory Evoked Responses

Mean RMS of responses to the SFG stimuli (collapsed over FA/FP conditions) under LL and HL. Scalp maps are given at peak topographies. Black horizontal bars at the bottom of the plot indicates periods when there was a significant difference between conditions.

There was a significant difference in amplitude between the responses under high versus low load from 192-260 ms (P2 response) and 385-465 ms (during the P3). In both cases the auditory evoked responses were reduced under high load, indicating that high visual load impacted both the latter components of the onset complex to the auditory stimuli and the P3 awareness response. This effect is similar to that shown in Chapter 2 for pure-tone stimuli, though the effect in the present data was not apparent quite as early, affecting only the P2 response rather than the aM100 as seen in Chapter 2.

Effect of Load on FRN

Figure 4.5 A illustrates the auditory evoked responses under low and high load, with the FRN for each load condition shown underneath the main plots. Though the responses presented here are somewhat noisier than those in Chapter 3 (due to the substantially lower number of trials in each condition, necessary to accommodate the load manipulation, and the fact that the auditory responses had been isolated from within the aggregate auditory-visual response), the responses show the same characteristic pattern as those in Chapter 3. There was a small FRN visible in low load that was significant from ~90-240 ms post stimulus onset, and also from 480-520 ms. In contrast, under high load the FRN was completely abolished, suggesting that high visual load impaired the auditory system’s ability to distinguish between FP and FA scenes.

The effect of load was confirmed explicitly by evaluating the interaction between FP/FA and HH/LL conditions: for each subject, a difference time-series (LL (FP-FA) – HL (FP – FA)) was computed and subjected to bootstrap resampling. Figure 4.5 B

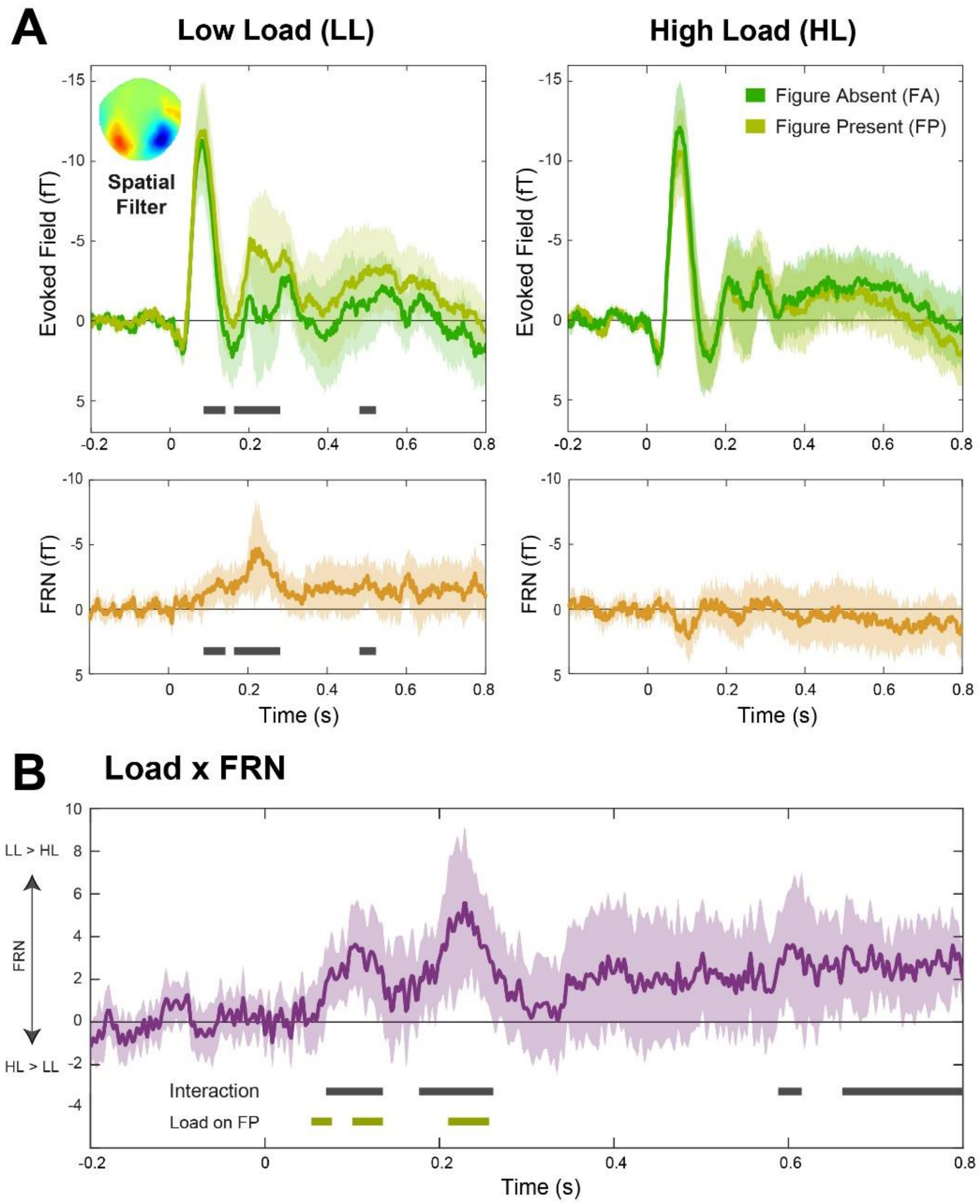


Figure 4.5: Experiment 4 - Effect of Load on FRN.

A: FRN under low load (left) and high load (right). *Top panel shows the evoked responses separately for FP and FA stimuli, while the bottom panel illustrates the FRN (difference timeseries of FP-FA). The spatial filter used to calculate the responses (see methods) is inset. Error bars reflect the SD of bootstrap resamplings for each condition, and significant differences between the conditions are marked by horizontal black bars at the bottom of the plots.* ***B: Load x FRN Interaction.*** *The difference timeseries, LL(FP-FA) – HL(FP-FA), quantifies the interaction between load and figure. Error bars show 2*SD of the bootstrap, for comparison with zero line; periods when the values differed significantly from zero are indicated by the black bars below the plot. Green bars indicate periods when load had a significant effect on responses to FP stimuli (no significant periods were found for the effect of load on FA responses).*

plots the resulting mean difference across subjects, confirming a significant effect of load on the FRN from 70-135 ms, 177-262 ms (both within the main body of the FRN), and from ~600 ms until the end of the epoch. To understand whether the effect of load on the FRN was driven by a load effect on FP trials, FA trials, or both, I compared HL and LL responses for FP and FA stimuli separately. This analysis demonstrated that load only had a significant effect on FP responses (green horizontal lines in Figure 4.5 B); load did not have a significant effect on responses to FA stimuli. This suggests that high load impacted auditory processing that was specifically related to detecting the figure (as opposed to affecting responses to sound overall).

4.2.3 Interim Discussion

In spite of the fact that the full cohort was not run for this experiment, and that a source analysis was not tenable, the ERF data suggest that the neural response to the figure within the SFG signals (the FRN) was reduced under high visual load. The data also suggest that load had a specific effect on evoked responses to FP stimuli, suggesting that the computations underlying figure-ground segregation were specifically affected. However, due to the small sample and the fact that the visual task may not have been a strong manipulation of load, these results should be held as preliminary.

4.3 EXPERIMENT 5: SFG UNDER LOAD – COLOUR-SHAPE TASK - MEG

For this experiment I designed a new visual task that (as established by pilot testing) was more demanding than the letter search used previously. It also offered an opportunity to address a potential confound of the stimuli used in the letter search task which was that the stimuli were not identical between load conditions.

4.3.1 Methods

4.3.1.1 Participants

Twenty paid participants (8 male; mean age of 24.5 years, SD = 4.3 years) took part in Experiment 5. All were right handed, had normal or corrected to normal vision and reported normal hearing and no history of neurological disorders. None had taken part in any of the previous experiments presented in this thesis, or in experiments which used a similar load task.

4.3.1.2 Apparatus, Stimuli and Procedure

The apparatus and recording methods were identical to those detailed for previous MEG experiments in this thesis. Auditory stimuli were as described in Experiment 4.

For the new visual task, I designed a search where the same displays could be used for both high and low load tasks. The visual search arrays consisted of five coloured shapes presented on a dark grey background, spaced equally around a (non-visible) circle centred at fixation and subtending 1.9° viewing angle (Figure 4.6). The five shapes comprised one each of a circle, triangle, square, diamond and pentagon, and the colours were assigned so that there were always two red items, two green, and one either blue or yellow (50% of trials each). A feature vs conjunction visual search task was used to implement different levels of perceptual load (Treisman and Gelade, 1980; Lavie, 1995): for low load the target was any blue shape (feature search), while in high load the targets were a red circle or a green square (conjunction search). Targets were present in 50% of arrays, and counterbalanced so that LL and HL targets did not correlate (i.e., if the LL target was present in an array, the likelihood of the HL target being present was 50%, and vice-versa). The positions of the shapes were pseudo-randomised on each trial so that the target had an equal probability of occurring in each position. Each combination of shape and colour was equiprobable across the stimulus array sets. Search arrays were displayed for 200 ms, to match the duration of the auditory SFG signals.

The session was comprised of 12 blocks (6 low load, 6 high load) consisting of 80 trials each; the order of the blocks was counterbalanced between participants. Feedback was not given during blocks, but at the end of each block participants were provided a score of percentage correct, to help maintain task engagement.

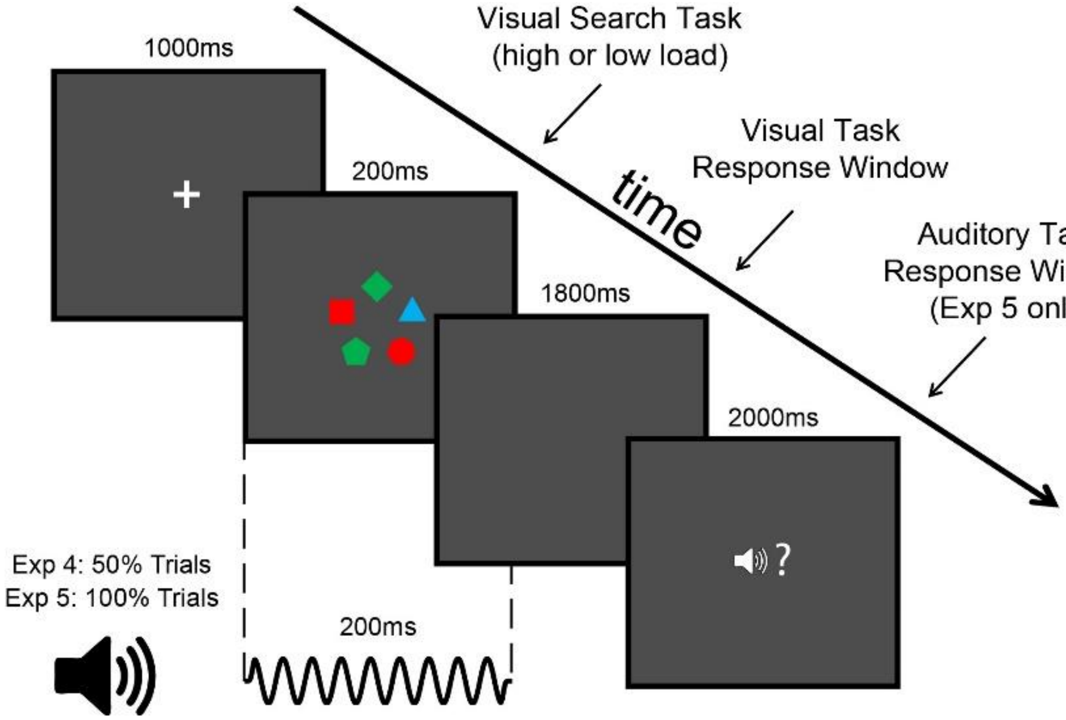


__Figure 4.6: Experiments 5 and 6 - Load Task Paradigm__

Low load task was a colour feature search; high load task was a colour-shape conjunction search. Auditory stimuli occurred on 50% of trials in Experiment 5 (MEG) and 100% of trials in Experiment 6 (Behavioural Dual Task). When present, auditory stimuli occurred at precisely the same time as the visual search array. The response window for the auditory target was displayed only during Experiment 5 when there was an active auditory task.

4.3.1.3 Analysis

Evoked data were analysed exactly as described above for Experiment 4. In addition, due to the better quality of the present dataset, I was able to perform some further analysis to explore both the source of the effects and individual differences.

Individual Differences

To assess the relationship between perceptual load and the process of figure-ground segregation on an individual basis, I ran a correlation analysis. For each participant, the decrement in visual task performance as load increased was quantified by subtracting the mean reaction time under low load from that under high load. The effect of load on the amplitude of the FRN was also calculated for each participant, by subtracting the mean amplitude of the FRN between 50-600 ms in high load from that in low load. A Spearman’s rank correlation analysis was used to assess the relationship between these two factors across subjects.

MEG Source Analysis

Source inversions were calculated using multiple sparse priors (MSP, Litvak & Friston, 2008) analysis. Inversions were based on all MEG channels and used a single shell head model and group constraints. For estimating sources of auditory activity, a soft prior over temporal and parietal areas was used, motivated by previous fMRI and MEG data for SFG stimuli (Teki et al., 2011; Teki et al., 2016), and the source results from Chapter 3, all of which indicate potential sources throughout the temporal lobe and IPS. The prior mask was created in FSLview (http://surfer.nmr.mgh.harvard.edu), based on combining the Harvard–Oxford Structural atlases for all temporal areas, and the Juelich histologic atlas for IPS, with a threshold of 5%. This resulted in a very broad prior, which was binarised so that the strength was equal over all regions. Note that solutions were not restricted to this mask, it served only as a prior to the source algorithm.

Second-level analyses consisted of paired t-contrasts to compare the visual and auditory responses between low and high load, and a full factorial repeated measures F-contrast to model the auditory responses, including main effects of load and figure, and the load x figure interaction and subject entered as a random factor. Results were thresholded at $p< .001$ at the peak level, and the subjected to a cluster-level FWE correction at $p < .05$.

4.3.2 Results

4.3.2.1 Visual Task

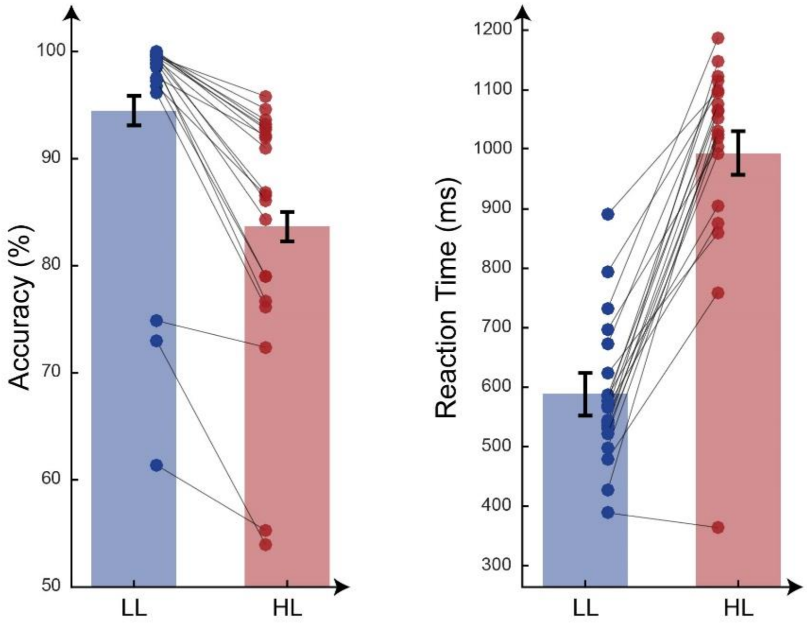


Figure 4.7: Experiment 5 - Visual Behavioural Data

Mean values for accuracy (percentage correct) and reaction times are plotted for low load (blue bars) and high load (red bars). Error bars indicate the standard error of the sample mean, corrected to reflect the within-subjects contrast. Individual data are plotted and connected by grey lines to illustrate change in performance for each participant between low and high load conditions.

A significant effect of load on performance in the visual task was observed (Figure 4.7). Increased load led to lower accuracy (Percentage Correct: *Mean*: LL = 94.5%, HL = 83.7%; *SD*: LL = 10.7, HL = 11.9; $z = -5.20$, $p < .001$. d'*: Mean:* LL = 4.3, HL = 2.5; *SD*: LL = 1.3, HL = 1.0; $t(19) = 9.9$, $p < .001$) and longer RTs (*Mean*: LL = 594 ms HL = 1000 ms; *SD*: LL = 128, HL = 184; $t(19) = -10.5$, $p < .001$). This suggests that the load manipulation was successful, and that, as hoped, the effect was larger than that seen using the previous visual task.

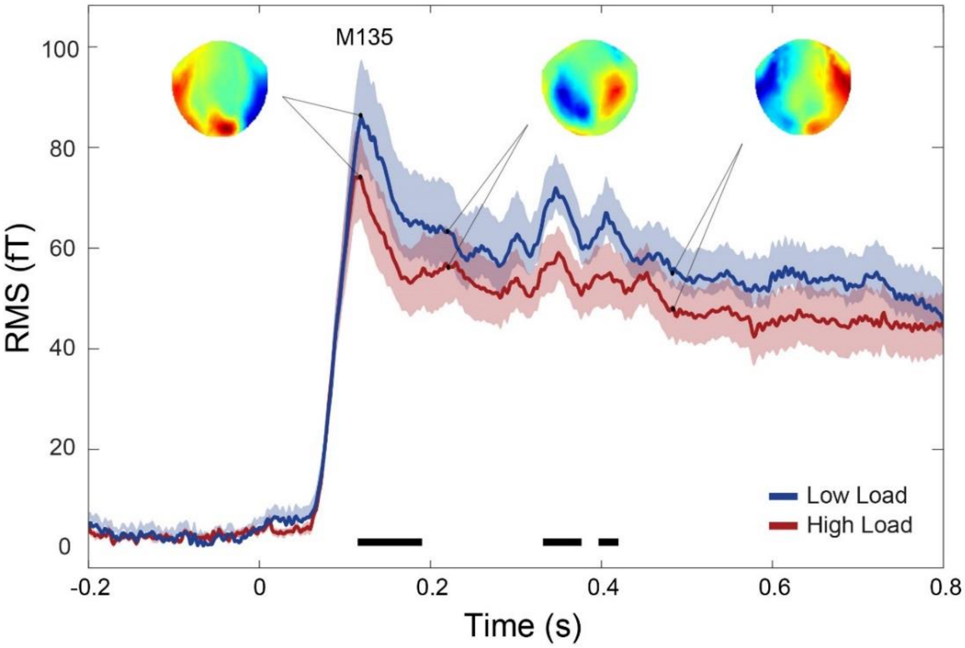


Figure 4.8: Experiment 5 – Visual Evoked Responses

Mean RMS (instantaneous power) of activity in visual-only trials contrasting responses in low load and high load, with scalp topographies to illustrate polarity reversals.

The evoked responses from the visual only trials (Figure 4.8) exhibited a clear onset peak with topography (inset) consistent with activity in occipital cortex. Subsequently, the group mean data apparently show a relatively flat sustained response; however, topographies indicate that there were in fact reversals in polarity within this response. Individual data (not shown) showed successive peaks which reversed in polarity, but due to relatively large variability in peak latencies between subjects, the group mean data do not reflect distinct peaks. When contrasted between low and high load, low load trials resulted in a larger amplitude throughout most of the response, which was significant from 115-190 ms and 330-420 ms post onset. The first period of significance occurred at a time which is associated with the P1-N1 complex in EEG, and multiple reports have shown that one or both of these peaks are stronger during feature search compared to conjunction search (Handy et al., 2001; Wolber and Wascher, 2003; Muller et al., 2006). The effects in the latter portion of the trial are consistent with late effects seen in the previous MEG data presented in this thesis, and likely reflect that under low load responses such as the N2pc and P3 (which occur in response to target selection) are stronger due to better task performance (Hopf et al., 2000; Molloy et al., 2015).

Interestingly, there was no effect of load on visual responses using the colour-shape (CS) task from 0-100 ms post onset. This is in contrast to the letter search (LS) task described previously, where high load led to stronger evoked responses in this early period (both in Experiment 4 here, and in Experiment 1, Chapter 2). This may reflect that the stimulus displays in the LS task had different complexity and luminance between low and high load in, while in the CS task I ensured the arrays were identical. However, when participants viewed the LS arrays but had no active visual task (recall the control group for Experiment 1) there was no difference in the evoked visual responses, suggesting that active attention to the stimuli is required in order for the effect to occur.

4.3.2.2 Auditory Processing

Effect of Load on Overall Auditory Responses

Figure 4.9 shows the auditory evoked responses, collapsed over FA and FP trials, as a function of visual task load. As in experiment 4, the activity isolated using the DSS analysis shows the standard auditory response peaks with reversing dipolar scalp topographies over temporal channels. This confirms that auditory activity was isolated successfully from the audio-visual response mixtures.

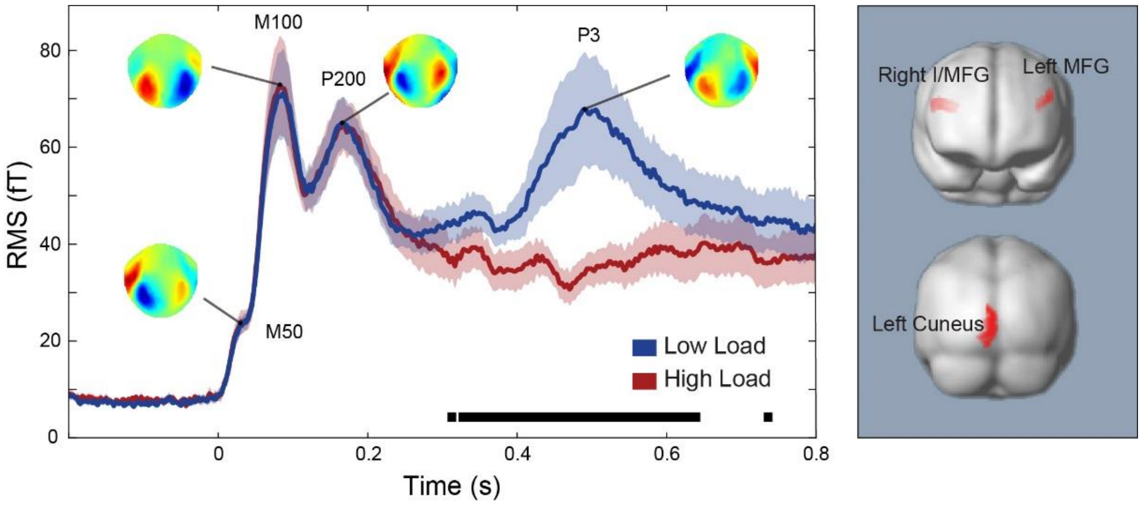


Figure 4.9: Experiment 5 - Auditory Evoked Responses

Mean RMS (instantaneous power) of responses to auditory stimuli (collapsed over FP and FA trials) in Experiment 5 under LL and HL, with scalp maps of peak topographies. Error bars reflect the SD of bootstrap iterations for each condition, and significant differences between the conditions are marked by horizontal black bars at the bottom of the plot. Right panel shows regions where activity was stronger under LL than HL. No regions were found to be significant in the opposite direction to that displayed.

Significant effects of load were observed from ~300-650 ms post stimulus, with a clear P3 apparent in the responses to the sounds under low but not high visual load.

This is in keeping with the findings from Experiments 1 and 4, where high visual load also led to the suppression of the P3 'awareness' response to the sounds. However, in contrast to both previous experiments, here there was no evidence of a suppression of the earlier onset response complex under high visual load. This suggests that the task used here induced load somewhat differently to the previous task, perhaps at a later stage in visual processing.

Source analysis revealed significantly stronger activity in frontal and occipital regions in low load compared to high (see Figure 4.9). The activity in bilateral middle frontal gyri is likely to be the source of the P3 response which was apparent under low but not high load; the P3 has reliably been shown to have a generator in the frontal lobe when it occurs in response to non-target stimuli (Comerchero & Polich, 1999; Simons et al. 2001; Goldstein et al., 2002; Polich 2007). The difference in activity within the left occipital lobe may indicate some residual visual activity.

Table 4.1: Experiment 5 – Source Contrast for Main Effect of Load

Source estimates for the difference between auditory responses in low and high load trials (regardless of figure presence).

Cluster		***Peaks***			
Cortical Structures	***p***(FWE-corr)	***t***	Co-ordinates		
			x	y	z
Left Occipital Lobe: Cuneus, Precuneus, Inferior Occipital Gyrus, BA 17, BA 18	0.001	4.46 3.92 3.70	-4 -6 -2	-92 -84 -100	-12 18 2
Right Frontal Lobe: Middle Frontal Gyrus, Inferior Frontal Gyrus	0.025	3.96 3.94	48 44	16 20	30 22
Left Frontal Lobe: Middle Frontal Gyrus	0.022	4.46	-48	20	30

Effect of load on FRN

Figure 4.10 A plots FP and FA responses separately for low and high load. Under low load there was a significant FRN from 33 ms post onset until the end of the epoch. Under high load there was a smaller FRN which was significant for relatively brief intervals from 185 ms post stimulus onset. As in Experiment 4, there was a significant effect of load on the FRN from ~50-75 ms post onset (during the onset of the FRN), and from ~200 ms post onset thereafter (during the peak of the FRN and onwards). Similarly, as in Experiment 4, load had a significant effect on FP responses (green bars in figure 4.10 B) but not on FA responses.

The fact that Experiment 5 replicated all the significant effects of load on auditory ERFs found in Experiment 4 gives some reassurance that these represent a genuine impact of load, in spite of the fact that behavioural results indicating that the task in Experiment 4 may have been fairly easy even in high load. In fact, interestingly, both the FRNs measured in Experiment 5 were stronger than those seen in Experiment 4, suggesting that overall, the visual task used in Experiment 5 was less disruptive to auditory processing in spite of it being a harder task behaviourally.

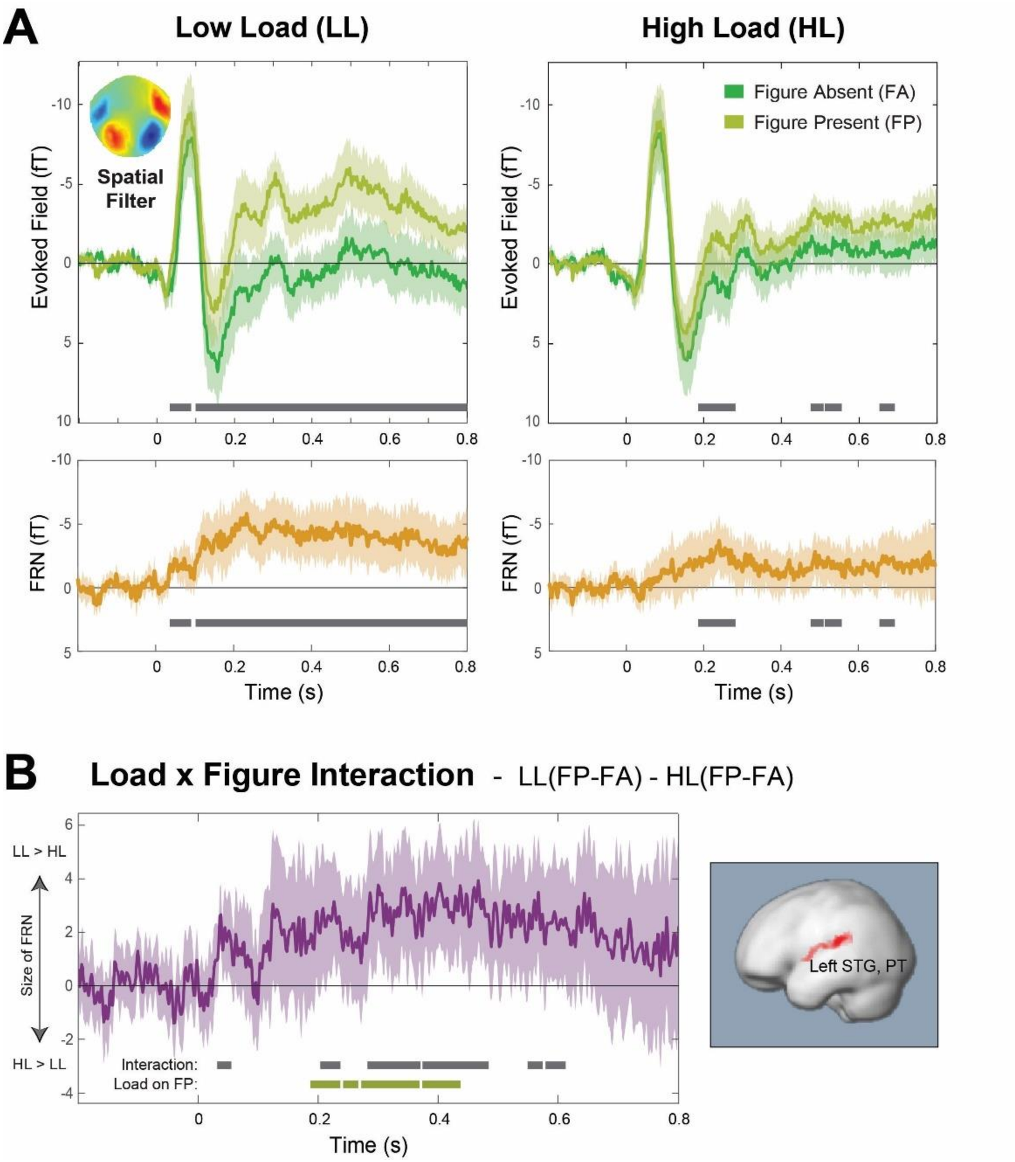


Figure 4.10: Experiment 5 - Effect of Load on FRN

A: FRN under low load (left) and high load (right). *Top panel shows the evoked responses separately for FP and FA stimuli, while the bottom panel illustrates the FRN (difference timeseries of FP-FA). The spatial filter used to calculate the responses (see methods) is inset. Error bars reflect the SD of bootstrap resamplings for each condition, and significant differences between the conditions are marked by horizontal black bars at the bottom of the plots.* ***B: Load x FRN Interaction.*** *The difference timeseries, LL(FP-FA) – HL(FP-FA), quantifies the interaction between load and figure. Error bars show 2*SD of the bootstrap, for comparison with zero line; periods when the values differed significantly from zero are indicated by the black bars below the plot. Green bars indicate periods when load had a significant effect on responses to FP stimuli (no significant periods were found for the effect of load on FA responses). Right panel indicates regions where the source analysis showed a significant interaction between responses in LL vs HL and FP vs FA.*

Individual Differences

There was a significant positive correlation between the effect of load on behavioural performance and the effect of load on the amplitude of the FRN, Spearman’s $r_s(20) = .572$, $p < .01$ (Figure 4.11). This indicates that participants who showed larger performance deficits on the visual task as load was increased also showed more substantial reductions in the neural signature of auditory figure-ground segregation in the higher compared to lower load conditions. This finding supports the critical role of demand on processing resources in determining the magnitude of the cortical response to auditory figure ground segregation.

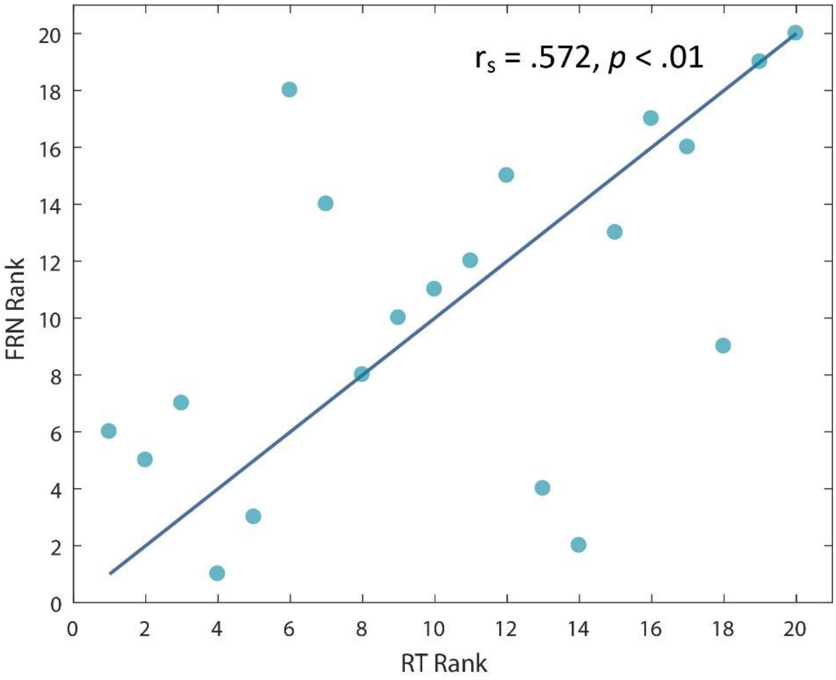


Figure 4.11: Experiment 5 – Individual Differences

The effect of load on reaction times in the visual task (x-axis), and the amplitude of the auditory FRN (y-axis) were calculated for each participant, and ranked within the group. Plotted is the correlation between the two ranks across individuals, which showed a significant relationship, $r_s(20) = .572$, $p < .01$.

Source Analysis

When collapsed over low and high load, source analysis identified areas in the right temporal and right parietal lobes that showed greater activity in response to FP vs FA scenes (Table 4.2). The temporal region covered the posterior portion of the right superior temporal gyrus, with some extension to middle temporal gyrus and planum temporale. This closely mirrors the bilateral temporal sources seen in Chapter 3 (when SFG signals were presented without a load task), and previous fMRI and MEG data (Teki et al., 2011; Teki et al., 2016). The parietal source covered regions of the superior and inferior parietal lobules. It was slightly superior and anterior to the source seen in Chapter 3, and overall more diffuse, but given the relatively poor spatial resolution for MEG, I believe both represent activity within the IPS. Both these loci are also consistent with the fMRI and MEG data discussed above (Teki et al., 2011; Teki et al., 2016). This further confirms that the DSS analysis successfully captured the relevant SFG evoked activity.

Table 4.2: Experiment 5 - Source Contrast for Main Effect of Figure

Source estimates for the difference between auditory responses in FP and FA conditions (regardless of load).

Cluster		***Peaks***				
Cortical Structures	***p***(FWE-corr)	***F***	***p***(uncorr)	Co-ordinates		
				x	y	z
Right Parietal Lobe:	< .001	21.64	<.001	48	-38	10
Inferior Parietal Lobule, Superior Parietal Lobule, BA 40		20.74	<.001	66	-26	2
		20.66	<.001	52	-24	2
Right Temporal Lobe:	< .001	19.31	<.001	38	-38	44
Superior Temporal Gyrus, Middle Temporal Gyrus, BA 22, BA 41, BA 42		17.11	<.001	42	-26	42
		16.74	<.001	38	-30	48

The source of the interaction between load and figure was localised to the left temporal lobe (Figure 4.10 B, Table 4.3), an area which extended down the left superior temporal gyrus and planum temporale, including Heschl's gyrus. This suggests that relatively early cortical processing of the SFG stimuli was affected, in keeping with the fact that the evoked data showed an early impact of load. The left-lateralised effect is commensurate with the main effect of figure described above, which was significant in the right but not left temporal gyrus.

Table 4.3: Experiment 5 – Source Contrast for Figure x Load Interaction

Source estimates for the interaction between load (low, high) and figure presence (FP, FA).

Cluster		***Peaks***				
Cortical Structures	***p***(FWE-corr)	***F***	***p***(uncorr)	Co-ordinates		
				x	y	z
Left Temporal Lobe:	0.006	25.82	<.001	-50	-30	18
Superior Temporal Gyrus,		14.69	<.001	-60	-2	6
Planum Temporale BA 41, BA 42		13.67	<.001	-62	-16	14

4.4 EXPERIMENT 6: SFG UNDER LOAD – CS TASK - BEHAVIOURAL

Having established that there was a reduced neural signature of auditory figure-ground segregation under load, I wanted to determine whether the behavioural percept of the figure within the SFG stimuli was similarly susceptible to a drain on processing resources. I ran a behavioural dual task experiment in which participants performed a primary visual load task but were also asked to determine whether a concurrently presented SFG stimulus contained a figure.

4.4.1 Methods

4.4.1.1 Participants

Thirteen paid participants took part in the behavioural study. One was excluded because their performance on the low-load task was extremely poor (61%; average of all included participants was 97.5%). For the remaining twelve participants (9 female), ages ranged from 18-24 years (mean = 20.6, SD = 2.1). All participants had normal or corrected to normal vision and reported normal hearing. None had taken part in any of the previous experiments presented in this thesis, or in experiments which used a similar load task.

4.4.1.2 Apparatus

The experiment was run on a Dell PC with a 13” monitor using Matlab 7.12 and Cogent 2000 (http://www.vislab.ucl.ac.uk/cogent.php). A viewing distance of 57cm was maintained throughout using a chin rest. Sounds were presented via tubephones (E-A-RTONE 3A 10 Ω, Etymotic Research, Inc) inserted into the ear-canal.

4.4.1.3 Stimuli and Procedure

Auditory and visual stimuli were identical to those used in Experiment 5, except that auditory stimuli were presented on every trial (50% FP, 50% FA. See Figure 4.6). As the stimuli were presented, participants made a speeded response (with their right hand) to the primary visual task. Then, unlike in the MEG experiment, participants were also asked to respond (with their left hand) to indicate whether there had been a figure within the auditory signal (Figure 4.6).

To acquaint participants with the two tasks and the dual task procedure, a series of short demo blocks preceded the main experimental session. These introduced the auditory and visual tasks separately, giving trial-by-trial feedback to ensure participants understood each task, and then combined them to illustrate the procedure for the dual task.

The experiment consisted of 12 blocks of 40 trials each, six low load and six high, with the order of blocks counterbalanced between participants. Trial-by-trial feedback was not given during experimental blocks, but at the end of each block participants were provided a score of percentage correct on the visual task, to boost engagement and maintain priority on the visual task. Blocks lasted for ~3 minutes each, and participants were encouraged to take breaks between blocks when needed.

Data was analysed exactly as described for Experiment 2 (see Section 2.3.1.3).

4.4.2 Results

Results are displayed in Table 4.4. Participants showed a significant effect of load on performance in the primary visual task. Increased load led to lower accuracy ($z = -11.14$, $p < .001$) and longer RTs ($t(1,11) = -12.5$, $p < .001$), indicating that the load manipulation was successful. Higher load in the primary task also led to reduced sensitivity to the target in the secondary auditory task – d' scores were significantly reduced in the high compared to the low load condition ($t(1,11) = 3.7$, $p = .004$). The change in sensitivity was not accompanied by a change in decision criterion (beta) or false alarm rates, indicating that the change of load in the visual task did not affect participants' response behaviour.

These data confirm that participants can successfully identify the figure within very short auditory scenes when they are actively attended. However, high demands in a concurrent (and prioritised) visual task reduced participants' ability to perceptually segregate the auditory figures from the scene, even though they were trying to detect them.

Table 4.4: Experiment 6 - Dual Task Behavioural Data.

	Auditory Task								***Visual Task***			
	d'		**Beta**		**FA Rate**		**Hit Rate**		**% Correct**		**RT (ms)**	
	LL	HL	LL	HL	LL	HL	LL	HL	LL	HL	LL	HL
Mean	1.8	1.5	0.8	0.91	31.1	32.8	85.5	80.5	97.5	89.9	714	1066
SD	0.27	0.23	0.16	0.15	5.7	5.9	3.3	3.6	0.5	1.4	42	31
***p* value**	0.004		0.178		0.44		0.001		<0.001		<0.001	

4.5 DISCUSSION

Increased visual load affected the processing of the SFG stimuli in two ways: first, there was an impact on the overall response to sound, with high load leading to a reduced P3 response in both MEG experiments, and a reduced P2 response in the Letter Search (LS) group. Second, there was a specific effect of load on responses to figure present stimuli such that the FRN was reduced (Colour-Shape group, CS) or even abolished (LS group) under high load. The effect of load on the FRN began as little as 50 ms after stimulus onset and was localised to auditory cortex, suggesting that early auditory processing was affected. The results of the behavioural experiment indicate that these effects of visual load result in poorer perceptual detection of the figure within the sounds, even when participants are explicitly listening for them. These results suggest that load on the visual system can deplete the computational capacity of the auditory system, and highlight that auditory figure-ground segregation is not strictly pre-attentive.

4.5.1 Figure-ground segregation is dependent on attention

Previously, data demonstrating brain or behavioural evidence of segregation in the absence of directed attention (Alain and Izenberg, 2003; Macken et al., 2003; Dyson et al., 2005) have been taken to indicate that segregation takes place automatically and pre-attentively, and that attention operates at a subsequent stage to enhance the representation of attended objects and suppress distracters (Bidet-Caulet et al., 2007; Elhilali et al., 2009a; Ding and Simon, 2012; Zion Golumbic et al., 2013; Akram et al., 2014). Specifically in the context of SFG segregation, accumulating work in humans has shown figure-evoked brain responses in naïve, distracted listeners, suggesting an early pre-attentive computation of temporal coherence (Teki et al., 2011; O'Sullivan et al., 2015; Teki et al., 2016) that is enhanced during active listening (O'Sullivan et al., 2015). In keeping with these reports, here under low load conditions (and in the 'passive' conditions used in Chapter 3), auditory evoked fields reliably showed an FRN in response to FP stimuli. The elicitation of the FRN in response to unattended stimuli supports the conclusion that figure-ground segregation process, even within the brief signals used here, does not require directed attention.

However, the load manipulations presented in this chapter illustrate that depleting resources can lead to an abrupt failure of the segregation process. In both MEG experiments (using two different visual load tasks), conditions of high visual load resulted in a reduction of the FRN so that there was little or no distinction between the neural responses to figure present and figure absent scenes. The critical effect of load on the FRN is further corroborated by the finding (from Experiment 5) of a significant correlation between the impact of perceptual load on individuals' visual task RT and its impact on the amplitude of their FRN: larger load effects on RT (indicative of a greater impact on capacity) were associated with a larger load effect on the FRN. This is consistent with a recent report establishing individual differences in perceptual capacity (Eayrs and Lavie, 2018). These results suggest that high load impairs the mechanisms which support auditory segregation, indicating that these computations are not performed in a truly automatic manner.

Though the FRN reported here shares many characteristics with the ORN, it was affected by task load whereas the ORN has, thus far, appeared robust to load manipulation (Alain & Izenberg, 2003; Dyson et al., 2005). This could indicate that the ORN reflects a distinct form of auditory segregation – one based on instantaneous harmonicity cues – which may be less susceptible to depletion of computational resources than segregation based on temporal coherence (Micheyl and Oxenham, 2010). However, it is also possible that the tasks used in the ORN studies did not involve a sufficiently high level of load to exhaust capacity; as noted in the introduction to this chapter, the tasks used in these studies were not classic perceptual load manipulations, and nor was the load timed precisely to the auditory stimuli. Thus, it remains unclear whether concurrent sound segregation based on harmonicity cues reflects a truly automatic process that is distinct from temporal coherence calculations, or whether it might require relatively little resources, and so was not disrupted by the high load task used in the ORN studies.

As discussed in Chapter 3 (Section 3.4.2), it is plausible that the FRN and the ARN reported by Gutschalk et al. (2008) are the same component. The finding that the FRN can be affected by the availability of processing resources provides another similarity with the ARN reported by Gutschalk et al. (2008). Both responses have been shown to be modulated by attention

4.5.2 Source of load effect on temporal coherence calculations

As discussed in Chapter 3, it is likely that the computations underlying auditory figure-ground segregation based on temporal coherence are elaborated along a processing hierarchy that involves both classic auditory cortical regions and broader network locations such as the IPS. Teki et al. (2016) suggest that coherence maps are computed within PT and then passed to the IPS, which integrates the information and reads out a signal if the sound is determined to consist of multiple objects. A priori, the IPS presents as likely target for an interaction between visual and auditory processing, since it is a region associated with relatively high level processes such as attention and perceptual organisation across both audition and vision. In particular, it has been found to be more active during conjunction than feature search (Shafritz et al., 2002), even when the tasks are matched for difficulty (Donner et al., 2002), suggesting a role for IPS in binding visual features together into separate objects. It has also been implicated in auditory segregation, using both the ABA- streaming stimuli (Cusack, 2005), and the SFG signals used here (Teki et al., 2011; Teki et al., 2016); in both cases, IPS activity correlated with the percept of multiple auditory objects. In fact, Cusack (2005) suggested that visual feature binding may be underwritten by similar neural processes to auditory segregation, since both are reliant on grouping low-level features into objects or groups based on cues such as temporal onset and spatial or frequency covariance (Humphreys, 2001; Müller et al., 2001). Thus, it seems plausible that the locus for the competition for resources between vision and audition in these experiments might be the IPS, which appears to perform a similar function in both tasks.

In contrast to this supposition, the data presented here indicate that load impacted processing in temporal areas: the effect of load on the FRN was localised to the upper bank of the left superior temporal gyrus, including the PT and Heschl’s gyrus. The fact that the effect of load was apparent within ~50 ms of stimulus onset is also consistent with an early effect on cortical processing. These findings illustrate that the competition for resources may in fact occur much lower in the processing hierarchy, being evident during even early sensory processing. It also suggests that the interaction seen in the present experiment was not an incidental effect that occurred as a consequence of the fact that the specific auditory and visual tasks used happen to share a common cortical locus in a ‘high level’ brain region, within which an interaction might occur subsequent to sensory processing.

Temporal coherence calculations that occur within auditory cortex have been proposed to be supported by rapid neuronal plasticity (Elhilali et al., 2009b; Krishnan et al., 2014; Lu et al., 2017); recent animal work has shown that multiple adaptive processes occur in auditory cortex in response to temporally coherent signals, including narrowing of spectro-temporal receptive fields and changes in both spiking rates and spiking correlations to the signal (Lu et al., 2017). However, in animals this rapid plasticity only took place when the auditory signals were explicitly attended (Lu et al., 2017); this does not tally with the human literature, where segregation has been reliably observed during passive listening (Teki et al., 2011; O'Sullivan et al., 2015; Teki et al., 2016). The data presented here clarify that the limiting aspect is not active attention per se, but rather the availability of general computational resources irrespective of the explicit focus of attention. Thus, a disparity in the size of resource pools could explain the apparent differences between the levels of attention required for these adaptive processes in humans versus animals.

It is worth noting that the source localisation carried out here does not necessarily rule out further impacts of load later in processing such as within the IPS. The source data for the interaction presented here were based on only a single component of the evoked data (isolated to highlight the FRN, see methods; though note that source inversions based on the full auditory evoked response did not show any regions where the interaction was significant). Given that even for the passive data (Chapter 3) a small volume correction was necessary for the parietal source to reach significance, it is possible that the signal in IPS in the load experiment (which would only have been present in the two FP conditions of the four total conditions, and, as illustrated by the amplitude of the FRNs, was weaker under load than during passive conditions) may not have been strong enough to draw the source algorithms used here. Future work using fMRI would be useful to fully explore the spatial extent of the interaction.

4.5.3 Visual load can impair auditory computational capacity

In Chapter 2 I illustrated that high visual load can lead to the suppression of early evoked responses in auditory cortex for near-threshold sounds, with the result that they were more likely to go undetected. In this Chapter I aimed to assess the potential impact of load specifically on auditory computation, and so needed to ensure that the auditory signals were robust enough to survive any gain-reduction effects. To this end, I chose SFG signals that produced a clear figure when processed passively (Chapter 3) and presented the stimuli at a clearly audible volume (as confirmed by the fact that participants in the dual-task group did not report any difficulty hearing the auditory stimuli under high load). Though the sounds remained audible, there was nonetheless an effect of load on the overall responses to sound: in both MEG experiments the P3 response was abolished by high load, and in the LS experiment the latter portion of the P2 response was reduced under load, suggesting that under high load the sounds were less likely to become the focal point of attention, or even reach subjective awareness. However, there was no effect of load on the early onset responses (compared to the effect on the aM100 seen in Chapter 2), suggesting that under these conditions the gain effects were confined to later stages of auditory processing.

In addition to the small dampening effect of load on auditory responses, isolating the component of the response which best captured the FRN illustrated a second effect of load that was specific to the processing of the figure. In both MEG experiments there was a substantial effect of load on the FRN that was characterised by an effect solely on responses to figure present stimuli (figure absent responses did not differ across load conditions). This effect was evident from as early as ~50 ms after onset (and by 70 ms post-onset in both experiments), which corresponds to the earliest portion of the FRN as measured during passive listening (Chapter 3). In fact, since the earliest latency at which the figure becomes technically extractable is at 25 ms (the onset of the second chord), and that the earliest MEG-recorded cortical responses to sound occur at a latency of ~30-50 ms, this timeframe captures the very earliest point at which an effect of load could be expected to be seen. This is in keeping with the localisation of the effect, which also suggests that early cortical processing was impacted. Since overall responses to sound were not affected by load at this stage (and, as noted above, the effect of load here was confined to FP responses), load appears have specifically affected the early cortical computations underlying the extraction of the figure from the background. The effect of load on the FRN persisted until at least 600 ms post onset in both experiments, suggesting that load also impaired subsequent processes, potentially including the P3 awareness response to the figure, which was apparent at ~400 ms post onset under passive listening conditions (Chapter 3). This suggests an extended temporal impact of load, though this may stem from the early impact on computations during crucial stages of auditory processing, which may be sufficient to disrupt the entire segregation process.

These data provide strong evidence that auditory computational processes, even ones which underlie basic functions of the auditory system, rely on the availability of attentional resources. Moreover, these resources appear to be shared between the visual and auditory systems. This highlights that processing in early auditory cortical regions is not encapsulated, and that a draw on domain-general resources can lead to both attenuation of evoked responses, and the failure of fundamental computational processes.

4.5.4 Contrast between visual tasks

Though both visual tasks resulted in qualitatively similar effects on auditory processing, the dynamics of the visual evoked activity differed between tasks. This is somewhat surprising, especially given that in essence both tasks manipulated load via a feature versus conjunction search: for the low load LS task only low level features needed to be processed to identify the target (which was formed of straight lines, while non-targets were formed of curves), while in high load all letters were formed of straight lines, so binding of the individual components into full letters was required to distinguish target from non-targets. One possible reason for the different effects is that colour was a relevant feature for the CS task, while all stimuli in the LS task were monochrome. Colour-driven activity is typically local to visual cortex, rather than some of the activity related to object identification, which occurs along a ventral pathway that extends bilaterally down the length of the inferior temporal gyrus (Grill-Spector, 2003). Additionally, early frequency-locked components of the ERP are enhanced in colour feature search relative to a conjunction search (Muller et al., 2006). Since the MEG analysis used here may have been more sensitive to activity emanating from occipital sources it may have been particularly sensitive to this colour-induced activity that was particularly prominent during low load conditions of the CS task.

The differences between tasks also appeared to have an effect on how auditory processing was affected under load. For the overall auditory evoked responses, the effect of load in the LS task was evident earlier than in the CS task, with the LS task resulting in an impact on both the P2 and P3, while the CS task impacted only the P3. Similarly, the LS task appeared generally more disruptive to figure processing than the CS task, with FRNs in both low and high load being substantially smaller in the LS task than the CS. Since the behavioural performance data suggest that the LS task was in fact *easier* overall than the CS task, the explanation for its efficacy may lie elsewhere: the mask presented at 100 ms in the LS task would have been a salient visual event which may have drawn attention and processing resources back to the visual system at a crucial time for auditory processing. This again emphasises the speed and flexibility with which resources are allocated, and highlights that the timing of load is an important factor to control.

CHAPTER 5: PUPILLARY RESPONSES

5.0 MOTIVATION

During collection of three of the MEG datasets presented previously in this thesis – those where irrelevant auditory stimuli were presented alongside an active visual task – pupil data were also recorded. Pupil diameter is known to vary systematically with the demands of a task, and responds to both visual and auditory stimuli, so it could potentially provide a useful physiological indicator of the degree of processing that occurred to both the task-relevant visual stimuli and the irrelevant auditory signals. Since MEG signals were recorded concurrently, this provided a rich dataset within which it was possible to assess the degree to which pupil dynamics can capture the separate effects of load on visual and auditory processing.

5.1 INTRODUCTION

The size of the human pupil is constantly in flux. The strongest drivers of pupil diameter are low-level factors such as the pupillary light reflex, which controls the amount of light that enters the eye (Ellis, 1981). However, a wealth of research has shown that pupil diameter also varies systematically with the demand that cognitive and perceptual tasks place on the system. Studies across a range of fields have shown that more demanding tasks lead to an increase in pupil dilation, from basic perceptual discriminations (Kahneman et al., 1967) to more complex processes such as working memory (Kahneman and Beatty, 1966; Peavler, 1974; Granholm et al., 1996), mental arithmetic (Hess and Polt, 1964; Bradshaw, 1968; Payne et al., 1968; Marshall, 2002; Klingner et al., 2011), language processing (Wright and Kahneman, 1971; Just and

Carpenter, 1993; Hyönä et al., 1995; Engelhardt et al., 2010; Seeber, 2013), recall (Sterpenich et al., 2006; Võ et al., 2008; Heaver and Hutton, 2011; Kafkas and Montaldi, 2011; Otero et al., 2011), reasoning (Bradshaw, 1968; Boersma et al., 1970), and cognitive control (Brown et al., 1999; Siegle et al., 2004; Siegle et al., 2008; Laeng et al., 2011; van Bochove et al., 2013; Van Steenbergen and Band, 2013). Interestingly, the effect occurs even when the demanding task is purely auditory (Kahneman et al., 1967), emphasising that it is not a stimulus-driven visual reflex but rather a response to the demands placed on the system as a whole.

5.1.1 Pupil dynamics and the LC-NE system

The link between pupil diameter and cognitive processing is believed to derive from activity in the Locus Coeruleus (LC), a nucleus within the brainstem that is the primary source of norepinephrine (NE) within the brain. Human neuroimaging has illustrated that pupil diameter correlates to BOLD activity in the LC (Alnæs et al., 2014; Murphy et al., 2014; Yellin et al., 2015), and pharmacological studies which manipulated tonic NE levels using α2-noradrenoceptor agonists (Clifford et al., 1982; Fanciullacci et al., 1988; Bitsios et al., 1998; Phillips et al., 2000) or antagonists (Morley et al., 1991; Phillips et al., 2000) have shown associated changes in pupil diameter. This relationship is illustrated more directly in several key animal studies. These have shown that the tonic level of activity in the LC is reflected by the baseline pupil diameter, with higher tonic rates being associated with larger pupil diameters (Rajkowski, 1993; Joshi et al., 2016; Reimer et al., 2016). Phasic patterns of firing in the NE-expressing neurons within the LC are also closely mirrored by short-term fluctuations in pupil diameter, effecting rapid task-related pupillary responses

(Rajkowski, 1993; Joshi et al., 2016; Reimer et al., 2016). These phasic changes in LC-NE activity have also been proposed to act as interrupt signals which are triggered by unexpected events within a task (Dayan & Yu, 2006). Though the mechanism underlying this correlation between LC-NE activity and pupil diameter is not entirely understood (Gilzenrat et al., 2010; Jepma and Nieuwenhuis, 2011; Costa and Rudebeck, 2016), there are plausible anatomical pathways such as an indirect inhibitory projection via the effect of NE on the Edinger-Westphal nucleus, which engages the parasympathetic nerve fibres that control the pupillary sphincter muscles (Loewenfeld and Lowenstein, 1993; Beatty and Lucero-Wagoner, 2000). While there are other factors that affect tonic changes in pupil diameter, such as the level of cortical acetylcholine (ACh) activity, which emanates from the basal forebrain, and is driven by both arousal and motor activity (Reimer et al., 2016), the short-term phasic changes in pupil diameter that are associated with stimulus processing, such as those measured in this series of experiments, are understood to be driven mainly by phasic changes in NE activity (Rajkowski, 1993; Joshi et al., 2016; Reimer et al., 2016).

The function that this correlation between activity in the LC and pupil dilation plays is not entirely clear. The LC –NE system responds to stress (Valentino et al., 1993; Aston-Jones et al., 1994; Valentino and Van Bockstaele, 2008), so it is possible that pupil dilation may be epiphenomenal, driven by simple changes in arousal which affect both neurotransmitter levels and muscle tone all over the body. However, there are dense noradrenergic projections from the LC to cortical areas associated with attention (Foote and Morrison, 1987), and the LC-NE system has been proposed to orchestrate the functional integration of all brain attention networks (Foote and

Morrison, 1987; Coull et al., 1999; Corbetta et al., 2008; Schneider and Kastner, 2009; Gilzenrat et al., 2010; Eldar et al., 2013). The LC is also active during memory retrieval (Sterpenich et al., 2006) and direct stimulation has been shown to increase plasticity during learning and to significantly heighten cortical responses to sensory input (Martins and Froemke, 2015). The role of LC-NE activity in these key cognitive functions suggests that the link between pupil dilation and demand on cognitive processes may not simply reflect arousal, but could have a more specific function. For example, since phasic NE is believed to act as a signal to interrupt or reset on-going processing in response to an unexpected event, the NE-associated pupil dilation may have evolved to maximise the visual sensory information available during important decision-making processes.

5.1.2 Pupillary responses under cognitive load

Though there has been extensive research into how pupillary dynamics vary with processing demand, the vast majority of the tasks employed fall into the category of cognitive load. The earliest study that showed a systematic link between pupil and task demand was conducted by Hess and Polt (1964): participants completed mental multiplications with four levels of difficulty, and pupil diameter during each trial was found to increase with multiplication difficulty. Though with only five participants the sample size in this original study was relatively small, many subsequent studies using similar mental arithmetic tasks have replicated the findings (Bradshaw, 1968; Payne et al., 1968; Marshall, 2002; Klingner et al., 2011). Other early studies used working memory tasks which are classic manipulations of cognitive load: Kahneman and Beatty (1966) found that during a serial recall task where the list items were presented one at a time, pupil diameter systematically increased as each new item appeared. Similar effects were shown by Peavler (1974) and Granholm and colleagues (1996), where pupil diameter was shown to increase up to the point when the number of items exceeded individual's capacity. Since these early studies, most research has explored more complex cognitive tasks such as conflict resolution (e.g. Laeng et al., 2011; van Bochove et al., 2013; Van Steenbergen and Band, 2013) or recall of long-term memories (e.g. Sterpenich et al., 2006).

5.1.3 Pupillary responses under perceptual load

In contrast to the extensive literature on how pupil dynamics reflect load on cognitive processing, relatively little work has explored how the pupil responds when the task load is exclusively perceptual. In 1967, Kahneman and Beatty followed up their working memory study with one which used an auditory pitch discrimination task. Participants were instructed to determine whether a comparison tone was higher or lower than a standard tone. During trials where the standard and comparison tone were closer in frequency (i.e., a more difficult discrimination), larger pupil dilations were observed. Perhaps because the results mirrored those seen using tasks which load cognitive resources, and because the distinction between cognitive and perceptual load was not formalised until much later (for example Lavie's 2000 paper on perceptual load and cognitive control), these effects were not explored further for some time. In 2007, Porter et al. conducted a study into how pupil diameter reflected demands during visual search. Their search task varied both the number of distractors and the degree of heterogeneity amongst distractors, giving a classic manipulation of perceptual load. When distractors are homogeneous, the search is considered ‘efficient’ (low load) and can be performed similarly quickly regardless of set size. In contrast, when the distractors are heterogeneous, search is ‘inefficient’ with each item being inspected in serial, and therefore the load (and response times) increase with set size. Their pupillometry data showed that pupil diameter increased with set size during inefficient but not efficient search, illustrating a pattern where dilation increased with perceptual load. Finally, a third study by Alnæs et al. in 2014 explored pupillary dynamics during a multiple object tracking task. Participants viewed 10 objects on screen which moved along random, independent vectors at constant speeds. Their task was to covertly track (i.e., while maintaining their gaze upon a fixation at the centre of the screen) a subset of the objects over a period of ten seconds. Load was manipulated by varying the set size of the target objects between two and five, and pupil diameter was found to increase systematically with the set size. These findings provide preliminary support for the notion that pupil diameter reflects the perceptual and cognitive demands of a task in a similar manner.

5.1.4 Pupil diameter varies systematically with load

One factor which makes pupil diameter particularly attractive as a physiological marker for task demand is that the precise magnitude of pupil dilation may offer a quantifiable measure of the difficulty of the task. Many studies have indicated that when tasks are parametrically varied, pupil diameter increases monotonically as a function of the task demand (Hess and Polt, 1964; Kahneman and Beatty, 1966; Kahneman et al., 1967; Bradshaw, 1968; Payne et al., 1968; Peavler, 1974; Granholm et al., 1996; Marshall, 2002; Klingner et al., 2011; Alnæs et al., 2014). Furthermore, as demands increase to the point when individual capacity is reached, this threshold is also reflected in the pupil diameter: in studies where working memory was loaded by sequentially presented items, pupil diameter increased with each object but plateaued at the point when individual capacity was reached (Peavler, 1974; Granholm et al., 1996). This means pupil diameter could be used as an objective index to quantify the relative demands of different tasks, or to compare between individuals how demanding a particular task is (for example as a proportion of their maximum capacity). While there are some limitations to this, such as the fact that some individual differences including age and ocular pathology can have effects on pupillary dynamics that are not associated with task demand, or that there are physical limits on maximum pupil size and speed of dilation, there are also many potential benefits. For example, since pupil diameter can be measured in real time, it can provide immediate feedback on how loaded an individual is, potentially alerting when capacity is close to being reached. As technology improves and makes portable pupillometry more feasible and affordable, it could be a powerful tool for real-world situations such as driving, when being overloaded can lead to dangerous mistakes.

5.1.5 Pupil diameter as a marker of perception

In addition to reflecting the level of processing load involved in a task, pupil diameter has been shown to reflect conscious perception of stimuli. For example, when presenting near-threshold light flashes, Hakerem and Sutton (1966) showed that there was pupil dilation in response to flashes which were detected, but not to undetected flashes of equal luminance. Similar findings have been shown using other near-threshold visual stimuli (Privitera et al., 2010; de Gee et al., 2014), for detected versus missed visual targets in attentional blink paradigms (Wierda et al., 2012;

Zylberberg et al., 2012), and for detected auditory targets (Beatty, 1982; Beatty and Lucero-Wagoner, 2000; O'Neill and Zimmerman, 2000). However, as far as I am aware pupil dilation has only been linked to the perception of task-relevant stimuli; it is not yet clear whether a similar dilation would occur in response to perceived but irrelevant stimuli.

5.1.6 Chapter Overview

In this chapter I present pupil data that were recorded during three of the MEG experiments described previously in this thesis (Experiments 1, 4 and 5). The data show the effect of perceptual load on pupil dilation using two different visual tasks, and explore how pupil dynamics reflect the degree to which irrelevant auditory signals are detected and processed. However, I note in advance that since the pupil data were recorded during sessions where the primary focus was MEG data collection, some experimental conditions were not optimised for pupillometry. In particular, pupil data were recorded directly into the MEG dataset and were therefore represented in arbitrary units, limiting some of the options for baselining during analysis. Additionally, the lighting conditions in the MEG chamber were dimmer than those typically preferred for pupillometry. Finally, since the recordings took place in an electrically shielded room, poor pupil signal could not be rectified until the end of a recording block, which led to a relatively large proportion of excluded data and participants. Because of these limitations, I present these findings with the explicit caveat that they should be treated as preliminary, and that the findings should be replicated before any firm conclusions are drawn. Nonetheless, the data suggest that pupillometry has the potential to simultaneously measure both the load of an active task and the processing of irrelevant sensory information.

5.2 METHODS

The data presented in this chapter were collected during three of the MEG experiments described in previous chapters: the Inattentional Deafness experiment (Chapter 2, Experiment 1) and the two SFG Load experiments (Chapter 4, Experiments 4 and 5). In this chapter I refer to these experiments in a shorthand based on the auditory stimuli used (tones, **T**, or **SFG** stimuli) and the visual task used (letter search task, **LS** or colour-shape search task, **CS**), giving **T-LS** (Experiment 1, Chapter 2), **SFG-LS** (Experiment 4, Chapter 4) and **SFG-CS** (Experiment 5, Chapter 4).

5.2.1 Participants

Across the three experiments, eight of the fifty participants were excluded from the pupil analysis due to missing pupil data (> 25% of recording time missing due to artefacts or bad signal) or if there were fewer than 60 trials remaining per condition after artefact rejection; see section 5.2.4 for further details. For the Inattentional Deafness experiment (**T-LS**), there were 2 exclusions, leaving 12 participants (7 female) with mean age 28.2 years (SD = 5.0). For the SFG Letter Search (**SFG-LS**) experiment there were three exclusions, leaving 13 participants (5 female), with mean age 25.4 years (SD = 5.8). For the SFG Colour-Shape Search (**SFG-CS**) experiment there were 3 exclusions, leaving 17 participants (9 female) with mean age 25.0 years (SD = 4.3).

5.2.2 Stimuli and Tasks

The stimuli and tasks were as described in Chapters 2 and 4. In short, all paradigms consisted of an active visual search task of high or low load, accompanied by incidental auditory stimuli that were matched in duration to the visual stimulus and presented concurrently. The auditory stimuli were either 100 ms pure tones set 12 dB above individual thresholds or 200 ms SFG stimuli (signals were either FP or FA) presented at ~70dB SPL. Two types of visual search task were used: a letter search and a colour-shape search.

The letter search task was to determine the identity of a target letter (X or Z) presented within a circular array of six letters. In the low load condition, all non-target letters were O (distinguished from the targets by the low-level feature of line curvature) and were in a smaller font than the target letters, resulting in a pop-out effect. In high load, the non-target letters were all different from one another (K, W, V, N, M), were formed by straight lines only, and were the same size as the target letter, resulting in a higher level of perceptual load during the search. Note that for the **SFG-LS** experiment, a 100 ms mask was presented directly after the 100 ms search array in order to match the duration of the auditory and visual stimuli.

The colour-shape task was to determine the presence or absence of a target among a circular array of five coloured shapes. In the low load condition, the target was any blue letter (identifiable by colour, a low-level visual feature, which resulted in a pop-out effect). In high load the targets were a red circle or a green square, so that the shape and colour of each object in the array had to be perceptually bound in order to determine whether one of the targets was present.

All stimuli were presented on a grey background, but the luminance of displays was not matched between experiments.

5.2.3 Pupil Data Acquisition

Pupil diameter was recorded concurrently with the MEG data acquisition using an Eyelink 1000 (http://www.sr-research.com/EL_1000.html) remote infrared eye tracking system. The eye tracking camera was mounted on a table below the projection screen within the MEG chamber, at a fixed distance of ~64 cm from participants (who had a limited range of movement within the MEG helmet). The eye-tracker was positioned to monitor the right eye for each participant, and was calibrated before each recording session using a standard 9-point calibration on the projection screen. The lighting in the room was dimmed during the experiment. The eye tracking system recorded pupil diameter, pupil position on the horizontal axis and pupil position on the vertical axes. These data were appended to the MEG dataset as additional channels during recording, and so were measured in arbitrary units (au) and sampled at 600Hz. Due to this conversion no explicit baseline calibration was possible between blocks or participants, and so the data have been baseline corrected on a trial-by-trial basis with respect to a pre-stimulus interval (see Section 5.2.4.3).

5.2.4 Data Analysis

5.2.4.1 Blink and artefact identification

Blinks were identified in the dataset using velocity threshold criteria as suggested by Mathôt (2018), which allows for the temporal asymmetry of blinks. First the data were smoothed using a 150 ms Hanning window. A velocity profile was then calculated for each block by taking the first derivative of the pupil diameter. Using a positive and negative velocity threshold (0.2*max velocity, arbitrary units), blinks were identified as periods when the velocity exceeded the negative threshold (eye closing), exceeded the positive velocity threshold (eye opening), and then returned to 0, in that order. Ten timepoints (~17 ms) were added either side of the on- and offsets of the blink, to ensure that the whole blink had been included in the window. Periods when there were fewer than 60 samples (100 ms) between blinks were also labelled as bad as they would not contain enough data for reconstruction. Figure 5.1A shows a 60 s portion of the raw pupil data from one participant which contained several blinks, visible as sharp downward peaks. Periods which the algorithm identified as blinks are shaded in grey.

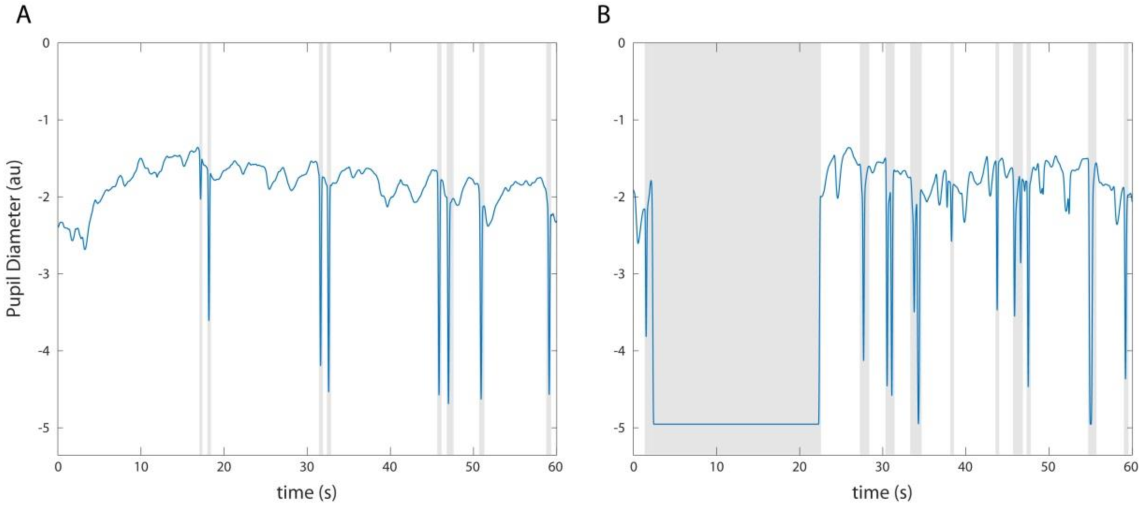


Figure 5.1: Raw Pupil Data

Two 60 second portions of single-subject data showing pupil diameter in arbitrary units. The times indicated by grey shading are those which the algorithm identified for removal. Blinks are visible as sharp downward spikes, while missing data appears as a sustained reading of -5.

Missing data were recorded by the eye tracking system as -5 in arbitrary units (this was the minimum value in all datasets, and only occurred during signal loss). This meant that times when the eye tracking signal was lost showed the same characteristic shape identified by the blink detection algorithm: a sharp decrease in the pupil diameter (negative velocity), a period of floor values, followed by a sharp increase in values (positive velocity) as the signal returned. Thus, the algorithm designed to identify blinks, as described above, also successfully identified periods when the eyetracking signal was lost for other reasons, and no further artefact identification was required. Figure 5.1B shows a 60 s portion of the raw data from a participant, which shows an extended period of signal loss.

5.2.4.2 Reconstruction or removal of missing data

For periods when the signal was lost for less than 3 seconds (characteristic of a blink or blink train) the pupil signal was reconstructed using cubic spline interpolation. For each blink, the on and offset times (calculated as described above) were labelled t2 and t3 respectively, and times t1 and t4, defined as t1=t2-t3+t2 (onset time – blink duration) and t4=t3-t2+t3 (offset time + blink duration) were calculated. Using the values of the pupil diameter at these four time points, a cubic spline fit was generated, and the values during the blink period were substituted for the interpolated data between t2 and t3.

For periods where the signal was lost for more than three seconds (and therefore spanned multiple trials), pupil reconstruction was not considered appropriate, and the data were discarded. Due to the fact that the pupil data were collected during an MEG recording which occurred in an electrically shielded room, signal loss on the eye tracker could not be addressed until the end of a recording block, and so there were several periods of poor quality signal within the datasets. Any participant whose data contained >25% signal loss overall, or where fewer than 60 trials per condition remained was not included in the group analysis (see Participants section above).

5.2.4.3 Epoching and baselining

The ITIs used in the experiments were very short to increase the yield of MEG data collection. Trials were 2000 ms long (including stimulus presentation and response window) followed by 1000 ms (plus a small jitter) during which a fixation cross was presented to signal the beginning of the next trial. However, since pupil diameter changes at a much slower rate than MEG evoked fields, some of the pupillary responses persisted well into the inter-trial interval. To account for this, epochs for the pupil analysis were defined with a relatively short pre-stimulus interval (200 ms) to reduce the impact of the previous trial on baselining, and extended to 3500 ms post stimulus onset, to capture the full response. Note that the epoch therefore included some of the next trial, but only effects within the trial are reported in the results and the onset of the next trial is indicated in all figures to ensure that the distinction is clear.

Data for each trial epoch (-200 to 3500 ms) were baseline-corrected using the average pupil diameter over the pre-stimulus period of -200 to 0 ms. This addressed any potential discrepancies in baseline values of the arbitrary units between blocks or participants (note that these units were produced on conversion from the eyetracking device to the MEG channels and could not be calibrated). However, the baseline correction meant that tonic differences in overall pupil diameter, for example between high and low load blocks, could not be assessed.

5.2.4.4 Statistical analysis

Data were averaged over trials for each participant and condition. To assess statistical differences between conditions, a Monte Carlo bootstrap procedure (1000 iterations with replacement) was applied to every timepoint of the difference waveforms, in the same manner as described for the MEG evoked responses. However, because the pupil changes more slowly than MEG fields, the autocorrelation of pupil diameter over time is stronger than that seen in MEG data. To account for this the criterion for significance was changed so that 150 adjacent samples (250 ms) with > 95% bootstrap iterations falling above or below the zero line were required for a period to be considered significant. All periods which met this criterion are marked in the figures in the results section using black bars located along the bottom of the plots.

5.3 RESULTS

Throughout the pupil datasets there was a large variability in the amplitude of responses between participants. This may reflect individual differences or simply differences in the lighting conditions during the experiment, as these were not explicitly controlled. Since the analyses were all within-subject designs (based on calculating the difference series between conditions for each participant) these effects did not impact the statistical results. However, when plotting the pupil dynamics for each condition separately the error bars appear large, and it is difficult to use them to judge significance as you typically can in the other figures presented in this thesis. Thus, for clarity, I have included additional plots within each figure (see for example, Figure 5.2, bottom) which show the within-subject difference timeseries that the statistical analysis was based on.

5.3.1 Inattentional Deafness (T-LS)

5.3.1.1 Effect of Visual Load

I first characterised the effect of visual load on pupil response in the trials where there were no sounds presented. Figure 5.2 (top) shows the mean pupil diameter for high and low load tasks, baseline corrected relative to the 200 ms pre-stimulus period. Overall, during the stimulus and response period, both the pupillary waveforms were characterised broadly by a dilation, which peaked at around 1400 ms post stimulus onset, and then a return towards the baseline. After the end of the trial there was a sharp constriction, which occurred in response to the fixation cross that preceded the next trial.

There were two periods when the pupil diameter differed significantly between load conditions. Early in the response, while the overall trajectory for both load conditions was a dilation, in high load there was a brief constriction and recovery around 500 ms post stimulus onset, resulting in a significant difference in pupil diameter between low and high load from ~450-700 ms. This swift pupil constriction shortly after presentation of a visual stimulus is characteristic of the pupillary light reflex (Ellis, 1981), which regulates the amount of light entering the eye. It is distinct from pupillary responses to stimulus processing, which are relative dilations and occur more slowly and with a much smaller magnitude than light-driven changes (Beatty and Lucero-Wagoner, 2000). The light-driven constrictions in the pupil depend on overall luminance, having lower latency and higher amplitude in response to large changes in brightness (Alpern et al., 1963). In the letter search task used here, the manipulation of load resulted in a display that was brighter in the high load than the low load condition

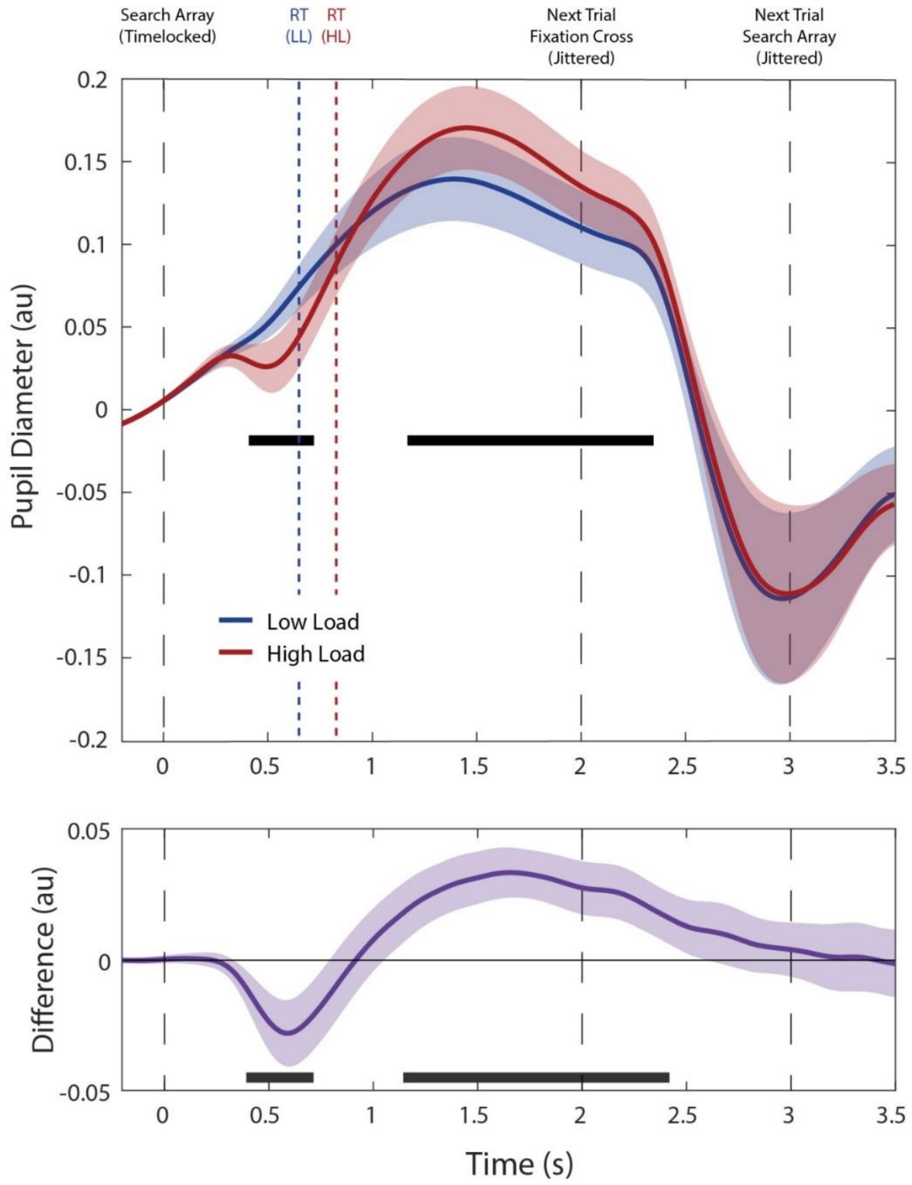


Figure 5.2: T-LS (Experiment 1) - Main Effect of Load on Pupil

TOP: Mean pupil diameters for high and low load (SA trials only). BOTTOM: Difference timeseries for HL-LL. In this and subsequent figures in this chapter, error bars illustrate the SD of the bootstrap across participants for the condition they are plotted about – this may be an individual condition or a difference timeseries, and in the latter case the error bars illustrate the bootstrap of the difference timeseries, reflecting a within-subject contrast. Horizontal black bars indicate timepoints when the pupil diameter showed a significant difference between displayed conditions, vertical dashed black lines indicate the beginning and end of each trial, and blue and red dashed lines mark the mean RTs for low and high load trials, respectively.

(low load non-targets were substantially smaller than high load non-targets), so this early difference in pupil diameter likely reflects more constriction in response to the high load display than the low load display.

It is important to note that the difference in pupil diameter driven by the luminance (more constriction in high load, i.e., smaller pupil diameter) is the opposite to that seen later in the trial, and therefore differences in the physical stimuli cannot account for the later effect. Further, in subsequent sections where I report the effects of load on responses to sound, all waveforms were calculated as a subtraction within load conditions (for example sound present - absent), so any pupillary response driven by the visual stimuli was removed and will not have impacted the results.

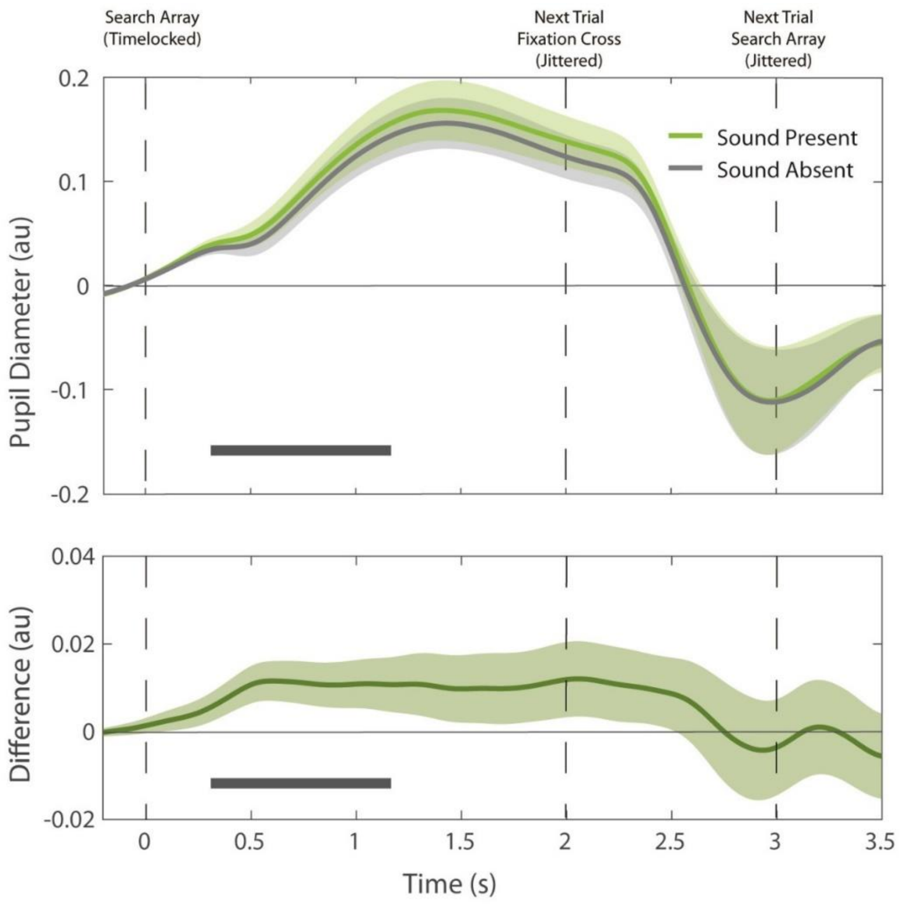
Later in the trial, during the peak of dilation, the high load task resulted in greater pupil dilation than the low load task from ~1200-2350 ms post-stimulus onset. This difference was largest at ~1650 ms post stimulus onset, when the pupil diameter in high load was 0.030 units larger than in low load. This effect of larger pupil dilation during a task with higher demand is consistent with the research outlined in the introduction to this chapter. However, since the effect occurred after participants had given the button response to the task (and the response times differed between low load, m=648 ms, SD=84, and high load, m=826 ms, SD=105) the difference in pupil diameter during this period could potentially be associated with the behavioural response rather than a difference in perceptual processing.

5.3.1.2 Effect of Sound

Here I contrast the pupil diameter in trials where a sound stimulus was present (SP) vs absent (SA), regardless of the load of the visual task. Figure 5.3 shows that the

SP trials (green) followed the same overall shape as the SA trials (grey; note that these are the trials described in the previous section but collapsed over LL and HL), with dilation and recovery over the stimulus and response period, followed by a constriction after the presentation of the fixation cross preceding the next trial. There was a significant difference in the pupil diameter between SP and SA trials during the main dilation response from ~300-1200 ms post stimulus onset, with the SP trials showing greater dilation during these periods than the SA trials. This effect of relative dilation in response to an auditory stimulus (versus no auditory stimulus) is similar to that seen in auditory detection paradigms described in the introduction to this chapter. However, in contrast to those experiments, the sounds used here were irrelevant to the participants and were presented in conjunction with a visual stimulus, so the interpretation of the pupillary response is not as straightforward.

The fact that the relative dilation in SP trials reached significance from very early in the trial suggests that it may in part reflect an orienting response. These responses are believed to be mediated by circuits in the intermediate layers of the superior colliculus (Wang and Munoz, 2015) and are characterised by dilations which occur earlier than those driven by load on cognitive processing. These responses scale with the saliency of the stimuli (Wang et al., 2014), so the fact that the effect seen here was small in magnitude likely reflects the fact that the auditory stimuli were very faint, simple, pure-tones, and that they were presented frequently and were task-irrelevant. However, since the dilations persisted throughout the trial, they may also reflect some degree of cognitive processing or perception.



<u>Figure 5.3: T-LS (Experiment 1) - Main Effect of Sound on Pupil</u>

TOP: Mean pupil diameters for SP and SA trials separately (collapsed over load conditions). BOTTOM: Difference timeseries for SP-SA trials.

5.3.1.3 Effect of Load on Response to Sound

To characterise the effect of load on the response to the tones, I calculated the difference between SP and SA trials separately for high and low load trials (Figure 5.4). While I refer to these difference timeseries as the ‘response to sound’, it is important to note that in fact that are the response to an additional auditory signal (on top of the visual stimulus), and thus may represent some activity which reflects multisensory stimulus processing.

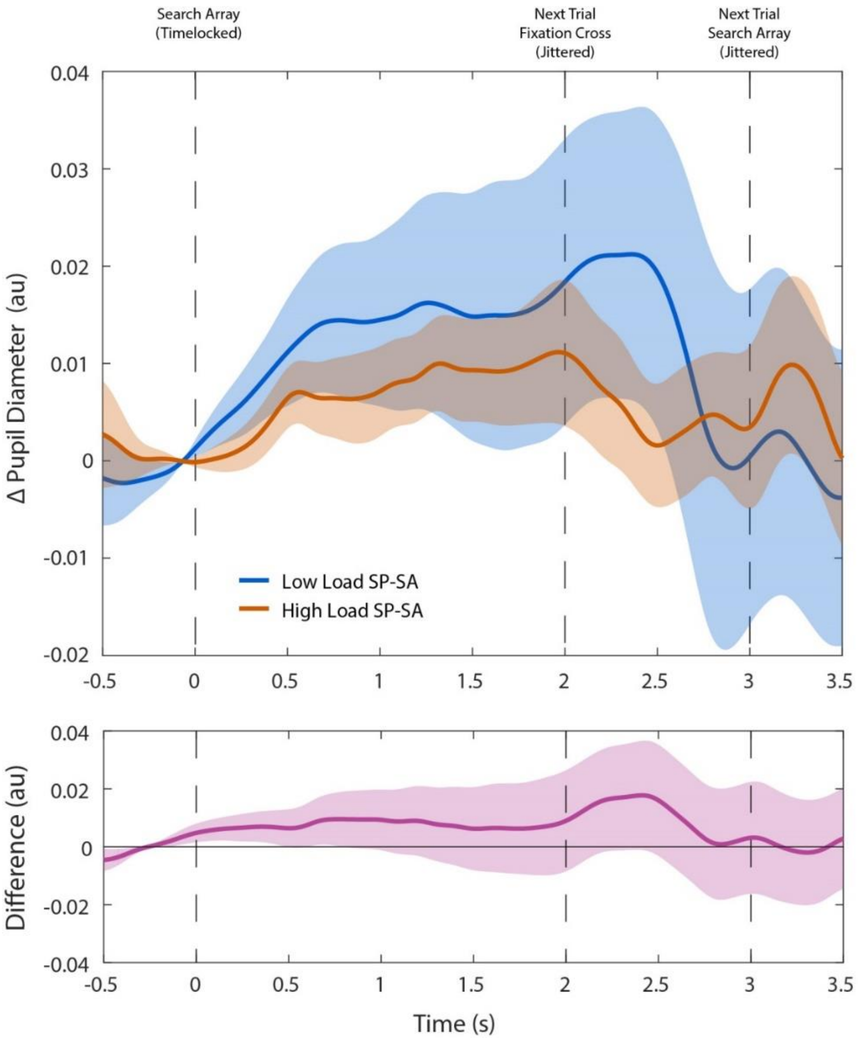


Figure 5.4: T-LS (Experiment 1) - Effect of Load on Pupil Response to Sound

TOP: Difference timeseries for SP - SA trials (i.e. the pupil response which was specific to sound) for low load and high load conditions separately. BOTTOM: Interaction timeseries for LL(SP-SA) – HL(SP-SA).

There were no significant differences between the low and high load pupil waveforms. This is in contrast to the MEG data from this experiment which suggested that high load suppressed cortical processing of the tones, and the behavioural data which indicated poorer behavioural sensitivity to the sounds under high load. The present data showed a similar effect numerically (with larger dilation to the auditory stimuli under low compared to high load), and since the pupil dilations to sound were overall relatively small (due to the very faint sounds used) it is unclear whether this result reflects a genuine null effect or that the data collected here were too noisy or the sample size too small to capture a small effect.

5.3.2 SFG Letter Search (SFG-LS)

5.3.2.1 Effect of Visual Load

Figure 5.5 shows the mean pupil diameter during the high and low load letter search tasks. The waveforms showed a different shape to those seen in the **T-LS** dataset: the initial response to the onset of the stimulus array was a substantial constriction which peaked at ~800 ms post stimulus onset. In contrast to the **T-LS** dataset, here the amplitude of the light-driven constriction was notably larger and did not differ between low and high load conditions (in spite of the slight difference between the stimulus arrays). Both differences are likely to be due to the addition of the 100 ms mask in the **SFG-LS** paradigm, which had a larger area and was brighter than the stimulus arrays. Photopupillary responses are dependent on the overall energy (i.e. both intensity and duration) of a light stimulus (Alpern et al., 1963), so the substantially larger constriction in the response here relative to the **T-LS** task likely reflects that the visual stimuli were twice the duration and overall brighter in this experiment. Similarly, since the constriction response here is likely to have been heavily influenced by the mask (which was identical between load conditions), the fact that the small difference in the search arrays was not reflected in the amplitude of the constriction response here is not unduly surprising.

Many pupillometry studies into the effects of cognitive processes on the pupil are performed in a well-lit room to minimise these stimulus-driven effects; however the present experimental set up required the use of a projector due to the electrically-shielded MEG chamber, and so to achieve good visual clarity on the display screen the lights were dimmed during the experiments.

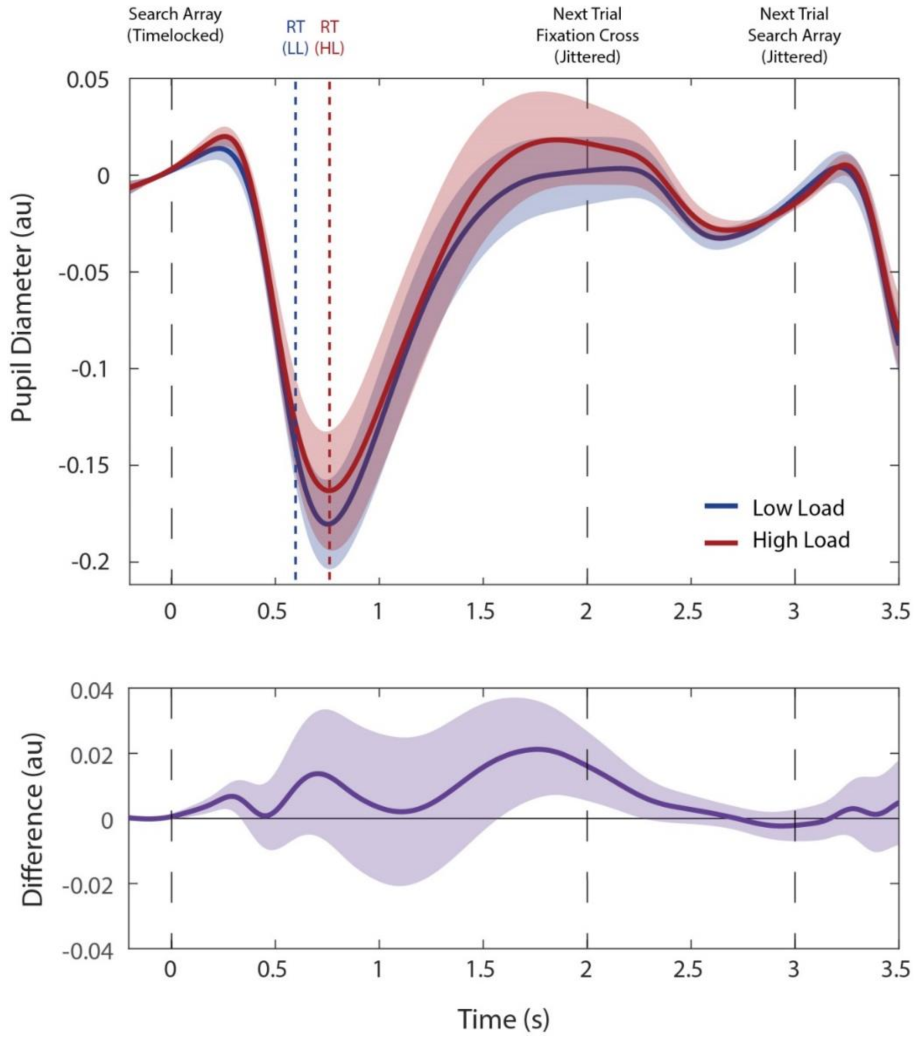


__Figure 5.5: SFG-LS (Experiment 4) - Main Effect of Load on Pupil__

TOP: Mean pupil diameters for HL and LL (SA trials only). BOTTOM: Difference timeseries for HL-LL.

Following the recovery from the early constriction, at about 1500ms responses in both conditions plateaued. It was during this timeframe that the effect of load was seen in the **T-LS data**, where there was more dilation under high load. However, while the same pattern is seen in the present data, the effect was not significant. Given that this experiment used the same visual task that did show a significant effect on pupil in the **T-LS** dataset, and, as I will describe below (Section 5.3.3.3), there was a significant effect of load on the response to sound within this dataset, it is possible that the difference seen during this period does reflect a real effect of load that the dataset was not sufficiently powered to confirm.

5.3.3.2 Effect of Sound

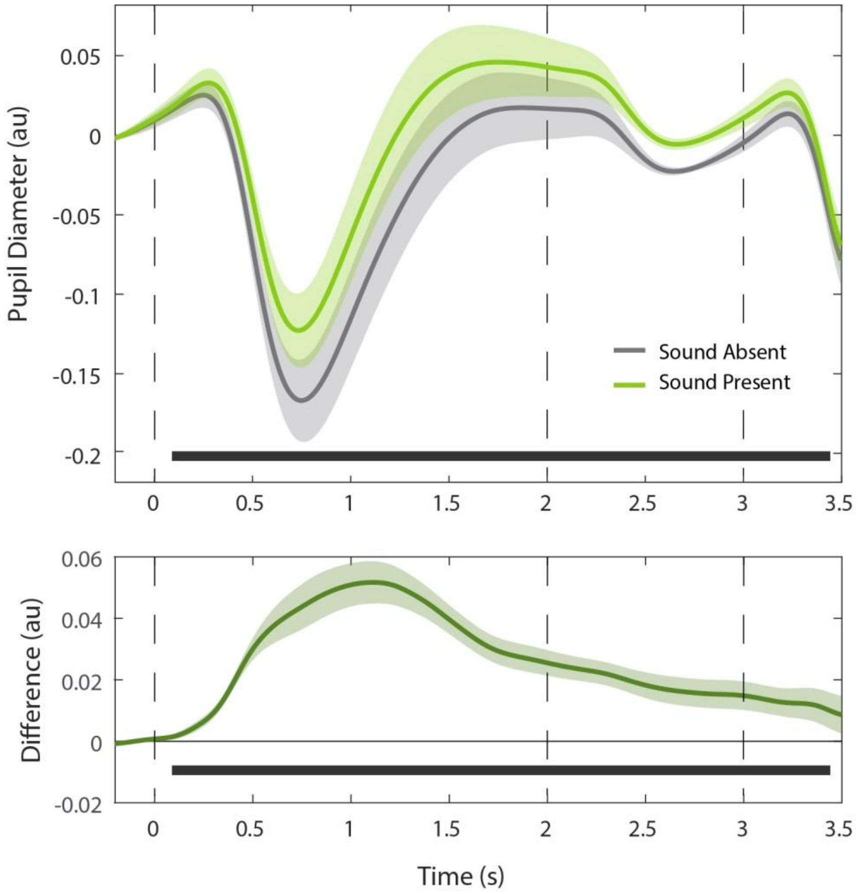


Figure 5.6: SFG-LS (Experiment 4) - Main Effect of Sound on Pupil Response

TOP: Mean pupil diameters for SP and SA trials separately (collapsed over load conditions). BOTTOM: Difference timeseries for SP-SA trials.

Contrasting the pupillary waveforms in trials when the SFG stimulus was present versus absent revealed a significant dilation in response to sound (Figure 5.6), with larger pupil diameter occurring from ~100-3500 ms post stimulus onset in trials when a sound was present relative to absent. This dilation effect in response to an auditory stimulus is consistent with the data from the **T-LS** experiment, though here it began earlier, was numerically larger, and persisted longer than the effect seen in the previous dataset. The stronger effect is likely to be due to differences in the auditory stimuli used: compared to the near-threshold pure tones used in the ID experiment, the SFG signals used in this experiment were considerably louder, more complex and were twice the duration.

5.3.3.3 Effect of Load on Response to Sound

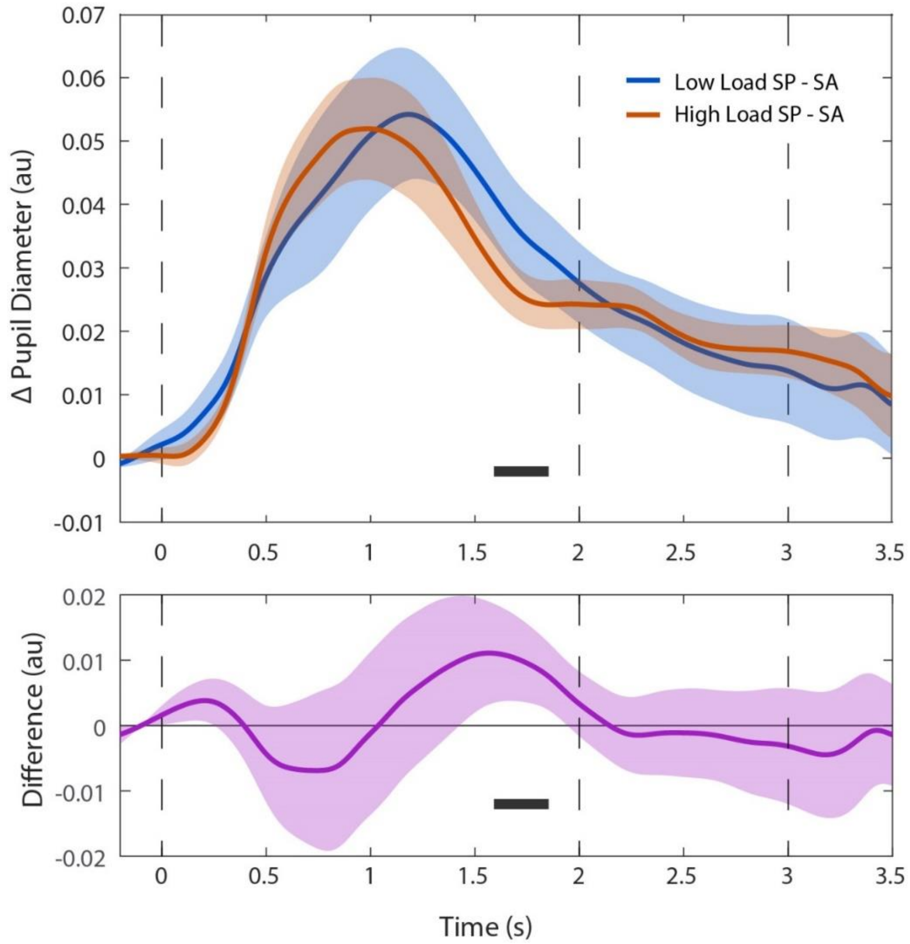


Figure 5.7: SFG-LS (Experiment 4) - Effect of Load on Pupil Response to Sound

TOP: Difference timeseries for SP - SA trials (i.e. the pupil response which was specific to SP trials) for low load and high load conditions separately. BOTTOM: Interaction timeseries for LL(SP-SA) – HL(SP-SA).

The pupil responses to sound (i.e. SP trials - SA trials) showed a significant effect of visual task load, with a larger dilation in response to sound under low load than high load from ~1600-1850 ms (Figure 5.7). The relatively long latency of the effect is consistent with cognitively-driven effects on pupil diameter (as opposed to luminance-driven activity), and so is likely to reflect reduced processing of auditory signals under high load. This is consistent with the MEG data collected simultaneously, which showed that there was an effect of load on overall response to sounds during both the P2 and P3 responses.

5.3.3.4 Effect of Load on Response to Auditory Figure

To explore the effect of load on the pupil response which occurred specifically to the repeated figure within the auditory stimuli, I calculated the difference timeseries of figure present (FP) trials – figure absent (FA) trials for high and low load separately (Figure 5.8A). There were two periods when the responses showed a significant difference between load conditions. The first was from -38 to 300 ms relative to the stimulus onset, and was likely an artefact of the fact that the LL and HL timeseries had opposite trajectories in the pre-stimulus interval: the mismatch between trajectories meant that even though the two waveforms were baselined based on their average values during this period, the responses diverged significantly even before the trial had begun. The difference in trajectories suggests that the inter-trial intervals used were not sufficient for the pupil diameter to stabilise fully between stimulus presentations, though this problem may have been confounded by the fact that the waveforms in the present analysis were difference timeseries rather than reflecting a concrete pupil diameter. The fact that the effect started before the stimuli indicated that it cannot reflect stimulus processing of the auditory signals in the 'current' trial.

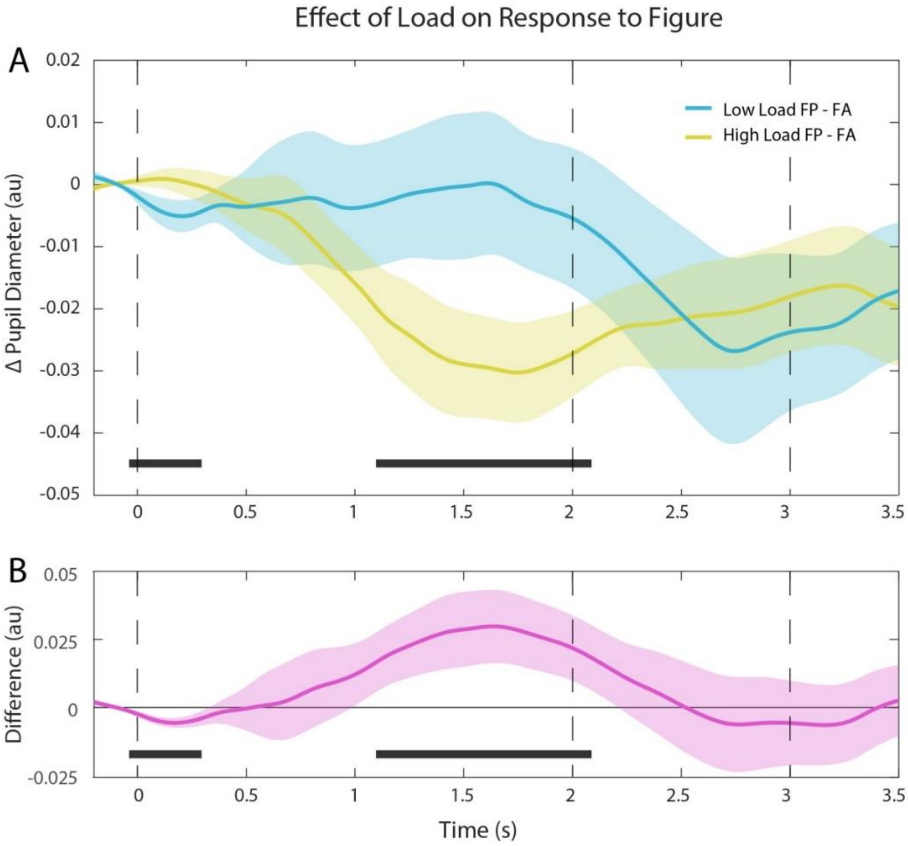


__Figure 5.8: SFG-LS (Experiment 4) - Effect of Load on Pupil Response to Auditory Figure__

A: *Difference timeseries for FP - FA trials (i.e. the pupil response which was specific to the figure) for low load and high load conditions separately.* ***B:*** *Interaction timeseries for LL(FP-FA) – HL(FP-FA).*

During the second significant period (~1100-2100 ms post stimulus onset) the presence of the figure within the auditory stimuli was associated with relatively more dilation under low load than high load (more positive waveform visible for low versus high load in Figure 5.8A), which may reflect increased processing of the figure in low load conditions. However, it is important to rule out the baselining difficulty as a contributor to the later effect. To explore whether the baselining could have had an effect, I ran the trial analysis using a range of different baseline periods, and the effect remained significant in all cases. Though this is not conclusive, the fact that a matching effect was shown in the next dataset (Section 5.2.3.4) lends some legitimacy to the effect.

5.3.3 SFG Colour-Shape Task (SFG-CS)

5.3.2.1 Effect of Visual Load

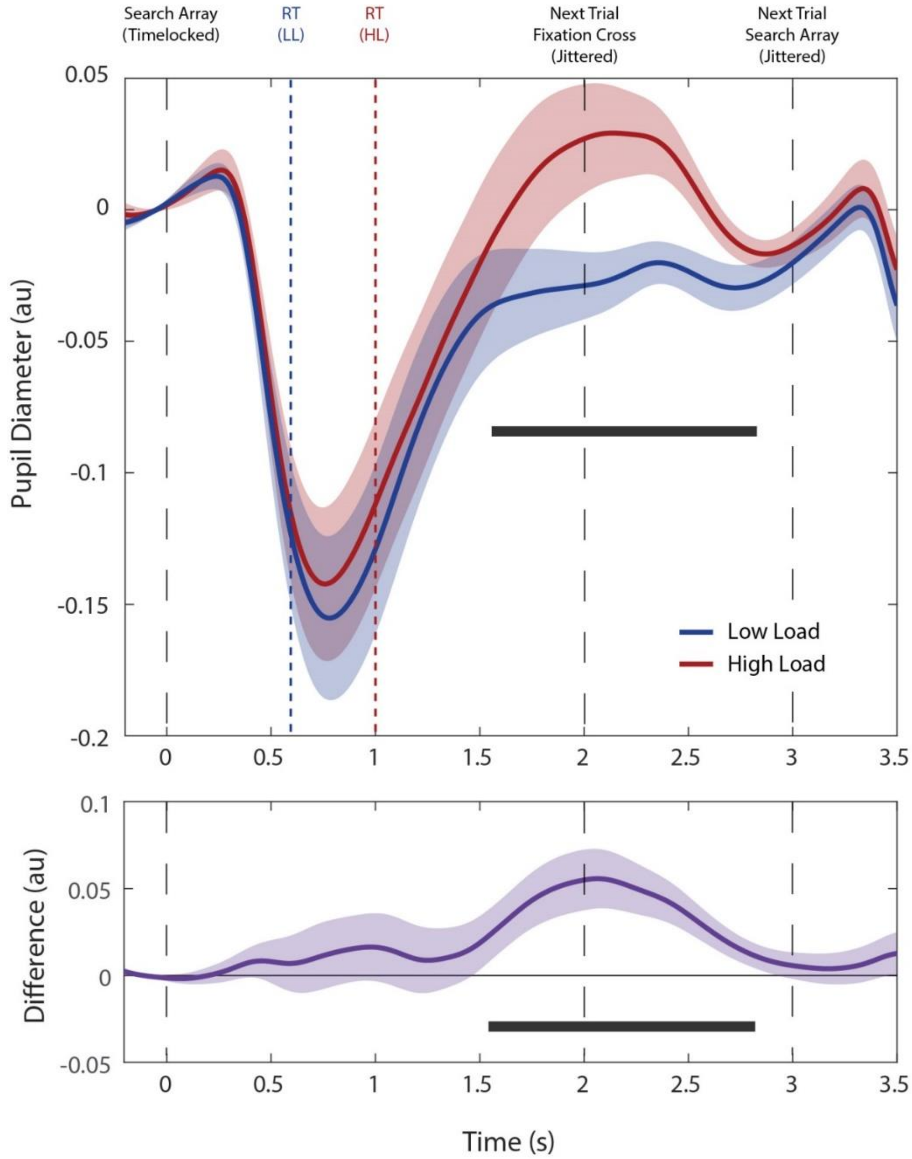


Figure 5.9: SFG-CS (Experiment 5) - Main Effect of Load on Pupil

TOP: Mean pupil diameters for high and low load (SA trials only). BOTTOM: Difference timeseries for HL-LL.

Figure 5.9 shows the mean pupil diameter during the high and low load colour-shape tasks. The overall waveforms were similar in character to those seen in the **SFG-LS** dataset, with a substantial luminance-driven constriction peaking at ~800 ms post stimulus onset, followed by recovery to baseline. Following this, at ~1500 ms the low load response plateaued, whereas in high load the pupil continued to dilate, leading to a significant difference between the responses from 1500-2800 ms post stimulus onset. This finding of relative pupil dilation in high versus low load late in the response is consistent with the effect seen in the **T-LS** experiment (though the difference between load conditions here was larger and reached significance roughly 500 ms later) and with the studies reviewed in the introduction to this chapter that show that pupil dilation increases with task demand.

Interestingly, as discussed in Chapter 4, the MEG data for the visual responses to the LS task and CS task showed opposite effects of load on early visual evoked activity. I have argued that these differences were due to the different processes involved in the letter search versus the colour-shape task, but that both reflected a manipulation perceptual load. Here, the pupil data appear to support this, with both experiments showing relative pupil dilation in high load. However, as with the effect in the **T-LS** data, this effect occurred after behavioural responses occurred, and so may have been driven to some extent by decision- or response-relates processes.

5.3.2.2 Effect of Sound

Contrasting the pupillary waveforms in trials when the SFG stimulus was present versus absent revealed a significant dilation in response to sound (Figure 5.10). Trials where an auditory stimulus was present showed significantly larger pupil diameter from ~200–3000 ms post stimulus onset relative to the trials when there was no auditory stimulus, an effect consistent with both previous datasets.

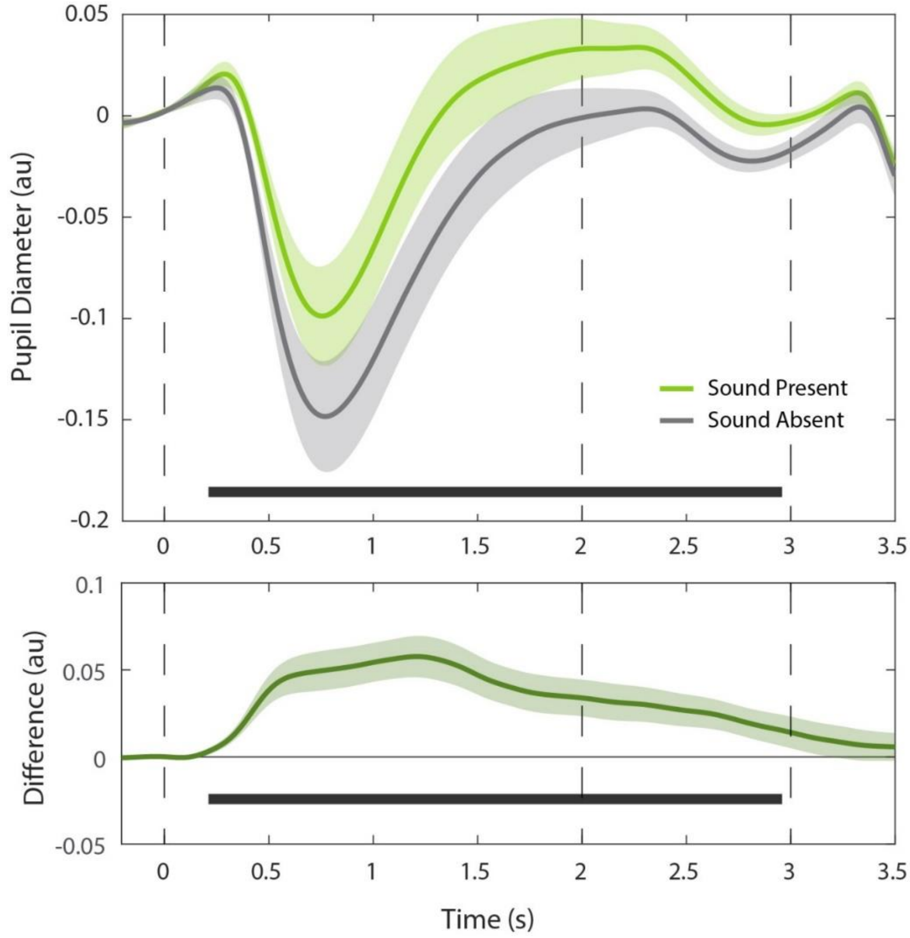


__Figure 5.10: SFG-CS (Experiment 5) - Main Effect of Sound on Pupil Response__

TOP: Mean pupil diameters for SP and SA trials separately (collapsed over load conditions). BOTTOM: Difference timeseries for SP-SA trials.

5.3.2.3 Effect of Load on Response to Sound

In contrast to the **SFG-LS** dataset, here there was no significant effect of load on the overall pupillary response to sound (Figure 5.11). Recalling the effects in the MEG datasets for the two experiments, both were found to have significant effects on the P3 responses (which index conscious perception), while only the **SFG-LS** design resulted in an effect on an earlier, sensory ERF component (M200). Speculatively, it is therefore possible that the pupil diameter measure here reflects sensory processing of the auditory signals, independent of later factors associated with awareness.

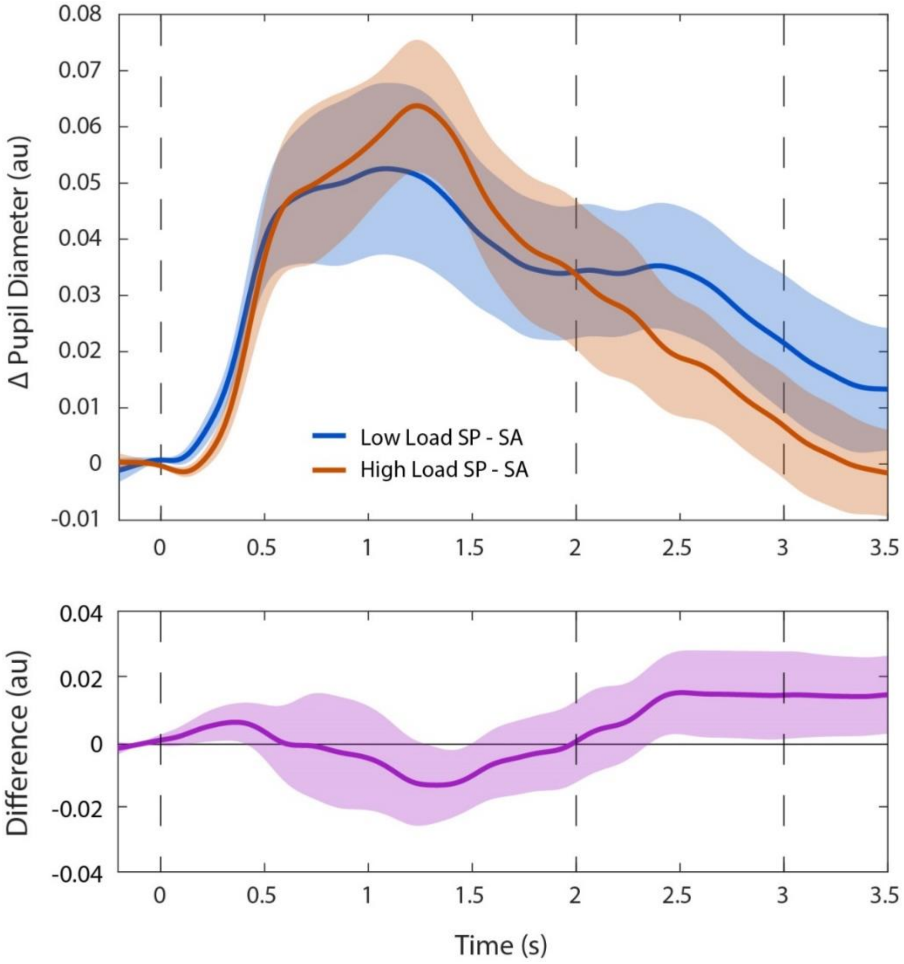


__Figure 5.11: SFG-CS (Experiment 5) - Effect of Load on Pupil Response to Sound__

TOP: Difference timeseries for SP - SA trials (i.e. the pupil response which was specific to SP trials) for low load and high load conditions separately. BOTTOM: Interaction timeseries for LL(SP-SA) – HL(SP-SA).

5.3.2.4 Effect of Load on Response to Auditory Figure

Figure 5.12A illustrates the pupil response to the auditory figure under low and high load. There was a significant effect of load on the response to the figure from ~1500-1900 ms post stimulus onset: the presence of the figure within the auditory stimuli was associated with more pupil dilation under low load than high load.

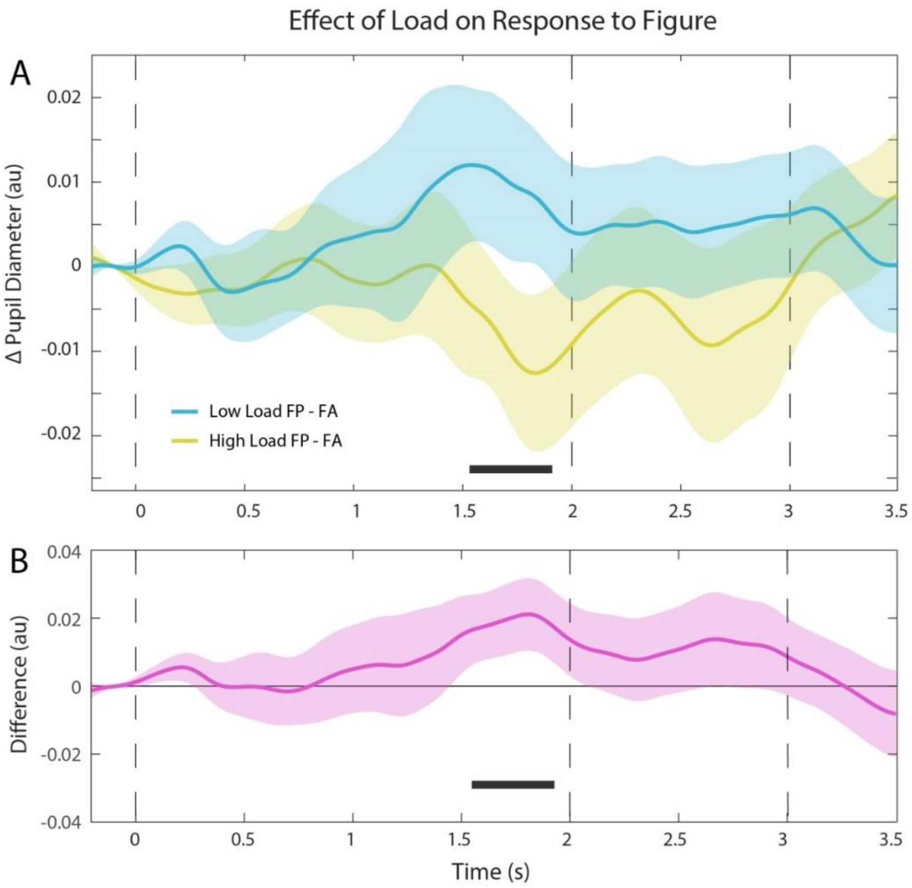


__Figure 5.12: SFG-CS (Experiment 5) - Effect of Load on Pupil Response to Auditory Figure__

***A:** Difference timeseries for FP - FA trials (i.e. the pupil response which was specific to the figure) for low load and high load conditions separately. **B:** Interaction timeseries for LL(FP-FA) – HL(FP-FA).*

This effect is consistent with the result seen in the **SFG-LS** dataset and with the MEG data from both experiments, which showed reduced processing of the figure under high load. Comparing the duration and size of the load effect on pupil response to the figure between the **SFG-CS** and **SFG-LS** datasets suggests that the load manipulation in the letter search task was more effective at disrupting auditory processing than the colour-shape search task, which is also consistent with the relative magnitude of the effects in the MEG data. This suggests that pupil diameter may serve as an index to reflect the level of computational processing that occurs to irrelevant sounds, and/or act as an index for the degree to which an auditory signal is perceptually segregated.

5.4 DISCUSSION

5.4.1 Summary of results

The pupillometry data from the three experiments showed results which (though noisy) broadly mirrored each other across the three datasets, including the MEG and behavioural data that were collected simultaneously. Pupil dilation in response to the visual tasks was systematically larger under high perceptual load conditions, lending support to previous reports that pupil diameter indexes perceptual demands and, interestingly, that it appears to index cognitive demands in a similar manner. Secondly, there was a significant pupillary dilation in response to the task-irrelevant auditory signals: dilation was generally larger under low-load conditions, in keeping with the MEG data which illustrate reduced processing of the auditory signals under high load. Finally, pupil diameter showed larger dilation responses to the auditory figure in low versus high load, mirroring the MEG indices of the degree to which auditory signals were segregated. Though some of the analyses suggested that certain data were underpowered, with two null results that may have been missed small effects, generally the data were consistent between experimental datasets.

5.4.2 Pupil dynamics under visual perceptual load

In this section I consider the data from trials where only visual stimuli were presented, and discuss how the load of the tasks affected pupil dilations. According to a host of previous work (reviewed in the introduction to this chapter) increased processing load in a task is typically reflected by larger pupil dilation. This pattern was seen numerically in all three datasets, though the effect did not reach significance in the **SFG-LS** experiment. This may have been due to the fact that the load manipulation was somewhat weak in this experiment (as noted in Chapter 4) which in fact led me to abort data collection slightly short of the planned cohort.

Interestingly, the pupil responses showed a consistent effect of load between the two visual tasks, in spite of the fact that the MEG data indicated that the two tasks had substantially different neural dynamics. This suggests that pupil diameter is a robust index of perceptual load, regardless of the precise processes which underlie the tasks. In support of this conclusion, the small number of previous studies which have explicitly investigated the effect of perceptual load on pupil dynamics (Kahneman et al., 1967; Porter et al., 2007; Alnæs et al., 2014), each of which used a different manipulation of perceptual load, all found that pupil dilation increased with perceptual load.

In addition to the general pattern of larger pupil dilations with increased demand, previous studies have found that the magnitude of the peak pupil dilation increases systematically with load (Hess and Polt, 1964; Kahneman and Beatty, 1966; Kahneman et al., 1967; Bradshaw, 1968; Payne et al., 1968; Peavler, 1974; Granholm et al., 1996; Marshall, 2002; Klingner et al., 2011; Alnæs et al., 2014). A similar effect was apparent in the present datasets: when the three load manipulations were ranked for the strength of the load manipulation (as indexed by the increase in reaction times from low to high load, where the **SFG-CS** design illustrated the strongest effect, followed by the **T-LS** design and then the **SFG-LS** design), this order corresponded to the magnitude of the load effect on pupil dilation.

5.4.3 Pupil response to irrelevant sounds

5.4.3.1 Early pupil dynamics reflect orienting responses

In all experiments, pupil diameters were larger in trials when a sound was present relative to those which contained only a visual stimulus. The dilation effects reached significance from very early in the trial, suggesting that they are partially ‘orienting responses’, which occur earlier than dilations driven by load on cognitive processing. These orienting responses occur in response to both auditory and visual stimuli, though at lower latencies and with smaller amplitudes for auditory stimuli (Bell et al., 2004), and are understood to be mediated in the superior colliculus (Wang and Munoz, 2015). For audiovisual stimuli there is a sub-additive effect, where the response amplitude is greater than that seen to either unimodal stimulus but smaller than the direct sum of their amplitudes, and the response occurs with the lower latency associated with auditory stimuli (Wang et al., 2014). This means that the responses calculated here (a subtraction of the responses in the visual trials from the audiovisual trials, which I refer to as the response to sound) should accurately represent the latency of responses which would have occurred to the auditory stimuli alone, but might have smaller amplitude. This is preferable to if there had been a supraadditive response, in which case some of the effect that was driven by the multisensory nature of the stimuli may have been falsely attributed to auditory processing.

5.4.3.2 Later pupil dynamics reflect sensory processing

The dilations in all datasets persisted to latencies at which dynamics are likely to reflect sensory processing. Since pupil diameter apparently indexes the processing load involved in active tasks, as established by the pupil responses to the visual tasks presented here, it is plausible that it would also act as an index for the degree to which other stimuli are processed, even if they are outside the explicit task set. The effect of load that was apparent in the pupil responses to the figures in both SFG experiments supports this idea: pupil dilation (which is associated with increased processing) that occurred in response to the figures was larger under low load than high load, suggesting that pupil diameter might be used as an index of processing even of irrelevant stimuli.

If this is the case, pupil diameter should also reflect the overall effect of load on sensory processing of the auditory signals. The MEG data from the experiments indicate that load had a significant effect on ERF components that reflect sensory cortical processing (i.e. M50, M100 and M200) in the **T-LS** experiment (in which the M100 was affected by load) and the **SFG-CS** experiment (where the M200 was affected), but not the **SFG-LS** experiment (in which there was an effect on the FRN and the P3, but not the onset peaks during the analysis of the overall response to sound – see Section 4.2.1.2). The pupil responses from the SFG datasets showed the same pattern, with a significant effect of load in the **SFG-LS** experiment but not the **SFG-CS** experiment. In the **T-LS** experiment, there was no significant effect of load on pupil response to sound. However, it is worth noting that the auditory signals used in the **T-LS** experiment were extremely faint, meaning that the pupillary response to them was, unsurprisingly, relatively small and noisy, and that the pupil data did show a pattern consistent with the expected effect, but which fell just short of the significance criterion I had pre-defined. Thus, it is possible that pupil diameter does typically reflect the degree of sensory processing of irrelevant sounds, but with less sensitivity than is provided by the rich data provided by MEG recordings.

The fact that there is strong evidence that figure processing (and, by proxy, auditory computation) was represented in pupil dynamics, and some evidence that pupil responses also reflected the degree to which auditory signals (regardless of whether a figure was present) were processed, suggests that the pupil could serve as a useful index not only of task-related processing, but also the degree to which concurrent irrelevant signals are processed.

5.4.3.3 Pupil responses and awareness of irrelevant stimuli

As discussed in the introduction to this chapter, there is evidence that pupil diameter reflects subjective awareness of auditory stimuli, for example showing more dilation in trials where a stimulus is detected (Beatty, 1982; Beatty and Lucero-Wagoner, 2000; O'Neill and Zimmerman, 2000).

From the datasets presented here it is not possible to determine the degree to which pupil diameter reflects explicit awareness of stimuli outside the focus of attention, since there were no behavioural responses to the sounds. Nonetheless, the degree of awareness associated with the irrelevant signals may broadly be illustrated by the amplitude of the P3 responses from the MEG data. However, the data from the P3 response do not appear to match the pupil responses particularly well: the MEG data illustrate a significant effect of load on the P3 response to sound in all three experiments, whereas the pupil data in response to overall sound only showed a significant effect in the **SFG-LS** dataset. While I have argued that in the T-LS dataset this might be a consequence of the very faint auditory signals, the SFG stimuli used in the SFG-CS task were much louder. Further, while the T-LS data show a pattern that is consistent with the expected effect, the SFG-CS dataset do not. The SFG-CS also showed the strongest effects of load on both the visual responses and to the figure, so it is difficult to dismiss the null effect in the other contrast. As I have argued above, it seems more likely that the pupil dynamics for overall responses to sound under load reflect sensory processing of signals.

One explanation for this pattern of effects would be that although there were P3 responses elicited by sounds under low load in all three experiments, this does not necessarily mean they rose to subjective awareness. As reviewed in the introduction, Load Theory proposes that along with automatic allocation of perceptual resources, cognitive control resources can be employed, if available, to effect a late selection filtering in which stimuli that have been processed with spare perceptual capacity but are task-irrelevant are prevented from progressing to the stage of awareness. Since cognitive load in all the tasks used here was kept intentionally low, it is likely that cognitive control resources would have been available to perform such a function. Thus, unlike in the dual task experiments, where the auditory signals were relevant, during the MEG experiments participants may not have experienced a high level of awareness to the auditory signals. This is supported by the fact that there was no evidence of attention capture in the experiments; behavioural performance on the visual task was equally good in trials when the auditory stimuli were present and absent.

5.4.4 Conclusions and limitations

In spite of the fact that pupil data presented here were collected as secondary measure, the results showed some interesting effects. The pupil appeared to be sensitive to a number of factors relating to the task and stimuli, including the physical aspects of the visual stimuli and the load involved in the visual task, and more surprisingly pupil diameter responded reliably to the presence of irrelevant sounds, and even provided some indication of the level of processing that occurred. Crucially, these factors were all reflected simultaneously and were dissociable from one another, even though pupil diameter is a single channel of data.

However, while these preliminary findings are interesting, it is vital that the effects are replicated using paradigms which are explicitly geared towards collecting high quality pupil data. As noted previously, there are several aspects of the present experiments that were not optimal for pupillometry, including the lighting conditions, the fact that trials were relatively short and so may not have allowed the pupil to fully return to baseline between trials, and that behavioural responses may have had an impact on pupil dynamics. Additionally, due to the hardware set up, differences in baseline pupil diameter between experiments and conditions could not be assessed. It is very likely that there was also a change in participants' overall state between low and high load blocks, which would be reflected in a change in baseline diameter of the pupil. The experiments presented in this thesis were not designed to capture these gross effects (a point which I will elaborate on in the general discussion of the thesis), and generally the short- and long-term effects of load can be investigated independently. However, since the pupil has physical limits on its size, when the baseline diameter is very large there is a relatively small potential range within which dilations can occur, leading to muted phasic dynamics (Rajkowski et al., 1993). Thus, it is important to bear in mind that the results presented here may have been confounded if there were large differences in baseline pupil diameter between experiments or blocks. These issues could be straightforwardly addressed by additional experiments in conditions designed for pupillometry, and the results would provide important clarification as to whether the conclusions presented here are valid.

CHAPTER 6: GENERAL DISCUSSION

6.1 SUMMARY OF FINDINGS

The data presented in this thesis establish that visual perceptual load can impact cortical sensory processing of unattended auditory signals. The data from Chapter 2 illustrate the neural underpinnings of load-induced inattentional deafness, establishing that load on visual perceptual resources can reduce the amplitude of auditory evoked responses at the level of both cortical sensory processing (as indexed by the M100 response) and during later, awareness-related stages. The timing of the effects suggest that resources are allocated dynamically based on demand in a 'push-pull' manner rather than acting as a pre-emptive suppression of all auditory activity. In Chapters 3 and 4 I explored whether the effect of load is limited to a straightforward dampening of evoked responses or whether it can also lead to a reduced capacity for computational processing within the auditory system. Using signals that allow for the process of auditory figure-ground segregation via temporal-coherence cues, I showed that conditions of high visual perceptual load result in reduced auditory computational processing, as indexed by the degree to which an evoked component elicited by the figure within the sounds was evident in the neural responses. This effect was replicated using two different visual load tasks. The temporal locus of the effect and the neural source estimates both suggest that visual load impacted processing within auditory cortex, rather than at higher levels of the processing hierarchy where auditory and visual information are known to be integrated, suggesting that auditory signal processing in early cortical areas is subject to limitations on domain-general resources.

In Chapter 5 I presented pupillometry data recorded during the MEG load experiments reported in previous chapters, and showed that pupil dilation can reflect both the level of processing load in the active visual task and potentially the degree to which the irrelevant auditory stimuli were processed. This highlights the potential use of pupillometry as an objective measure of cortical processing in situations where neural imaging is not viable.

6.2 REVISITING THE UNDERSTANDING OF 'AUTOMATIC' AUDITORY PROCESSING

Many auditory processes have been shown to occur passively, without the need for attention to be explicitly engaged. A popular view based on these findings is that audition acts as an early warning system that pre-attentively processes the auditory scene and can trigger the capture of attention if this analysis suggests a behaviourally relevant change has occurred. However, in vision many aspects of processing have been shown to be able to benefit from the automatic allocation of spare perceptual capacity, allowing for the processing of stimuli that are outside the explicit focus of attention. These processes may appear automatic, but in fact are susceptible to resource depletion in conditions of high perceptual demand. The preferred methodologies used in previous auditory research had meant it remained unclear whether the auditory system can benefit from spare capacity in this manner, and this potential confound undermines much of the evidence upon which previous claims of automatic auditory processing have been based.

The results presented here suggest that in fact many fundamental aspects of auditory processing are not automatic, with evidence that both the overall strength of auditory evoked responses, and the capacity of the auditory system to process information can be impaired when the load of a concurrent visual task is high. It appears that the conditions necessary for these effects to occur (at least to the extent where they are measurable in experimental conditions) require precise co-incidence between the visual load task and the critical auditory stimuli. In practice, in everyday listening situations this is unlikely to occur, and the auditory system may be able to make up for moments of reduced processing capacity due to its ability to make use of 'glimpses' of information when available. However, the effects reported here highlight the need for precise experimental control of resource availability in studies that aim to explore the reliance of specific auditory process on attention.

Though I have demonstrated the reliance that the mechanism for auditory segregation via temporal coherence has on the availability of resources, the question of whether this applies to other aspects of auditory processing remains open. To establish whether similar limitations apply to other auditory processes requires the systematic application of load tasks paired with irrelevant auditory signals that characterise a variety of auditory processes. Although this approach has been used in a few previous studies (e.g. Alain and Izenberg, 2003; Dyson et al., 2005), revealing that there is apparently no effect of load on processes such as segregation via harmonicity cues or repetition cues, and detection of the violation of a regular sequence (Chait et al., 2012), many of these studies used tasks where the level of perceptual load may not have been sufficiently high to maintain resource depletion consistently over an extended period. In particular, the temporal integration mechanisms via which segregation is achieved in the paradigm used by Masutomi et al. (2016), in which an effect of auditory perceptual load, but not visual perceptual load was illustrated, appears to share many characteristics with the mechanism for segregation via temporal coherence that was shown here to be susceptible to resource depletion by a visual task. It may therefore be worthwhile to see whether the null findings reported based on previous load manipulations are replicated using alternative load tasks. Another question which may be worth further investigation is whether auditory segregation based on concurrent cues (e.g., harmonicity) is governed by a distinct process to segregation of on-going sound sources, as indicated by the apparently different effects of load on the ORN and FRN. It is possible that rather than being distinct mechanisms, the features that cue segregation for concurrent stimuli require relatively little processing compared to the spectro-temporal integration required to segregate the SFG stimuli used in the present experiments; this could mean, for example, that segregation via harmonicity cues requires few resources and so is not readily disrupted in experimental settings, but is nonetheless still fundamentally resource-dependent.

6.3 PUPILLOMETRY AS A MEASURE OF PERCEPTUAL LOAD AND PROCESSING

The pupillometry data presented here support the findings from the few previous investigations of pupillary dynamics under perceptual load (as opposed to the well-established link between pupil dilation and cognitive load), which suggest that high perceptual load is reflected by increased pupil dilation in the latter stages of the pupil response. One interesting aspect of this is that the pupil data showed similar effects of task load over two experimental paradigms which had shown markedly differing task-evoked MEG responses, suggesting that pupil dilation reflects demand regardless of the precise underlying neural activity. This is in keeping with the fact that load on both cognitive and perceptual resources appear to have similar effects on pupil dynamics in spite of the fact that the systems underlying these two aspects of cognition are relatively distinct. This characteristic of pupil dilation makes it potentially useful in situations when it is important to have an objective physiological measure that can reflect the level of demand a person is experiencing regardless of the precise neural dynamics of the underlying processes.

The sensitivity of the pupil response to stimuli outside the explicit focus of attention is also potentially useful. At the moment it is difficult to investigate processing of ‘ignored’ stimuli without using neural imaging or resorting to a behavioural dual task where the secondary stimuli are not actually ignored, merely deprioritised. Yet there are numerous situations where having an objective marker of the degree to which an unexpected stimulus was processed would be useful: as technology develops there are likely to be more and more systems which aim to enhance perception, for example by highlighting important information that may have been missed due to high task load. If pupil can simultaneously measure task demand and index processing of unattended stimuli this could facilitate substantial improvements in the ability of technology to perform in a ‘smart’ manner.

6.4 QUANTIFYING PERCEPTUAL LOAD

Perceptual Load has been manipulated via a number of different methods within vision. Neuroimaging studies of how perceptual load affects visual activity have typically revealed increased activity in visual cortex in response to the load stimuli, while responses to irrelevant stimuli are reduced (Rees et al., 1997; Pinsk et al., 2004). However, the two tasks used here to manipulate perceptual load resulted in substantially different neural dynamics based on the evoked MEG signals, in spite of the fact that both were visual search tasks that manipulated load via a feature versus conjunction search. This sensitivity of MEG responses to the subtle differences in the neural dynamics involved in load-inducing tasks suggests that it may not be straightforward to establish an objective physiological measure that indexes the level of perceptual load. While the pupillary dilation response appears to serve the purpose of reflecting the perceptual load that is induced using different tasks in a coherent way, it also has the same response to increased cognitive load. Since perceptual and cognitive load have opposite effects on the degree to which irrelevant signals are processed, this means that pupil may not be ideal as an index in situations when both cognitive and perceptual systems are loaded.

The lack of a current option for an objective index that reflects load specifically on perceptual resources means that it has so far been difficult to strictly quantify the relative effects that various manipulations of load have. Ideally it would be possible to develop a firm operational formula to quantify the level of load a task induces based on its parameters, akin to the way in which videos of real-life scenes have recently been analysed to quantify the degree of load involved (Palmer et al., 2017). A formula of this sort or a physiological objective measure of the level of load could provide substantial facilitation to the efforts to establish auditory manipulations of load. Currently, with little consensus in how load should be manipulated and whether it can be expected to apply to auditory processing, the question of whether Load Theory is a good model for within-modality allocation of resources in auditory processing remains open. However, establishing a robust manipulation of auditory load may be vital in order to probe the resilience of a variety of auditory processes (most of which occur over long time periods that are not compatible with the brief load task I have established here) to depletion of perceptual resources.

6.5 IMPLICATIONS FOR THE MECHANISMS OF SELECTIVE ATTENTION

The data presented here suggest that the allocation of perceptual resources occurs dynamically according to momentary variations in task demands. This is in contrast with the conclusions drawn by Carmel et al. (2011), who explored detection of critical visual stimuli that were presented at a variety of stimulus onset asynchronies relative to a visual perceptual load task. The visual task was a feature versus conjunction RSVP search presented at the centre of the visual field and was participants' primary task. They were also asked to simultaneously detect the appearance of a critical stimulus that was presented in one of four peripheral locations. They showed that while there was main effect of perceptual load on detection of distractors, there was neither an effect of timing nor an interaction between the factors. They conclude that in vision perceptual load leads to a reduction in the excitability of visual cortex for locations outside the fovea. This is consistent with fMRI data reported by Parks et al. (2013) in which the same central task was used and cortical activity was recorded in response to concurrently presented stimuli at a variety of peripheral angles. They found that load had the largest effect on responses that were located close to the central task, with effects reducing as eccentricity was increased. This pattern of centre-surround enhancement and suppression is characteristic of low-level effects on cortical excitability.

However, the effects of load reported here cannot be explained by low-level effects, since stimuli were presented in separate modalities, indicating that there is some aspect to the effects of perceptual load that is not based on low-level effects. A lack of low-level interactions in cross-modal designs may explain why load induced by RSVP tasks appears to be effective at reducing sensitivity to visual distractors, but not to auditory distractors; if the RSVP load effect is partially based on reduced excitability of visual cortex to non-attended locations, it would not be expected to affect auditory processing in the same way as visual processing. If cross-modal effects of load are time sensitive, as suggested by the data here, then the stimulus asynchronies that exist within RSVP tasks between the load stimuli and critical auditory stimuli would mean that load would not affect processing of the auditory signals. This once again raises the question of whether prolonged auditory load tasks might provide a better control of resources to an unattended auditory signal; if part of the effect of load is a low-level suppression of signals that do not share a cortical locus with the task stimuli (i.e. for low levels of the auditory pathway, signals that do not overlap spectrally) using an auditory task to manipulate load might help ensure that resources are not available for processing irrelevant auditory signals. There are several plausible low-level effects in auditory processing that might recreate the suppression patterns seen in the visual studies described above. The sensitivity of auditory processing in frequency space is determined by the auditory bandwidths of the neurones in auditory cortex. Engaging selective attention to a frequency narrows the relevant bandwidths, giving a more precise representation in the selected frequency region, but reducing sensitivity outside it; similarly, there is a brief temporal window within which auditory signals in the same frequency band can mask one another even though they did not occur at precisely the same time (termed backwards and forwards masking), and engaging attention also improves the sensitivity of the temporal representation of signals (Fritz et al., 2003; Fritz et al., 2007). These effects of suppression of spectral and temporal regions around the focal point of attention is akin to the centre-surround suppression covering short temporal windows that has been described by Carmel et al. (2011) and Parks et al. (2013), suggesting that a similar effect to that seen in vision using RSVP tasks may occur for auditory processing. Indeed, one study which contrasted the effect of load in RSVP and rapid serial auditory presentation (RSAP) on the processing of irrelevant auditory signals found an effect of load in the auditory but not the visual task (Chait et al., 2012).

The automatic allocation of spare resources across modalities described in this thesis also raises the question of how the automatic process decides which irrelevant stimuli receive processing. Is capacity allocated evenly, resulting in equal but small levels of resources applied to processing all signals, or are there some biases built into the system? One possibility that seems likely is that unattended auditory signals might receive priority over unattended visual signals, in keeping with the idea that the auditory system performs a lot of passive processes in order to provide an early warning system. However, even within auditory processing there may be some processes that are always given prioritised access to leftover resources. For instance, change detection mechanisms might be particularly valuable in terms of the utility of information provided and the relatively low level of processing required to determine factors such as changes in low-level stimulus parameters. This seems plausible given that the MMN can be abolished in cases of extreme resource depletion, but is typically elicited by default even during situations such as anaesthesia or coma, when stimulus processing would be guided entirely based on any in-built prioritisation hierarchy. There is also the question of how stimulus characteristics, such as those which govern saliency, interact with any in-built default resource allocation hierarchy. These questions could provide the basis for a programme of research that could provide further valuable insight into the factors that govern the degree to which unattended auditory signals are processed in conditions of load on perceptual resources.

6.6 CONCLUSIONS

The data presented here provide insight into the degree to which auditory processes occur automatically, mechanisms of selective attention, and models of perceptual load. I hope that my work helps advance understanding in these fields and encourage the use of paradigms in which visual load is used as a tool to probe the effects of limited global capacity on auditory processing.

REFERENCES

Akram S, Englitz B, Elhilali M, Simon JZ, Shamma SA (2014) Investigating the neural correlates of a streaming percept in an informational-masking paradigm. PloS one 9:e114427.

Alain C, Woods DL (1997) Attention modulates auditory pattern memory as indexed by event-related brain potentials. Psychophysiology 34:534-546.

Alain C, Arnott SR (2000) Selectively attending to auditory objects. Frontiers in bioscience 5:D202-212.

Alain C, Izenberg A (2003) Effects of attentional load on auditory scene analysis. Journal of cognitive neuroscience 15:1063-1073.

Alain C, McDonald KL (2007) Age-related differences in neuromagnetic brain activity underlying concurrent sound perception. The Journal of Neuroscience 27:1308-1314.

Alain C, Woods DL, Ogawa KH (1994) Brain indices of automatic pattern processing. Neuroreport 6:140.

Alain C, Achim A, Woods DL (1999a) Separate memory-related processing for auditory frequency and patterns. Psychophysiology 36:737-744.

Alain C, Cortese F, Picton TW (1999b) Event-related brain activity associated with auditory pattern processing. Neuroreport 10:2429-2434.

Alain C, Arnott SR, Picton TW (2001) Bottom-up and top-down influences on auditory scene analysis: evidence from event-related brain potentials. Journal of experimental psychology: human perception and performance 27:1072-1089.

Alain C, Schuler BM, McDonald KL (2002) Neural activity associated with distinguishing concurrent auditory objects. The Journal of the Acoustical Society of America 111:990-995.

Alnæs D, Sneve MH, Espeseth T, Endestad T, van de Pavert SHP, Laeng B (2014) Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the locus coeruleus. Journal of Vision 14:1-1.

Alpern M, McCready DW, Barr L (1963) The dependence of the photopupil response on flash duration and intensity. The Journal of General Physiology 47:265-278.

Andreou L-V, Kashino M, Chait M (2011) The role of temporal regularity in auditory segregation. Hearing research 280:228-235.

Anstis SM, Saida S (1985) Adaptation to auditory streaming of frequency-modulated tones. Journal of experimental psychology: human perception and performance 11:257.

Aston-Jones G, Rajkowski J, Kubiak P, Alexinsky T (1994) Locus coeruleus neurons in monkey are selectively activated by attended cues in a vigilance task. The Journal of Neuroscience 14:4467-4480.

Beatty J (1982) Phasic not tonic pupillary responses vary with auditory vigilance performance. Psychophysiology 19:167-172.

Beatty J, Lucero-Wagoner B (2000) The pupillary system. Handbook of psychophysiology. Cambridge University Press.

Bell AH, Fecteau JH, Munoz DP (2004) Using auditory and visual stimuli to investigate the behavioral and neuronal consequences of reflexive covert orienting. Journal of neurophysiology 91:2172-2184.

Bendixen A (2014) Predictability effects in auditory scene analysis: a review. Frontiers in Neuroscience 8:60.

Bidet-Caulet A, Fischer C, Besle J, Aguera PE, Giard MH, Bertrand O (2007) Effects of selective attention on the electrophysiological representation of concurrent sounds in the human auditory cortex. The Journal of Neuroscience 27:9252-9261.

Billig AJ, Carlyon RP (2016) Automaticity and primacy of auditory streaming: Concurrent subjective and objective measures. Journal of experimental psychology: human perception and performance 42:339-353.

Bitsios P, Szabadi E, Bradshaw C (1998) The effects of clonidine on the fear-inhibited light reflex. Journal of Psychopharmacology 12:137-145.

Boersma F, Wilton K, Barham R, Muir W (1970) Effects of arithmetic problem difficulty on pupillary dilation in normals and educable retardates. Journal of Experimental Child Psychology 9:142-155.

Bradshaw J (1968) Pupil size and problem solving. Quarterly Journal of Experimental Psychology 20:116-122.

Broadbent D (1958) Perception and communication. Elmsford, NY, US.

Brown GG, Kindermann SS, Siegle GJ, Granholm E, Wong EC, Buxton RB (1999) Brain activation and pupil response during covert performance of the Stroop Color Word task. Journal of the International Neuropsychological Society 5:308-319.

Calvert GA (2001) Crossmodal processing in the human brain: insights from functional neuroimaging studies. Cerebral cortex 11:1110-1123.

Carlyon RP, Cusack R, Foxton JM, Robertson IH (2001) Effects of attention and unilateral neglect on auditory stream segregation. Journal of experimental psychology: human perception and performance 27:115-127.

Carmel D, Thorne JD, Rees G, Lavie N (2011) Perceptual load alters visual excitability. Journal of experimental psychology: human perception and performance 37:1350.

Cartwright-Finch U, Lavie N (2007) The role of perceptual load in inattentional blindness. Cognition 102:321-340.

Chait M, Ruff CC, Griffiths TD, McAlpine D (2012) Cortical responses to changes in acoustic regularity are differentially modulated by attentional load. Neuroimage 59:1932-1941.

Clifford J, Day M, Orwin J (1982) Reversal of clonidine induced miosis by the alpha 2-adrenoreceptor antagonist RX 781094. British journal of clinical pharmacology 14:99-101.

Comerchero MD, Polich J (1999) P3a and P3b from typical auditory and visual stimuli. Clinical Neurophysiology 110:24-30.

Cooke M (2006) A glimpsing model of speech perception in noise. The Journal of the Acoustical Society of America 119:1562-1573.

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

Costa VD, Rudebeck PH (2016) More than meets the eye: the relationship between pupil size and locus coeruleus activity. Neuron 89:8-10.

Coull J, Büchel C, Friston K, Frith C (1999) Noradrenergically mediated plasticity in a human attentional neuronal network. Neuroimage 10:705-715.

Craik FI, Lockhart R (1972) Levels of processing: A framework for memory research. Journal of verbal learning and verbal behavior 11:671-684.

Cusack R (2005) The intraparietal sulcus and perceptual organization. Journal of cognitive neuroscience 17:641-651.

Cusack R, Deeks J, Aikman G, Carlyon RP (2004) Effects of location, frequency region, and time course of selective attention on auditory scene analysis. Journal of experimental psychology Human perception and performance 30:643-656.

Dalton P, Lavie N (2004) Auditory attentional capture: effects of singleton distractor sounds. Journal of experimental psychology Human perception and performance 30:180.

Darwin C, Hukin R, Al-Khatib BY (1995) Grouping in pitch perception: Evidence for sequential constraints. The Journal of the Acoustical Society of America 98:880-885.

de Cheveigné A, Parra LC (2014) Joint decorrelation, a versatile tool for multichannel data analysis. NeuroImage 98:487-505.

de Gee JW, Knapen T, Donner TH (2014) Decision-related pupil dilation reflects upcoming choice and individual bias. Proceedings of the National Academy of Sciences, 111:E618-E625.

Dehaene S, Changeux J-P (2011) Experimental and theoretical approaches to conscious processing. Neuron 70:200-227.

Demany L, Semal C, Cazalets J-R, Pressnitzer D (2010) Fundamental differences in change detection between vision and audition. Experimental Brain Research 203:261-270.

Deutsch JA, Deutsch D (1963) Attention: Some theoretical considerations. Psychological review 70:80.

Ding N, Simon JZ (2012) Emergence of neural encoding of auditory objects while listening to competing speakers. Proceedings of the National Academy of Sciences 109:11854-11859.

Donner TH, Kettermann A, Diesch E, Ostendorf F, Villringer A, Brandt SA (2002) Visual feature and conjunction searches of equal difficulty engage only partially overlapping frontoparietal networks. NeuroImage 15:16-25.

Driver J (2001) A selective review of selective attention research from the past century. British Journal of Psychology 92:53-78.

Dyson BJ, Alain C (2004) Representation of concurrent acoustic objects in primary auditory cortex. The Journal of the Acoustical Society of America 115:280-288.

Dyson BJ, Alain C, He Y (2005) Effects of visual attentional load on low-level auditory scene analysis. Cognitive, affective & behavioral neuroscience 5:319-338.

Dyson BJ, Alain CJC (2008) It all sounds the same to me: Sequential ERP and behavioral effects during pitch and harmonicity judgments. Cognitive, affective & behavioral neuroscience 8:329-343.

Eayrs J, Lavie N (2018) Establishing Individual Differences in Perceptual Capacity. Journal of Experimental Psychology: Human Perception and Performance 44:1240-1257.

Efron B, Tibshirani RJ (1993) An Introduction to the Bootstrap: Monographs on Statistics and Applied Probability, Vol. 57. New York and London: Chapman and Hall/CRC.

Eldar E, Cohen JD, Niv Y (2013) The effects of neural gain on attention and learning. Nature neuroscience 16:1146.

Elhilali M, Xiang J, Shamma SA, Simon JZ (2009a) Interaction between attention and bottom-up saliency mediates the representation of foreground and background in an auditory scene. PLoS biology 7:e1000129.

Elhilali M, Ma L, Micheyl C, Oxenham AJ, Shamma SA (2009b) Temporal coherence in the perceptual organization and cortical representation of auditory scenes. Neuron 61:317-329.

Ellis C (1981) The pupillary light reflex in normal subjects. British Journal of Ophthalmology 65:754-759.

Engelhardt PE, Ferreira F, Patsenko EG (2010) Pupillometry reveals processing load during spoken language comprehension. Quarterly Journal of Experimental Psychology 63:639-645.

Eramudugolla R, Irvine DR, McAnally KI, Martin RL, Mattingley JB (2005) Directed attention eliminates ‘change deafness’ in complex auditory scenes. Current Biology 15:1108-1113.

Escera C, Corral M (2007) Role of mismatch negativity and novelty-P3 in involuntary auditory attention. Journal of Psychophysiology 21:251-264.

Escera C, Alho K, Winkler I, Näätänen R (1998) Neural mechanisms of involuntary attention to acoustic novelty and change. Journal of cognitive neuroscience 10:590-604.

Fanciullacci M, Pietrini U, Fusco B, Alessandri M, Marabini S, Sicuteri F (1988) Does anisocoria by clonidine reflect a central sympathetic dysfunction in cluster headache? Clinical neuropharmacology 11:56-62.

Foote SL, Morrison JH (1987) Extrathalamic modulation of cortical function. Annual review of neuroscience 10:67-95.

Forster S, Lavie N (2008) Failures to ignore entirely irrelevant distractors: The role of load. J Exp Psychol-Appl 14:73-83.

Fritz J, Shamma S, Elhilali M, Klein D (2003) Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. Nature neuroscience 6:1216.

Fritz JB, Elhilali M, David SV, Shamma SA (2007) Does attention play a role in dynamic receptive field adaptation to changing acoustic salience in A1? Hearing research 229:186-203.

Fu K-MG, Foxe JJ, Murray MM, Higgins BA, Javitt DC, Schroeder CE (2001) Attention-dependent suppression of distracter visual input can be cross-modally cued as indexed by anticipatory parieto–occipital alpha-band oscillations. Cognitive Brain Research 12:145-152.

Gilzenrat MS, Nieuwenhuis S, Jepma M, Cohen JD (2010) Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. Cognitive, affective & behavioral neuroscience 10:252-269.

Granholm E, Asarnow RF, Sarkin AJ, Dykes KL (1996) Pupillary responses index cognitive resource limitations. Psychophysiology 33:457-461.

Grill-Spector K (2003) The neural basis of object perception. Current opinion in neurobiology 13:159-166.

Gutschalk A, Micheyl C, Oxenham AJ (2008) Neural correlates of auditory perceptual awareness under informational masking. PLoS biology 6:e138.

Gutschalk A, Oxenham AJ, Micheyl C, Wilson EC, Melcher JR (2007) Human cortical activity during streaming without spectral cues suggests a general neural substrate for auditory stream segregation. The Journal of neuroscience : the official journal of the Society for Neuroscience 27:13074-13081.

Gutschalk A, Micheyl C, Melcher JR, Rupp A, Scherg M, Oxenham AJ (2005) Neuromagnetic correlates of streaming in human auditory cortex. Journal of Neuroscience 25:5382-5388.

Hakerem G, Sutton S (1966) Pupillary response at visual threshold. Nature 212 (5061):485.

Handy TC, Soltani M, Mangun GR (2001) Perceptual load and visuocortical processing: Event-related potentials reveal sensory-level selection. Psychological Science 12:213-218.

Hansen JC, Hillyard SA (1983) Selective attention to multidimensional auditory stimuli. Journal of experimental psychology: human perception and performance 9:1.

Hari R (1990) The neuromagnetic method in the study of the human auditory cortex. Advances in audiology 6:222-282.

Hashimoto T, Kashii S, Kikuchi M, Honda Y, Nagamine T, Shibasaki H (1999) Temporal profile of visual evoked responses to pattern-reversal stimulation analyzed with a whole-head magnetometer. Experimental brain research 125:375-382.

Heaver B, Hutton SB (2011) Keeping an eye on the truth? Pupil size changes associated with recognition memory. Memory 19:398-405.

Hess EH, Polt JM (1964) Pupil size in relation to mental activity during simple problem-solving. Science 143:1190-1192.

Hillyard SA, Hink RF, Schwent VL, Picton TW (1973) Electrical signs of selective attention in the human brain. Science 182:177-180.

Hoormann J, Falkenstein M, Hohnsbein J (2000) Early attention effects in human auditory-evoked potentials. Psychophysiology 37:29-42.

Hopf JM, Luck SJ, Girelli M, Hagner T, Mangun GR, Scheich H, Heinze HJ (2000) Neural sources of focused attention in visual search. Cerebral Cortex 10:1233-1241.

Humphreys GW (2001) A multi-stage account of binding in vision: Neuropsychological evidence. Visual Cognition 8:381-410.

Hyönä J, Tommola J, Alaja A-M (1995) Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. Quarterly Journal of Experimental Psychology 48:598-612.

Isreal J, Chesney G, Wickens C, Donchin E (1980) P300 and tracking difficulty: Evidence for a multiple capacity view of attention. Psychophysiology 17:259-273.

Jenkins R, Lavie N, Driver J (2005) Recognition memory for distractor faces depends on attentional load at exposure. Psychonomic Bulletin and Review 12:314-320.

Jepma M, Nieuwenhuis S (2011) Pupil diameter predicts changes in the exploration–exploitation trade-off: Evidence for the adaptive gain theory. Journal of cognitive neuroscience 23:1587-1596.

Jiang Y, Chun MM (2001) The influence of temporal selection on spatial selection and distractor interference: An attentional blink study. Journal of experimental psychology: human perception and performance 27:664.

Joshi S, Li Y, Kalwani RM, Gold JI (2016) Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. Neuron 89:221-234.

Jovicich J, Peters RJ, Koch C, Braun J, Chang L, Ernst T (2001) Brain areas specific for attentional load in a motion-tracking task. Journal of cognitive neuroscience 13:1048-1058.

Just MA, Carpenter PA (1993) The intensity dimension of thought: pupillometric indices of sentence processing. Canadian Journal of Experimental Psychology 47:310.

Kafkas A, Montaldi D (2011) Recognition memory strength is predicted by pupillary responses at encoding while fixation patterns distinguish recollection from familiarity. Quarterly Journal of Experimental Psychology 64:1971-1989.

Kahneman D, Beatty J (1966) Pupil diameter and load on memory. Science 154:1583-1585.

Kahneman D, Beatty J, Pollack I (1967) Perceptual deficit during a mental task. Science 157:218-219.

Kelly SP, Gomez-Ramirez M, Foxe JJ (2008) Spatial attention modulates initial afferent activity in human primary visual cortex. Cerebral Cortex 18:2629-2636.

Kelly SP, Lalor EC, Reilly RB, Foxe JJ (2006) Increases in alpha oscillatory power reflect an active retinotopic mechanism for distracter suppression during sustained visuospatial attention. Journal of neurophysiology 95:3844-3851.

Klingner J, Tversky B, Hanrahan P (2011) Effects of visual and verbal presentation on cognitive load in vigilance, memory, and arithmetic tasks. Psychophysiology 48:323-332.

Kluge C, Bauer M, Leff AP, Heinze H-J, Dolan RJ, Driver J (2011) Plasticity of human auditory-evoked fields induced by shock conditioning and contingency reversal. Proceedings of the National Academy of Sciences 108:12545-12550.

Koivisto M, Kainulainen P, Revonsuo A (2009) The relationship between awareness and attention: evidence from ERP responses. Neuropsychologia 47:2891-2899.

Kok A (2001) On the utility of P3 amplitude as a measure of processing capacity. Psychophysiology 38:557-577.

Kramer A, Schneider W, Fisk A, Donchin E (1986) The effects of practice and task structure on components of the event-related brain potential. Psychophysiology 23:33-47.

Krishnan L, Elhilali M, Shamma S (2014) Segregating complex sound sources through temporal coherence. PLoS computational biology 10:e1003985.

Laeng B, Ørbo M, Holmlund T, Miozzo M (2011) Pupillary stroop effects. Cognitive processing 12:13-21.

Lakatos P, Barczak A, Neymotin SA, McGinnis T, Ross D, Javitt DC, O'Connell MN (2016) Global dynamics of selective attention and its lapses in primary auditory cortex. Nature neuroscience 19:1707.

Lavie N (1995) Perceptual load as a necessary condition for selective attention. Journal of experimental psychology Human perception and performance 21:451-468.

Lavie N (2005) Distracted and confused?: selective attention under load. Trends in cognitive sciences 9:75-82.

Lavie N (2010) Attention, Distraction, and Cognitive Control Under Load. Curr Dir Psychol Sci 19:143-148.

Lavie N, Cox S (1997) On the efficiency of visual selective attention: Efficient visual search leads to inefficient distractor rejection. Psychological Science 8:395-396.

Lavie N, Fox E (2000) The role of perceptual load in negative priming. Journal of experimental psychology Human perception and performance 26:1038.

Lavie N, De Fockert JW (2003) Contrasting effects of sensory limits and capacity limits in visual selective attention. Perception and Psychophysics 65:202-212.

Lavie N, De Fockert J (2005) The role of working memory in attentional capture. Psychonomic Bulletin and Review 12:669-674.

Lavie N, Torralbo A (2010) Dilution: A theoretical burden or just load? A reply to Tsal and Benoni (2010). Journal of experimental psychology Human perception and performance 36(6):1657-1664.

Lavie N, Hirst A, De Fockert JW, Viding E (2004) Load theory of selective attention and cognitive control. Journal of Experimental Psychology: General 133:339.

Lavie N, Lin Z, Zokaei N, Thoma V (2009) The role of perceptual load in object recognition. Journal of experimental psychology Human perception and performance 35:1346.

Levitt H (1971) Transformed up-down methods in psychoacoustics. The Journal of the Acoustical Society of America 49:467-477.

Lipp R, Kitterick P, Summerfield Q, Bailey PJ, Paul-Jordanov I (2010) Concurrent sound segregation based on inharmonicity and onset asynchrony. Neuropsychologia 48:1417-1425.

Litvak V, Friston K (2008) Electromagnetic source reconstruction for group studies. NeuroImage 42:1490-1498.

Loewenfeld I, Lowenstein O (1993) The pupil: Anatomy, physiology, and clinical applications. Detroit: Wayne State University Press.

Lu K, Xu YB, Yin PB, Oxenham AJ, Fritz JB, Shamma SA (2017) Temporal coherence structure rapidly shapes neuronal interactions. Nature Communications 8:13900.

Lütkenhöner B, Steinsträter O (1998) High-precision neuromagnetic study of the functional organization of the human auditory cortex. Audiology and Neurotology 3:191-213.

Macdonald JS, Lavie N (2008) Load induced blindness. Journal of experimental psychology: human perception and performance 34:1078.

Macdonald JS, Lavie N (2011) Visual perceptual load induces inattentional deafness. Attention, Perception & Psychophysics 73:1780-1789.

Macken WJ, Tremblay S, Houghton RJ, Nicholls AP, Jones DM (2003) Does auditory streaming require attention? Evidence from attentional selectivity in short-term memory. Journal of experimental psychology Human perception and performance 29:43-51.

Marshall SP (2002) The index of cognitive activity: Measuring cognitive workload. In: Proceedings of the IEEE 7th conference on Human Factors and Power Plants, pp 7-7: IEEE.

Martins ARO, Froemke RC (2015) Coordinated forms of noradrenergic plasticity in the locus coeruleus and primary auditory cortex. Nature neuroscience 18:1483.

Masutomi K, Barascud N, Kashino M, McDermott JH, Chait M (2016) Sound segregation via embedded repetition is robust to inattention. J Exp Psychol Human 42:386.

Mathôt S, Fabius J, Van Heusden E, Van der Stigchel S (2018) Safe and sensible preprocessing and baseline correction of pupil-size data. Behavior research methods 50:94-106.

McCarthy G, Luby M, Gore J, Goldman-Rakic P (1997) Infrequent events transiently activate human prefrontal and parietal cortex as measured by functional MRI. Journal of Neurophysiology 77:1630-1634.

McDermott JH, Wrobleski D, Oxenham AJ (2011) Recovering sound sources from embedded repetition. Proceedings of the National Academy of Sciences 108:1188-1193.

McDonald KL, Alain C (2005) Contribution of harmonicity and location to auditory object formation in free field: evidence from event-related brain potentials. Journal of the Acoustical Society of America 118:1593-1604.

Mecklinger A, Maess B, Opitz B, Pfeifer E, Cheyne D, Weinberg H (1998) A MEG analysis of the P300 in visual discrimination tasks. Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section 108:45-56.

Micheyl C, Oxenham AJ (2010) Pitch, harmonicity and concurrent sound segregation: psychoacoustical and neurophysiological findings. Hearing research 266:36-51.

Molloy K, Lavie N, Chait M (2019) Auditory figure-ground segregation is impaired by high visual load. Journal of Neuroscience 39:1699-1708.

Molloy K, Griffiths TD, Chait M, Lavie N (2015) Inattentional Deafness: Visual Load Leads to Time-Specific Suppression of Auditory Evoked Responses. The Journal of neuroscience : the official journal of the Society for Neuroscience 35:16046-16054.

Moore BC, Glasberg BR, Peters RW (1986) Thresholds for hearing mistuned partials as separate tones in harmonic complexes. The Journal of the Acoustical Society of America 80:479-483.

Morley M, Bradshaw C, Szabadi E (1991) Effects of clonidine and yohimbine on the pupillary light reflex and carbachol-evoked sweating in healthy volunteers. British journal of clinical pharmacology 31:99-101.

Muller-Gass A, Stelmack RM, Campbell KB (2006) The effect of visual task difficulty and attentional direction on the detection of acoustic change as indexed by the mismatch negativity. Brain research 1078:112-130.

Muller-Gass A, Macdonald M, Schröger E, Sculthorpe L, Campbell K (2007) Evidence for the auditory P3a reflecting an automatic process: elicitation during highly-focused continuous visual attention. Brain research 1170:71-78.

Müller BW, Achenbach C, Oades RD, Bender S, Schall U (2002) Modulation of mismatch negativity by stimulus deviance and modality of attention. Neuroreport 13:1317-1320.

Müller D, Widmann A, Schröger E (2005) Auditory streaming affects the processing of successive deviant and standard sounds. Psychophysiology 42:668-676.

Müller HJ, Elliott MA, Herrmann CS, Mecklinger A (2001) Neural binding of space and time: An introduction. Visual Cognition 8:273-285.

Muller MM, Andersen S, Trujillo NJ, Valdes-Sosa P, Malinowski P, Hillyard SA (2006) Feature-selective attention enhances color signals in early visual areas of the human brain. Proceedings of the National Academy of Sciences 103:14250-14254.

Murphy PR, O'connell RG, O'sullivan M, Robertson IH, Balsters JH (2014) Pupil diameter covaries with BOLD activity in human locus coeruleus. Human brain mapping 35:4140-4154.

Murphy S, Spence C, Dalton P (2017) Auditory perceptual load: A review. Hearing research 352:40-48.

Näätänen R, Michie PT (1979) Early selective-attention effects on the evoked potential: a critical review and reinterpretation. Biological psychology 8:81-136.

Näätänen R, Paavilainen P, Rinne T, Alho K (2007) The mismatch negativity (MMN) in basic research of central auditory processing: a review. Clinical neurophysiology 118:2544-2590.

Näätänen R, Astikainen P, Ruusuvirta T, Huotilainen M (2010) Automatic auditory intelligence: An expression of the sensory–cognitive core of cognitive processes. Brain research reviews 64:123-136.

Nager W, Teder-Sälejärvi W, Kunze S, Münte TF (2003) Preattentive evaluation of multiple perceptual streams in human audition. Neuroreport 14:871-874.

Nakamura A, Kakigi R, Hoshiyama M, Koyama S, Kitamura Y, Shimojo M (1997) Visual evoked cortical magnetic fields to pattern reversal stimulation. Cognitive Brain Research 6:9-22.

Neisser U, Becklen R (1975) Selective looking: Attending to visually specified events. Cognitive psychology 7:480-494.

Norman DA, Bobrow DG (1975) On data-limited and resource-limited processes. Cognitive psychology 7:44-64.

O'Connor DH, Fukui MM, Pinsk MA, Kastner S (2002) Attention modulates responses in the human lateral geniculate nucleus. Nature neuroscience 5:1203-1209.

O'Sullivan JA, Shamma SA, Lalor EC (2015) Evidence for Neural Computations of Temporal Coherence in an Auditory Scene and Their Enhancement during Active Listening. Journal of neuroscience 35:7256-7263.

O’Neill W, Zimmerman S (2000) Neurologial interpretations and the information in the cognitive pupillary response. Methods of information in medicine 39:122-124.

Opitz B, Mecklinger A, Friederici AD, von Cramon DY (1999) The functional neuroanatomy of novelty processing: integrating ERP and fMRI results. Cerebral cortex 9:379-391.

Otero SC, Weekes BS, Hutton SB (2011) Pupil size changes during recognition memory. Psychophysiology 48:1346-1353.

Otten LJ, Alain C, Picton TW (2000) Effects of visual attentional load on auditory processing. Neuroreport 11:875-880.

Paavilainen P, Simola J, Jaramillo M, Näätänen R, Winkler I (2001) Preattentive extraction of abstract feature conjunctions from auditory stimulation as reflected by the mismatch negativity (MMN). Psychophysiology 38:359-365.

Palmer L, Bialkowski A, Brostow GJ, Ambeck-Madsen J, Lavie N (2017) Predicting the Perceptual Demands of Urban Driving with Video Regression. In: 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), pp 409-417: IEEE.

Pantev C, Eulitz C, Hampson S, Ross B, Roberts L (1996) The auditory evoked" off" response: sources and comparison with the" on" and the" sustained" responses. Ear and hearing 17:255-265.

Parasuraman R (1978) Auditory evoked potentials and divided attention. Psychophysiology 15:460-465.

Parks NA, Hilimire MR, Corballis PM (2009) Visual perceptual load modulates an auditory microreflex. Psychophysiology 46:498-501.

Parks NA, Hilimire MR, Corballis PM (2011) Steady-state signatures of visual perceptual load, multimodal distractor filtering, and neural competition. Journal of cognitive neuroscience 23:1113-1124.

Parks NA, Beck DM, Kramer AF (2013) Enhancement and suppression in the visual field under perceptual load. Frontiers in psychology 4:275.

Payne DT, Parry ME, Harasymiw SJ (1968) Percentage of pupillary dilation as a measure of item difficulty. Perception and Psychophysics 4:139-143.

Peavler WS (1974) Pupil size, information overload, and performance differences. Psychophysiology 11:559-566.

Phillips M, Szabadi E, Bradshaw C (2000) Comparison of the effects of clonidine and yohimbine on spontaneous pupillary fluctuations in healthy human volunteers. Psychopharmacology 150:85-89.

Pinsk MA, Doniger GM, Kastner S (2004) Push-pull mechanism of selective attention in human extrastriate cortex. Journal of Neurophysiology 92:622-629.

Plack CJ, Moore BC (1990) Temporal window shape as a function of frequency and level. The Journal of the Acoustical Society of America 87:2178-2187.

Polich J, Comerchero MD (2003) P3a from visual stimuli: typicality, task, and topography. Brain topography 15:141-152.

Porter G, Troscianko T, Gilchrist ID (2007) Effort during visual search and counting: Insights from pupillometry. Quarterly Journal of Experimental Psychology 60:211-229.

Privitera CM, Renninger LW, Carney T, Klein S, Aguilar M (2010) Pupil dilation during visual target detection. Journal of Vision 10:3-3.

Puvvada KC, Simon JZ (2017) Cortical Representations of Speech in a Multitalker Auditory Scene. Journal of neuroscience 37:9189-9196.

Rajkowski J (1993) Correlations between locus coeruleus (LC) neural activity, pupil diameter and behavior in monkey support a role of LC in attention. In: Society for Neuroscience. Washington, DC, USA.

Rauss KS, Pourtois G, Vuilleumier P, Schwartz S (2009) Attentional load modifies early activity in human primary visual cortex. Human brain mapping 30:1723-1733.

Raveh D, Lavie N (2015) Load-induced inattentional deafness. Attention, Perception & Psychophysics 77:483-492.

Rees G, Frith CD, Lavie N (1997) Modulating irrelevant motion perception by varying attentional load in an unrelated task. Science 278:1616-1619.

Rees G, Frith C, Lavie N (2001) Processing of irrelevant visual motion during performance of an auditory attention task. Neuropsychologia 39:937-949.

Reimer J, McGinley MJ, Liu Y, Rodenkirch C, Wang Q, McCormick DA, Tolias AS (2016) Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. Nature Communications 7:13289.

Restuccia D, Della Marca G, Marra C, Rubino M, Valeriani M (2005) Attentional load of the primary task influences the frontal but not the temporal generators of mismatch negativity. Cognitive Brain Research 25:891-899.

Roberts TP, Ferrari P, Stufflebeam SM, Poeppel D (2000) Latency of the auditory evoked neuromagnetic field components: stimulus dependence and insights toward perception. Journal of Clinical Neurophysiology 17:114-129.

Saarinen J, Paavilainen P, Schöger E, Tervaniemi M, Näätänen R (1992) Representation of abstract attributes of auditory stimuli in the human brain. NeuroReport 3:1149-1151.

Sabri M, Humphries C, Binder JR, Liebenthal E (2013) Neural events leading to and associated with detection of sounds under high processing load. Human brain mapping 34:587-597.

Sabri M, Liebenthal E, Waldron E, Medler DA, Binder JR (2006) Attentional modulation in the detection of irrelevant deviance: a simultaneous ERP/fMRI study. Journal of Cognitive Neuroscience 18:689-700.

Schadwinkel S, Gutschalk A (2010) Activity associated with stream segregation in human auditory cortex is similar for spatial and pitch cues. Cerebral Cortex 20:2863-2873.

Scherg M, Vajsar J, Picton TW (1989) A source analysis of the late human auditory evoked potentials. Journal of Cognitive Neuroscience 1:336-355.

Schneider KA, Kastner S (2009) Effects of sustained spatial attention in the human lateral geniculate nucleus and superior colliculus. Journal of Neuroscience 29:1784-1795.

Schröger E, Wolff C (1998) Attentional orienting and reorienting is indicated by human event-related brain potentials. Neuroreport 9:3355-3358.

Schulz JF, Fischbacher U, Thöni C, Utikal V (2014) Affect and fairness: Dictator games under cognitive load. Journal of Economic Psychology 41:77-87.

Schwartz S, Vuilleumier P, Hutton C, Maravita A, Dolan RJ, Driver J (2005) Attentional load and sensory competition in human vision: modulation of fMRI responses by load at fixation during task-irrelevant stimulation in the peripheral visual field. Cerebral cortex 15:770-786.

Sculthorpe LD, Collin CA, Campbell KB (2008) The influence of strongly focused visual attention on the detection of change in an auditory pattern. Brain research 1234:78-86.

Seeber KG (2013) Cognitive load in simultaneous interpreting: Measures and methods. Target International Journal of Translation Studies 25:18-32.

Shafritz KM, Gore JC, Marois R (2002) The role of the parietal cortex in visual feature binding. Proceedings of the National Academy of Sciences 99:10917-10922.

Shamma SA, Micheyl C (2010) Behind the scenes of auditory perception. Current opinion in neurobiology 20:361-366.

Shamma SA, Elhilali M, Micheyl C (2011) Temporal coherence and attention in auditory scene analysis. Trends in neurosciences 34:114-123.

Shigeto H, Tobimatsu S, Yamamoto T, Kobayashi T, Kato M (1998) Visual evoked cortical magnetic responses to checkerboard pattern reversal stimulation: a study on the neural generators of N75, P100 and N145. Journal of the neurological sciences 156:186-194.

Siegle GJ, Steinhauer SR, Thase ME (2004) Pupillary assessment and computational modeling of the Stroop task in depression. International Journal of Psychophysiology 52:63-76.

Siegle GJ, Ichikawa N, Steinhauer S (2008) Blink before and after you think: Blinks occur prior to and following cognitive load indexed by pupillary responses. Psychophysiology 45:679-687.

Simons DJ, Chabris CF (1999) Gorillas in our midst: Sustained inattentional blindness for dynamic events. Perception 28:1059-1074.

Snyder JS, Alain C, Picton TW (2006) Effects of attention on neuroelectric correlates of auditory stream segregation. Journal of cognitive neuroscience 18:1-13.

Snyder JS, Gregg MK, Weintraub DM, Alain C (2012) Attention, awareness, and the perception of auditory scenes. Frontiers in psychology 3:15.

Sterpenich V, D'Argembeau A, Desseilles M, Balteau E, Albouy G, Vandewalle G, Degueldre C, Luxen A, Collette F, Maquet P (2006) The locus ceruleus is involved in the successful retrieval of emotional memories in humans. Journal of Neuroscience 26:7416-7423.

Sussman E, Ritter W, Vaughan Jr HG (1998a) Attention affects the organization of auditory input associated with the mismatch negativity system. Brain research 789:130-138.

Sussman E, Ritter W, Vaughan Jr HG (1998b) Predictability of stimulus deviance and the mismatch negativity. Neuroreport 9:4167-4170.

Sussman E, Ritter W, Vaughan HG (1999) An investigation of the auditory streaming effect using event-related brain potentials. Psychophysiology 36:22-34.

Sussman E, Winkler I, Wang W (2003) MMN and attention: competition for deviance detection. Psychophysiology 40:430-435.

Sussman E, Ceponiene R, Shestakova A, Näätänen R, Winkler I (2001) Auditory stream segregation processes operate similarly in school-aged children and adults. Hearing research 153:108-114.

Sussman E, Winkler I, Huotilainen M, Ritter W, Näätänen R (2002) Top-down effects can modify the initially stimulus-driven auditory organization. Cognitive Brain Research 13:393-405.

Sussman ES (2007) A new view on the MMN and attention debate: the role of context in processing auditory events. Journal of Psychophysiology 21:164-175.

Sussman ES, Horvath J, Winkler I, Orr M (2007) The role of attention in the formation of auditory streams. Perception and Psychophysics 69:136-152.

Sussman ES, Chen S, Sussman-Fort J, Dinces E (2014) The five myths of MMN: redefining how to use MMN in basic and clinical research. Brain topography 27:553-564.

Teki S, Chait M, Kumar S, von Kriegstein K, Griffiths TD (2011) Brain bases for auditory stimulus-driven figure-ground segregation. Journal of neuroscience 31:164-171.

Teki S, Chait M, Kumar S, Shamma S, Griffiths TD (2013) Segregation of complex acoustic scenes based on temporal coherence. eLife 2:e00699.

Teki S, Barascud N, Picard S, Payne C, Griffiths TD, Chait M (2016) Neural Correlates of Auditory Figure-Ground Segregation Based on Temporal Coherence. Cerebral Cortex 26:3669-3680.

Thompson SK, Carlyon RP, Cusack R (2011) An objective measurement of the build-up of auditory streaming and of its modulation by attention. Journal of experimental psychology Human perception and performance 37:1253-1262.

Tipper SP (1985) The negative priming effect: Inhibitory priming by ignored objects. Quarterly Journal of Experimental Psychology 37:571-590.

Todd JJ, Fougnie D, Marois R (2005) Visual short-term memory load suppresses temporo-parietal junction activity and induces inattentional blindness. Psychological science 16:965-972.

Toth B, Kocsis Z, Haden GP, Szerafin A, Shinn-Cunningham BG, Winkler I (2016) EEG signatures accompanying auditory figure-ground segregation. NeuroImage 141:108-119.

Treisman AM, Gelade G (1980) A feature-integration theory of attention. Cognitive psychology 12:97-136.

Trejo LJ, Ryan-Jones DL, Kramer AF (1995) Attentional modulation of the mismatch negativity elicited by frequency differences between binaurally presented tone bursts. Psychophysiology 32:319-328.

Tsal Y, Benoni H (2010) Diluting the burden of load: perceptual load effects are simply dilution effects. Journal of experimental psychology: human perception and performance 36:1645.

Valentino RJ, Van Bockstaele E (2008) Convergent regulation of locus coeruleus activity as an adaptive response to stress. European journal of pharmacology 583:194-203.

Valentino RJ, Foote SL, Page ME (1993) The Locus Coeruleus as a Site for Integrating Corticotropin-Releasing Factor and Noradrenergic Mediation of Stress Responsesa. Annals of the New York Academy of Sciences 697:173-188.

van Bochove ME, Van der Haegen L, Notebaert W, Verguts T (2013) Blinking predicts enhanced cognitive control. Cognitive, Affective and Behavioral Neuroscience 13:346-354.

van Noorden LPAS (1975) Temporal coherence in the perception of tone sequences: Institute for Perceptual Research Eindhoven, The Netherlands.

Van Steenbergen H, Band GP (2013) Pupil dilation in the Simon task as a marker of conflict processing. Frontiers in human neuroscience 7:215.

Võ MLH, Jacobs AM, Kuchinke L, Hofmann M, Conrad M, Schacht A, Hutzler F (2008) The coupling of emotion and cognition in the eye: Introducing the pupil old/new effect. Psychophysiology 45:130-140.

Wang C-A, Munoz DP (2015) A circuit for pupil orienting responses: implications for cognitive modulation of pupil size. Current opinion in neurobiology 33:134-140.

Wang C-A, Boehnke SE, Itti L, Munoz DP (2014) Transient pupil response is modulated by contrast-based saliency. Journal of Neuroscience 34:408-417.

Weintraub DM, Metzger BA, Snyder JS (2014) Effects of attention to and awareness of preceding context tones on auditory streaming. Journal of experimental psychology Human perception and performance 40:685-701.

Wickens C, Kramer A, Vanasse L, Donchin E (1983) Performance of concurrent tasks: a psychophysiological analysis of the reciprocity of information-processing resources. Science 221:1080-1082.

Wierda SM, van Rijn H, Taatgen NA, Martens S (2012) Pupil dilation deconvolution reveals the dynamics of attention at high temporal resolution. Proceedings of the National Academy of Sciences 109:8456-8460.

Wilson EC, Melcher JR, Micheyl C, Gutschalk A, Oxenham AJ (2007) Cortical FMRI activation to sequences of tones alternating in frequency: relationship to perceived rate and streaming. Journal of neurophysiology 97:2230-2238.

Wolber M, Wascher E (2003) Visual search strategies are indexed by event-related lateralizations of the EEG. Biological psychology 63:79-100.

Woldorff M, Hansen JC, Hillyard SA (1987) Evidence for effects of selective attention in the mid-latency range of the human auditory event-related potential. Electroencephalography and clinical neurophysiology Supplement 40:146-154.

Woldorff MG, Hackley SA, Hillyard SA (1991) The effects of channel-selective attention on the mismatch negativity wave elicited by deviant tones. Psychophysiology 28:30-42.

Woldorff MG, Hillyard SA, Gallen CC, Hampson SR, Bloom FE (1998) Magnetoencephalographic recordings demonstrate attentional modulation of mismatch-related neural activity in human auditory cortex. Psychophysiology 35:283-292.

Woldorff MG, Gallen CC, Hampson SA, Hillyard SA, Pantev C, Sobel D, Bloom FE (1993) Modulation of early sensory processing in human auditory cortex during auditory selective attention. Proceedings of the National Academy of Sciences 90:8722-8726.

Woods KJ, McDermott JH (2015) Attentive tracking of sound sources. Current Biology 25:2238-2246.

Wright P, Kahneman D (1971) Evidence for alternative strategies of sentence retention. Quarterly Journal of Experimental Psychology 23:197-213.

Yellin D, Berkovich-Ohana A, Malach R (2015) Coupling between pupil fluctuations and resting-state fMRI uncovers a slow build-up of antagonistic responses in the human cortex. NeuroImage 106:414-427.

Yi D-J, Woodman GF, Widders D, Marois R, Chun MM (2004) Neural fate of ignored stimuli: dissociable effects of perceptual and working memory load. Nature neuroscience 7:992-996.

Zhang Y, Chen Y, Bressler SL, Ding M (2008) Response preparation and inhibition: the role of the cortical sensorimotor beta rhythm. Neuroscience 156:238-246.

Zion Golumbic EM, Ding N, Bickel S, Lakatos P, Schevon CA, McKhann GM, Goodman RR, Emerson R, Mehta AD, Simon JZ, Poeppel D, Schroeder CE (2013) Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party". Neuron 77:980-991.

Zylberberg A, Oliva M, Sigman M (2012) Pupil dilation: a fingerprint of temporal selection during the “attentional blink”. Frontiers in Psychology 3:316.