



A University of Sussex PhD thesis

Available online via Sussex Research Online:

<http://sro.sussex.ac.uk/>

This thesis is protected by copyright which belongs to the author.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Please visit Sussex Research Online for more information and further details

Multisensory Interactions and Impacts on Attention and Awareness

Jessica Lunn

Thesis submitted for the degree of Doctor of Philosophy

School of Psychology

University of Sussex

September 2020

Declaration

I hereby declare that this thesis has not been and will not be, submitted in whole or in part to another University for the award of any other degree.

Signature:.....

Jessica Lunn

Acknowledgements

First and foremost, I would like to thank my supervisor, Dr Sophie Forster. Thank you for your endless guidance and reassurance, for inspiring me, and for dedicating so much time to helping me develop as a researcher, I could not have asked for a better supervisor. I am so thankful for the time I have spent in the Sussex Attention Lab, you make it a wonderful one to be a member of. Thank you also to my second supervisor, Professor Jamie Ward, for all your advice, positivity, and encouragement, I am so very grateful for your support.

Thank you to Professor Salvador Soto-Faraco and Dr Nick Berggren, for the time you have generously taken to provide advice and give me feedback, I have learnt a lot from you both. I would also like to thank Dr Graham Hole, for encouraging me to start this process in the first place.

Thank you to all the lab members, office mates, and other colleagues I have had the pleasure of spending time with over the last four years. Particularly to my fellow 'Triple J's' Jenny and Jo, and my PhD sisters Paloma and Rachel, for sharing this experience with me and offering encouragement when I needed it most. As well as to Chris, with additional special thanks for the guidance you have given on fMRI. My PhD experience would not have been nearly as fun without you all.

Thank you to my TGS girls Roxane and Sarah, and to Ruth, Jodie, Chris and Peter, you are all truly wonderful friends and have kept me (somewhat) sane over my many years in education. To my amazing parents, thank you for supporting me throughout every endeavour, with endless love and care. Thank you for encouraging me to keep, and to love, learning, and for believing in me even when I haven't believed in myself. To Hannah, thank you for always being there for me as a friend as well as a

sister. Your incredible work ethic combined with always being ready to have fun is inspiring. And last, but by no means least, to Devin. Thank you for your love, your patience, and your support, in this process as in everything. My life would not be the same without you.

Contents

Declaration	2
Acknowledgements.....	3
Summary	8
Preface.....	10
Chapter 1: General Introduction	11
Multisensory interactions & integration	12
Methods of measuring multisensory integration.....	13
Multisensory attention	18
Methods of measuring attentional capture by, and awareness of, multisensory stimuli	21
Perceptual Load Theory and multisensory attention.....	25
Mind-wandering and mental imagery	29
Thesis aims and overview.....	33
Chapter 2: Multisensory enhancement of attention depends on whether you are already paying attention	36
Chapter Summary.....	36
Abstract.....	37
Introduction.....	38
Experiment 1	46
<i>Materials and Methods</i>	46
<i>Results and Discussion</i>	50
Experiment 2.....	56
<i>Materials and Methods</i>	56
<i>Results and Discussion</i>	59
Experiment 3.....	65
<i>Materials and Methods</i>	65
<i>Results and Discussion</i>	67
Experiment 4.....	71
<i>Materials and Methods</i>	72
<i>Results and Discussion</i>	73
General Discussion.....	77
Chapter 3: Irrelevant sights and sounds require spatial suppression: ERP evidence	81
Chapter Summary.....	81

Abstract.....	82
Introduction.....	83
Methods	88
Results	94
General Discussion.....	100
Supplementary Materials.....	104
Chapter 4: Load-induced inattentional deafblindness	110
Chapter Summary.....	110
Abstract.....	111
Introduction.....	112
Materials and Methods	116
Results	120
Discussion.....	124
Supplementary Materials.....	129
Chapter 5: Establishing the effects of multisensory perceptual load.....	132
Chapter Summary.....	132
Abstract.....	133
Introduction.....	134
Behavioural Experiment 1	139
<i>Methods</i>	139
<i>Results</i>	143
<i>Discussion</i>	148
Behavioural Experiment 2	150
<i>Methods</i>	151
<i>Results</i>	154
<i>Discussion</i>	158
Experiment 3.....	159
<i>Methods</i>	159
<i>Results</i>	164
<i>Discussion</i>	173
General Discussion.....	175
Supplementary Materials.....	180
Chapter 6: Blinded or deafened by our thoughts? Testing for dissociable effects of auditory and visual mental imagery on early perceptual processing	186
Chapter Summary.....	186

Abstract.....	187
Introduction.....	188
Methods	192
Results	199
Discussion.....	206
Supplementary Materials.....	211
Chapter 7: General Discussion	216
Implications	223
<i>Theoretical implications for the attention literature</i>	223
<i>Theoretical implications for the multisensory integration literature</i>	224
<i>Real-world implications</i>	226
Future Directions.....	227
Conclusions.....	231
References	232

Summary

In our daily lives, we constantly receive input from multiple senses, including sights, sounds, odours, tactile sensations, and tastes. Information from one sensory modality may interact with, and affect the processing of, information from another sense, and stimuli from different senses may be integrated and perceived as one multisensory event.

Multisensory stimuli have been argued to be particularly effective at capturing attention, and the overarching aim of the first part of this thesis was to investigate the proposed ‘special’ attentional status of external multisensory stimuli, and their effects on both attentional capture and awareness. Using well-established manipulations of perceptual load, multisensory stimuli presented as search targets were demonstrated to offer an advantage over unisensory stimuli. Whilst still being subject to modulation by visual perceptual load, this advantage presented both in terms of speed of detection (Chapters 2 & 5) and awareness (Chapter 4) and persisted even when the load itself was multisensory (Chapter 5). On the other hand, multisensory stimuli presented as irrelevant distractors were found to be no more distracting than unisensory stimuli (Chapter 2), and ERP evidence suggests that the two sensory modalities involved were actively suppressed through two independent processes, with no multisensory integration (Chapter 3). Additionally, multisensory stimuli do not appear to impose on perceptual capacity any differently than unisensory stimuli in a primary task (Chapter 5). The final study looks at a different type of multisensory interaction, in terms of the impact of internal mental imagery on external sensory processing. Chapter 6 presents fMRI evidence of differing effects on early perceptual processing of visual and auditory

external stimuli, by visual and auditory internal imagery. Theoretical and practical implications of all findings are considered throughout, and in Chapter 7.

Preface

This thesis conforms to an article format, whereby the five empirical chapters are written as discrete articles. Each paper is referenced and formatted in APA style, and all references are presented in alphabetical order at the end of the thesis. An introduction chapter precedes the empirical chapters, which are then followed by a discussion chapter summarising the research undertaken. I present a short summary at the beginning of each empirical chapter to tie them together.

I have been the lead author on all papers regarding empirical work and writing. Dr Sophie Forster and Professor Jamie Ward have supervised the empirical work and writing of this thesis, and Dr Sophie Forster has provided feedback on the general introduction and discussion. Additional advice and feedback was provided by Professor Salvador Soto-Faraco (Chapters 2, 3 and 5), Dr Nick Berggren (Chapter 3), and Dr Chris Racey (Chapter 6). I was responsible for the programming of, and data collection for, all experiments in this thesis with three exceptions: Chapter 2 Experiment 1 and Chapter 5 Experiment 1, where data collection was conducted by Dr Amanda Sjöblom for her Masters dissertation, and Chapter 4 where Rosie Tucker and Laura Perryman provided assistance with data collection.

Chapter 2 has been published in *Cognition* as: Lunn, J., Sjöblom, A., Ward, J., Soto-Faraco, S., & Forster, S. (2019). Multisensory enhancement of attention depends on whether you are already paying attention. *Cognition*, 187, 38-49.

Author contributions are as follows: I was responsible for programming and collecting data for Experiments 2-4 - A.S. programmed and collected data for Experiments 1a and 1b for her Masters dissertation. I conducted all data analysis, and drafted the manuscript with feedback from S.F., S.S-F. and J.W.

Chapter 1: General Introduction

Every day, we constantly receive input from multiple senses, including sights, sounds, odours, tactile sensations, and tastes. Information from one sensory modality may interact with, and affect the processing of, information from another sense, and stimuli from different senses may be integrated and perceived as one multisensory event. These interactions between the senses occur in even the most typical of daily activities, such as having a conversation, driving a car, crossing a road, or watching television. Multisensory stimuli have been argued to be particularly effective at capturing attention, and the overarching aim of the first part of this thesis was to investigate the proposed ‘special’ attentional status of external multisensory stimuli, and their effects on both attentional capture and awareness, through the application of Perceptual Load Theory (Lavie, 1995). In this chapter, I will first outline the factors which influence multisensory integration, as well as describing the methods commonly used to measure this phenomenon. Next, I will give an overview of the evidence demonstrating the special attentional status of multisensory stimuli, along with that suggesting that integration itself may depend on attention. I will then describe the methods commonly used to measure attentional capture and awareness, before outlining research in support of Perceptual Load Theory, with particular reference in both sections to studies conducted either crossmodally or in the context of multisensory attention. Chapter 6 of this thesis examines a different type of multisensory interaction, in terms of the impact of internal mental imagery on external sensory processing. I will present research on crossmodal interactions occurring between internally generated sensory imagery, and external perceptual processing, given growing evidence regarding the similarities between these processes. Finally, I will outline the specific research aims of this thesis.

Multisensory interactions & integration

Interactions between multiple senses may cause differing perceptual experiences than would occur if the individual senses were processed independently. For example, the concurrent presentation of a visual stimulus with a sound may alter where the sound was perceived to have occurred from, even when instructions are given to ignore the visual modality (e.g. Bertelson et al., 2000; Bertelson & Radeau, 1981), indicating the automaticity of crossmodal interactions. Additionally, The McGurk effect (McGurk & Macdonald, 1976), is a multisensory illusion whereby conflicting visual and auditory streams cause the observer to perceive a ‘fused’ event between the two (for example hearing the syllable ‘ba’ but seeing someone mouth ‘ga’ may be perceived as ‘da’), and has previously been used as supporting evidence for the idea that multisensory integration, at least in terms of speech, is an automatic and unavoidable process (e.g. Soto-Faraco et al., 2004). Typically however, multisensory integration - the process by which input from two or more senses at a time are combined - gives rise to a more accurate perception of an event, and thus is likely to have been extremely important throughout evolutionary history (Stein et al., 2014).

There are a number of principles, or ‘rules’, which have been identified regarding the regulation of multisensory integration. First, the principle of inverse effectiveness, which states that neuronal response enhancements to multisensory stimuli maximally occur where the unisensory constituent parts are minimally effective at eliciting a response independently (Stein & Meredith, 1993). This effect has also been demonstrated behaviourally, whereby facilitatory effects on reaction times to multisensory stimuli increased as stimulus intensity decreased (Diederich & Colonius, 2004), with ERPs (Senkowski et al., 2011), and using fMRI (Stevenson & James, 2009). The ‘spatial rule’ asserts that responses to multisensory stimuli are enhanced when the

unisensory elements are presented at approximately the same location (Holmes & Spence, 2005; Wallace et al., 1992; see Spence, 2013 for a review and discussion of indications for a task-dependency of this rule), and the ‘temporal rule’ contends that responses are enhanced when the constituent stimuli are presented at approximately the same time (Holmes & Spence, 2005; Meredith et al., 1987). In addition to these three main rules, the effect of semantic congruency on multisensory integration has also been examined, suggesting that where the unisensory elements are congruent with each other (e.g. a picture of a dog with the sound of a dog barking), integration is more likely to occur. This has been demonstrated both in terms of a behavioural facilitation effect (Laurienti et al., 2004), and an enhanced neuronal response (Hein et al., 2007; see Doehrmann & Naumer, 2008 for a review), with recent evidence suggesting that task relevance and perceptual load modulate the facilitatory effect of semantic congruence (Kvasova & Soto-Faraco, 2019).

Methods of measuring multisensory integration

Early investigations measuring neuronal indicators of multisensory integration were conducted on the superior colliculus of cats. Merideth & Stein (1983) documented how the responses of neurons in this region of the brain to a particular sensory stimulus were altered by the presence of an additional sensory input. Responses in individual cells to a multisensory stimulus were compared with the responses to each unisensory component, and two patterns were observed in different cells. Some cells were found to show ‘response enhancement’, whereby the individual senses may have elicited only a small response, or none at all, but when two were combined a dramatically increased response was observed. On the other hand, the ability of other cells to respond to a particular sense was inhibited by the simultaneous presentation of a stimulus in another

sense, and the authors theorised that these cells act as response inhibitors, preventing inappropriate responses to stimuli. Whether or not incoming information from two senses elicits response enhancement or response depression relates to the principles of multisensory integration described above. In conditions where the two senses are spatially or temporally aligned, the responses in multisensory neurons are enhanced, but where they are separated they often elicit response depression, or fail entirely to be integrated (Kadunce et al., 2001; Meredith et al., 1987; Meredith & Stein, 1986, 1996; Stein et al., 2014). Putting this into perspective in terms of real-world events, we can see that this is highly logical. Sensory information conveyed by the same event, e.g. the sound and sight of a barking dog, will come from the same spatial location, and occur simultaneously, and thus an enhanced response to this one event is an advantage. However, stimuli from two separate events, occurring either at two different locations or with temporal misalignment, should not be integrated and it would be inappropriate, and likely less beneficial, to perceive these as one event rather than two separate ones.

Studies measuring multisensory integration in the superior colliculus of cats also revealed two further effects which are notable for this thesis. First, parallel patterns of this response enhancement were observed not only in the neuronal responses of the cats, but also in their behaviour. Stein et al., (1988) trained cats to move towards a visual or auditory stimulus for a food reward, and found that multisensory stimuli which had previously been found to enhance the responses in superior colliculus neurons (i.e. spatially and temporally aligned) also enhanced the overt behaviour of the cats, resulting in more correct responses. This multisensory enhancement of spatial localisation and orientation of attention has since been demonstrated in numerous studies in humans, which will be presented later in this chapter.

Secondly, whilst Meredith and Stein (1983) defined multisensory enhancement as being when the multisensory response was larger than the largest response to either of the two unisensory components, it has been demonstrated that these neuronal responses can be nonlinear, or 'superadditive', whereby the response to the multisensory stimulus is larger than the sum of the two unisensory responses (e.g. Meredith & Stein, 1986). This criterion for determining whether multisensory integration has taken place has since been used in event related potential (ERP) studies in humans. Giard and Peronnet (1999) recorded ERPs while participants performed a forced-choice categorisation task, confirming that, behaviourally, multisensory objects were identified faster and more accurately than unisensory ones. They then compared the ERP response to a multisensory (AV) object, with the sum of responses to the visual and auditory objects presented independently (A+V). This AV-(A+V) comparison revealed early differences indicative of integration at scalp regions where potentials are typically evoked in response to visual (parietal-occipital) and auditory (fronto-central) stimuli. These non-linear enhancements to ERP responses have been found in subsequent studies (e.g. Fort et al., 2002; Foxe et al., 2000; Molholm et al., 2002; Santangelo & Spence, 2007; Talsma et al., 2007; Talsma & Woldorff, 2005; Van der Burg et al., 2011) on audiovisual, as well as auditory-somatosensory integration, suggesting that these multisensory interactions affect early sensory processing.

Using the same principle, Calvert et al., (2001) provided one of the first demonstrations of multisensory integration using fMRI. They found that a superadditive response was elicited in the voxels in the superior temporal sulcus (STS) when semantically congruent visual and auditory stimuli were presented, whereas when the stimuli were incongruent the response was subadditive. The BOLD response in the STS has also been demonstrated to follow the principle of inverse effectiveness, with

responses to multisensory stimuli being relatively larger in comparison to the summed unisensory counterparts as signal-to-noise ratio decreases (Stevenson & James, 2009).

Whilst the criterion of superadditivity for inferring the occurrence of multisensory integration is commonly used in neuroimaging studies, Stanford and Stein (2007) have since highlighted that multisensory behavioural advantages are not only due to superadditive responses, since even linear summation of the two unisensory stimuli would represent an increase in superior colliculus activity compared to a unisensory stimulus alone, and could be expected to enhance behavioural responding (Stanford et al., 2005). Additionally, Besle et al. (2009) highlight that whilst single cell recording research examines superadditivity in multisensory neurons, the activation of these cannot be inferred from ERP or fMRI studies, which examine activation at the population level (see also, Giard & Besle, 2010).

These authors additionally highlight the importance of, and provide recommendations to assist with, avoiding confounds when applying the criterion of superadditivity to ERP research. One such example is to avoid analysing time periods that may contain neural responses to other, non-sensory processes, for example the P3 component reflecting target processing (Besle et al., 2004, 2009; see also Hillyard et al., 1998). A second example is to ensure the balance of attention between conditions. If a design is blocked such that the unisensory condition is measured when participants only need to attend to one modality, sensory cortices of other, unattended, modalities may be deactivated (e.g. Laurienti et al., 2002). Subtracting this deactivation from the multisensory response when testing for non-linearity would therefore reveal an effect that is not in fact due to integration (Besle et al., 2004). Furthermore, the authors note that since ERP responses consist of both positive and negative voltages, both

superadditivity and subadditivity reflect a nonlinear response to a multisensory stimulus, from which it cannot be determined whether this reflects enhancement or suppression of the response. They therefore advocate for comparing latency and topographic maps of the responses, in order to distinguish between the two (Besle et al., 2009; Giard & Besle, 2010).

Stein et al., (2009) argue for the use of multiple measures for a more ‘complete perspective’ of a dataset, and the merit of behavioural measures has also been noted (Laurienti et al., 2005). Behavioural analyses of multisensory interactions have typically demonstrated that these stimuli elicit faster responses than unisensory stimuli, known as the redundant signals effect (Hershenson, 1962; Kinchla, 1974; Miller, 1982). It has been suggested that this may merely be due to participants responding to the first stimulus detected (Raab, 1962), but this ‘race model’ can be tested on reaction time distributions, to investigate whether reaction times to multisensory stimuli exceed the statistical facilitation which would be predicted by probability summation based on two unisensory signals (Miller, 1982). Violations of the race model strongly suggest that it is the integration of the two unisensory stimuli that facilitates response times, and this has been widely demonstrated throughout the literature (Crosse et al., 2019; Freeman et al., 2018; Hughes et al., 1994; Laurienti et al., 2004; Molholm et al., 2002; Murray et al., 2005; Pannunzi et al., 2014; though see Otto & Mamassian, 2012).

Throughout the empirical chapters, a number of the techniques presented above are employed. In Chapters 2 and 5, response times to multisensory stimuli are compared with unisensory stimuli, and violations of the race model inequality tested for. In Chapter 5 electrophysiological data is recorded, and ERPs are examined for evidence of superadditivity. ERPs are also utilised in Chapter 3, with direct comparisons between

multisensory and unisensory responses. Finally, in Chapter 6, fMRI is used to examine the impact of crossmodal sensory mental imagery on the perceptual processing of external sensory stimuli, in the context of sensory interactions between internally generated and externally presented stimuli.

Multisensory attention

Evidence of a special attention status for multisensory stimuli

As well as demonstrating neuronal enhancements resulting from multisensory integration, seminal research using single cell recordings from the superior colliculus of cats also revealed that these enhancements came with a facilitatory effect on orientation and detection responses (e.g. Stein et al., 1989). Activity in this same region in humans has also been implicated in a multisensory enhancement of spatial orientation to a target (Leo et al., 2008), and faster behavioural and saccadic reaction times to multisensory targets than unisensory targets has been consistently demonstrated throughout the literature (Colonius & Arndt, 2001; Crosse et al., 2019; Freeman et al., 2018; Frens et al., 1995; Hughes et al., 1994, 1998; Laurienti et al., 2004; Molholm et al., 2002; Murray et al., 2005; Pannunzi et al., 2014).

Further evidence in support of a special attentional status for multisensory stimuli has been observed through the ‘pip and pop’ effect, a phenomenon in visual search whereby the presence of an auditory ‘pip’ in time with a target colour change significantly speeds up search times (produces ‘pop out’) in an otherwise difficult (serial) search task (Van der Burg et al., 2008). Additionally, enhanced spatial cuing by multisensory stimuli has been demonstrated (see Spence & Santangelo, 2009 for a review). Two studies have provided evidence suggesting that multisensory cues (both audiovisual and audiotactile) can capture attention more effectively, and therefore

produce stronger cueing effects, than unisensory cues (Santangelo et al., 2008; Santangelo & Spence, 2007). Whilst all cues were able to capture attention when participants were not engaged in another task, unisensory cues (but not multisensory cues) were rendered ineffective when participants were also performing a demanding central Rapid Serial Visual Presentation (RSVP) task. Note that the authors of the original study have since provided further evidence to suggest that the ineffectiveness of the unisensory cues during the simultaneous RSVP task was due to the transients from the central RSVP causing a faster disengagement of attention from the cued location prior to target presentation (Santangelo et al., 2011). It is as yet unclear why the central transients did not similarly disrupt multisensory cues, and hence whether the preservation of multisensory cuing effects during the RSVP task reflects enhanced attentional capture, or delayed disengagement. Nevertheless, these findings do appear to reflect some kind of ‘special’ attentional status for multisensory stimuli, and Spence and colleagues have pointed to promising implications regarding the application of multisensory cues during demanding tasks in real life context. For example, a multisensory warning signal in a driving simulator appears to be particularly effective at capturing attention and eliciting faster braking responses in emergency situations (Ho et al., 2007).

Evidence that multisensory integration depends on attention

The suggestion that multisensory stimuli could be particularly effective in capturing attention during demanding tasks, supported by the studies above, is challenged by work indicating that multisensory integration may be strongly limited when the stimuli occur away from the focus of attention. Alsius et al. (2005) examined attentional influences on the McGurk effect, and showed that when participants were

involved in a dual task, they were less likely to hear the fused event. This implies that multisensory integration was disturbed by the additional task demands, and has been further supported through related ERP evidence, with demonstration that multisensory interactions shown in ERP components are disrupted under higher attentional load (Alsius et al., 2014). Similarly, in an fMRI study Fernández et al. (2015) demonstrated that a benefit of multisensory integration (i.e. enhanced behavioural and neuronal response) to visual and auditory speech streams was only observed when both modalities were attended.

Attentional modulation of multisensory integration has not only been demonstrated in complex, linguistic stimuli, but also in experiments which utilised more simplistic ‘flash and beep’ events (e.g. Pápai & Soto-Faraco, 2017; Senkowski et al., 2005; Talsma et al., 2007; Talsma & Woldorff, 2005). Talsma and Woldorff (2005) demonstrated an early superadditive response to attended multisensory stimuli, comprising of visual gratings with spatially and temporally matched auditory ‘pips’, which was eliminated when the stimuli were unattended. This finding was furthered in a subsequent study using the same multisensory stimuli, and which demonstrated that a superadditive early integration effect was observed only in conditions where both the visual and the auditory modalities were attended at the same time (Talsma et al., 2007). Given the evidence that, in some cases, the application of top-down attention does not appear to be a requirement for multisensory integration (Bertelson et al., 2000), a review by Talsma et al. (2010) suggests that the degree of competition from co-occurring events (e.g. perceptual load, see section below for more detail) plays a large role in determining whether top-down attention is a requirement for integration.

Attention may be allocated to a stimulus for a number of different reasons, for example due to being targets or matching the attentional set of a search task in some way (e.g. Bacon & Egeth, 1994; Folk & Remington, 1998), being presented at a task-relevant location (i.e. inside our ‘attentional window’, Belopolsky & Theeuwes, 2010), or because we have remaining perceptual capacity when a task is less demanding (see Lavie, 1995). This therefore raises the interesting question of what degree of attention may be sufficient for multisensory integration to occur.

Methods of measuring attentional capture by, and awareness of, multisensory stimuli

Perhaps the most commonly used measure of attentional capture, widely used in purely behavioural studies as well as in combination with neuroimaging methods, is reaction time. Attentional capture can be measured both through facilitatory effects on reaction time, whereby a target is identified more rapidly when its features capture attention, as well as distractor interference, whereby a salient but irrelevant distractor disrupts (e.g. slows down) performance from a main task because it summons attention automatically. One such example of this is the ‘irrelevant singleton’ paradigm (Theeuwes, 1991, 1992), where participants search for a target amongst an array of items, and one of the items in the display may be a unique and salient singleton which is irrelevant to the task. By demonstrating that reaction times are slower when the distractor singleton is present compared to when it is absent, it can be assumed that attention is first captured by the distractor before the target is subsequently located. This measure has been applied to demonstrations of attentional capture that occur in modalities other than vision, for example reaction time to an auditory search task is slowed by the presence of irrelevant auditory singletons (e.g. Dalton & Lavie, 2004), and also in demonstrations of attentional capture by multisensory stimuli (e.g. Van der

Burg et al., 2008), where the search for a visual target is faster with the presentation of an auditory stimulus. Another example of measuring attentional capture with reaction times, comes from the pre-cueing paradigm, where participants are shown an irrelevant and uninformative spatial cue before a search array (see Simons, 2000). If reaction times to the target are faster if the pre-cue had appeared at its location, or slowed if it had appeared at an invalid, non-target location, then attentional capture by the cue is assumed to have occurred (e.g. Folk et al., 1992, 1994; Folk & Remington, 1998). This measure of attentional capture has also been applied to studies of multisensory integration, for example the aforementioned study by Santangelo and Spence (2007), examining the cueing effects elicited by visual, auditory and multisensory (audiovisual) cues under different conditions of primary task demand.

Reaction times are an implicit measure of attentional capture, as we infer that it has occurred due to the increase or decrease in search times (Simons, 2000). A similar measure is that of oculomotor capture, where a stimulus is deemed to have captured attention if an eye movement is made towards it (Simons, 2000). Using the irrelevant singleton paradigm, Theeuwes et al. (1998) demonstrated an involuntary saccade to the onset of an irrelevant distractor before the goal driven eye movement to a target, arguing that this provides evidence for it having captured attention. Faster oculomotor capture by visual targets has been demonstrated when combined with a spatially or temporally matched auditory (Colonius & Arndt, 2001; Frens et al., 1995; Hughes et al., 1998) or tactile (Amlôt et al., 2003; Diederich et al., 2003) stimulus (see Colonius & Diederich, 2004, for a review).

Within the attentional capture literature, one indirect measure that may be particularly useful, particularly where behavioural interference may not be sensitive

enough to detect fast acting capture, are electrophysiological indices of the processes of spatial orientation of attention and distractor suppression. Two lateralised ERP components are primarily used in this area of research; the N2pc and the P_D. The N2pc component is held to reflect attentional selection of items in visual space (Eimer, 1996; Kiss et al., 2008; Luck & Hillyard, 1994). This component is reflected at posterior electrodes, characterised by a negative deflection in the ERP waveform contralateral to a stimulus presented at an attended location, and elicited 200-250ms post stimulus onset. The N2pc has been demonstrated to be elicited when an auditory signal aids the search for a visual target (Van der Burg et al., 2011) and an analogous component has also been documented for the selection of information within the auditory domain, occurring at anterior electrodes – the N2ac (Gamble & Luck, 2011). The P_D component is the opposite of the N2pc - a contralateral positivity, occurring at the same electrode sites - which is believed to index spatially localised suppression of a distractor stimulus in response to a spatial ‘attend-to-me’ signal (Hickey et al., 2008; Sawaki & Luck, 2010). The size of the P_D has been found to correlate with behavioural distractor interference (Gaspelin & Luck, 2018), and it predicts whether or not a distractor elicits an N2pc (McDonald et al., 2013). Thus, a distractor can elicit an ‘attend-to-me’ signal and be actively suppressed, despite not producing a behavioural cost. As of yet, no auditory analogue of the P_D has been established, neither has it been demonstrated to occur in the presence of entirely irrelevant multisensory distractors.

As well as attentional capture, another way of measuring potentially prioritised processing is through the awareness of a stimulus (Simons, 2000). This measurement can be merely asking participants if they observed anything unexpected during an event, as in Simons and Chabris (1999)’s classic study of the ‘Gorillas in our Midst’. After watching a video, during which participants were instructed to count the number of

passes made by an attended team of basketball players, the percentage of participants who noticed a gorilla on the screen was surprisingly low, indicating the absence of attention to this event. Measuring the conscious perception of irrelevant stimuli through the rate of awareness has been applied many times in the inattention blindness and deafness literature, where people fail to notice a visual or auditory event due to having their attention engaged in a primary task (e.g. Cartwright-Finch & Lavie, 2007; Mack et al., 1998; Most et al., 2001; Newby & Rock, 1998). This has since been extended to utilise measures from Signal Detection Theory (e.g. Macdonald & Lavie, 2008). By presenting an irrelevant stimulus in multiple trials throughout an experiment, both correction detection (hit) rate, and false alarm rate can be recorded, and thus a measure of detection sensitivity (d') established for the stimulus. This measure shows how sensitive a participant is in responding to the presence of a signal (or, critical stimulus), taking into account not only how often they detected the signal, but also when they were able to correctly deduce that it was not there (correct rejection). Whilst the initial research involved a single trial with an unexpected event, this measure differs in that it examines awareness of expected stimuli presented in the context of a secondary task, and with responses being made after each individual trial rather than being measured retrospectively, it is less dependent on memory (see Macdonald & Lavie, 2008). This measure has been applied within the visual domain, whereby detection sensitivity of irrelevant visual stimuli was low when participants were engaged in a demanding visual task (Macdonald & Lavie, 2008), and also crossmodally, with attention to a visual task also found to disrupt conscious awareness of auditory (Raveh & Lavie, 2015), as well as tactile (Murphy & Dalton, 2016), stimuli, but has not, as of yet, been applied to multisensory stimuli.

In this thesis, I will be applying the majority of the above methods to measure attentional capture by multisensory stimuli, including those that have not yet been applied to this body of research. In Chapters 2, 3 and 5, reaction times are used to examine both facilitatory effects and distractor interference of multisensory compared to unisensory stimuli. In Chapter 3, the P_D is used to investigate the spatial suppression of entirely irrelevant visual, auditory, and audiovisual, distractors, and in Chapter 4 conscious awareness of multisensory stimuli will be measured using d' .

Perceptual Load Theory and multisensory attention

Lavie (1995) proposed Load Theory in order to resolve the long-standing conflicts between early- and late- selection hypotheses. Treisman's attenuation model (Treisman, 1960) is one such example of an early selection theory, arguing that irrelevant information is filtered out at the early, perceptual stage, whilst only relevant information is selected for later processing. This is supported by evidence from the dichotic listening task (Cherry, 1953) where participants are presented with a different auditory speech stream to each ear and are asked to attend only to one of them. In this task, participants are typically able to report surface level features such as the physical characteristics of the unattended voice, but not the content of the stream, showing that they successfully filtered it out before semantic processing. However, MacKay (1973) used this task to show that if the attended stream contained ambiguous information, such as a word that could mean two different things in different contexts, the unattended stream was able to bias the interpretation of the attended phrase. Additionally, in the response competition flanker task (Eriksen & Eriksen, 1974), an incongruent distractor presented in an irrelevant location slowed down reaction times to a central letter search task, indicating that its identity was processed. Thus, late selection models, such as

those proposed by Deutsch and Deutsch (1963), Kahneman (1973) and Duncan (1980), argue that no information is filtered out until a later stage of processing.

These conflicting findings suggest that both early and late selection are possible depending on the particular circumstances, and Load Theory provides a useful theoretical framework to predict the contexts in which selection, and thus attentional capture, is more likely to occur. The theory posits that attention works with a limited perceptual capacity, automatically processing stimuli until capacity is depleted. During tasks which involve only low perceptual load, spare capacity remaining after processing relevant information spills over to allow processing of other, less relevant, stimuli (i.e. late selection occurs). On the other hand, under high perceptual load conditions, all processing capacity must be fully devoted to the relevant task and therefore stimuli irrelevant to the primary task are typically not processed (i.e. early selection occurs).

Load Theory has been supported by a large body of evidence using various different manipulations of load, and various measures of task-irrelevant processing. Manipulations of perceptual load fall largely into two categories. One type of manipulation involves performing the same task in conditions with varying amounts of information, and early evidence in support of Perceptual Load Theory comes from studies using this approach. In these initial experiments, participants were required to search for a target letter whilst ignoring a letter distractor. In low perceptual load conditions the target letter was presented alone (Lavie, 1995), or with small place holder 'o's (Lavie & Cox, 1997), whereas under conditions of high perceptual load it was presented amongst five other non-target letters. In the low load task, distractor letters that were congruent to the target increased reaction times, but where they were incongruent reaction times were slowed. This effect was eliminated under high load

when perceptual capacity was filled by the letter search task and the distractor was not processed. Forster and Lavie (2008a) further developed this paradigm to show the effect of perceptual load on entirely irrelevant distractors, using the same letter search task presented centrally, with distractor stimuli being cartoon characters presented at irrelevant peripheral locations, and a wide body of literature has further supported this approach (Cunningham & Egeth, 2018; Forster et al., 2014; Forster & Lavie, 2011, 2016; He & Chen, 2010; Morris et al., 2020). However, the extent to which these effects occur crossmodally is less clear cut, with evidence even for the opposite pattern of results when distractors are presented in the auditory domain. Tellinghuisen & Nowak (2003) demonstrated that auditory distractors elicited larger compatibility effects during high visual perceptual load compared to low, proposing that it may be the inhibition of cross-modal processing that is disrupted under conditions of high perceptual load, thus resulting in increased interference. In terms of multisensory distractors, Matusz et al. (2015) found no difference in the level of interference from multisensory audiovisual distractors versus unisensory auditory distractors, with both being modulated by perceptual load, in a response competition flanker task using coloured shapes. As of yet, the effect of entirely irrelevant multisensory distractors, being presented in irrelevant locations and sharing no features with the target, has not been examined.

The second type of perceptual load manipulation involves using the same stimuli in both conditions, but changing the task so that it becomes more or less perceptually demanding – such as searching an RSVP stream for a target defined by either a single feature (low load) or a conjunction of features (high load). For example, Schwartz et al., (2005) presented a stream of coloured ‘T’s, which could be either upright or inverted. Under conditions of low perceptual load targets were any red T, regardless of orientation, and under conditions of high load targets were defined by a conjunction of

features – either an upright yellow T or an upside-down green T. Analysis of the BOLD response in the visual cortex showed that high perceptual load reduced activations to irrelevant peripheral stimuli as early as V1, with effects increasing across successive visual areas (V1-V4). A similar study found comparable effects for peripheral stimuli presented subconsciously (Bahrami et al., 2007), and Rees et al., (1997) demonstrated reduced motion processing of an irrelevant background array of moving stars in V5 whilst participants were engaged in a high perceptual load compared to low perceptual load task.

The effect of perceptual load has also been demonstrated in the context of awareness. Cartwright-Finch & Lavie (2007) showed that participants were less likely to report detection of a critical visual stimulus whilst engaged in a high perceptual load vs low perceptual load task. This suggests that visual perceptual load modulates inattention blindness, and was further supported by demonstrations that increased load reduces detection sensitivity of a visual peripheral stimulus presented in the context of a secondary task (Macdonald & Lavie, 2008). There have been a small number of demonstrations of crossmodal effects of visual perceptual load on awareness of stimuli in other sensory modalities, for example visual load induced inattention to auditory (Macdonald & Lavie, 2011; Raveh & Lavie, 2015), tactile (Murphy & Dalton, 2016), and olfactory stimuli (Forster & Spence, 2018), but no demonstrations of the effect on multisensory stimuli. Within applied contexts, high visual perceptual load in a driving task has been shown to decrease accuracy in recalling when the sound of braking in a car crash occurred (Murphy & Greene, 2016), and to reduce awareness of both driving-relevant, and driving-irrelevant, visual and auditory stimuli (Murphy & Greene, 2015). It therefore appears that the effects of perceptual load can also occur crossmodally in real-life scenarios.

Whilst the vast majority of demonstrations in support of Load Theory have involved effects of load in a visual primary task, recent studies have also examined the effect of auditory perceptual load, with conflicting results. For example, one extensive investigation using a range of perceptual load manipulations in an auditory primary task found no modulation of an auditory distractor effect (Murphy et al., 2013). On the other hand, Fairnie et al., (2016) more recently looked at the effect of auditory perceptual load on awareness, and found that detection of a critical stimulus was reduced when the number of sounds in an audio-spatial search task was increased, consistent with the perceptual load literature in the visual domain. The variability of evidence obtained for the effects of auditory compared to visual perceptual load is discussed in detail in a review by Murphy et al., (2017), theorising that one's full attentional capacity is unlikely to be allocated exclusively to task-relevant auditory stimuli. Given these differences between visual and auditory perceptual load, and the fact that in our daily lives we rarely encounter tasks which are presented solely in one modality, it would be of interest to determine the effect of a multisensory perceptual load.

Mind-wandering and mental imagery

All the research documented above concerns multisensory interactions and integration when both unisensory counterparts come from the external world. An interesting line of evidence suggests that the neural substrates involved in creating a conscious percept of an external sensory stimulus substantially overlap with those involved in generating mental imagery (e.g. Chen et al., 1998; Ganis et al., 2004; Halpern et al., 2004; Ishai et al., 2002; Klein et al., 2000; Lambert et al., 2002; Le Bihan et al., 1993; Naselaris et al., 2015), and this thesis will additionally consider crossmodal

interactions between external perceptual processing and internally generated sensory events.

Using multivariate pattern analysis (MVPA) to investigate crossmodal interactions in sensory processing, Vetter et al., (2014) found that a classifier could successfully discriminate three types of natural sounds in the auditory cortex, as well as in the early visual cortex. In a second experiment, successful classifications in both cortical regions were again found when participants were required just to imagine the sounds, rather than listening to them. Additionally, the feedback of information from the perception of both auditory external sounds and auditory internal imagery to the visual cortex was mediated by the posterior STS. Given the role of the STS in multisensory integration (Calvert et al., 2000; Stevenson & James, 2009), this may suggest that we are able to integrate our percept of an external stimulus occurring in one modality with internally generated imagery occurring in another, as well as integrating imagery in two different modalities.

Whilst an internally generated visual representation of a particular event may integrate with a congruent auditory mental representation (either generated internally or through external perceptual processing), competition may occur if they are incongruent, or if one is irrelevant to our primary task. Following their investigations into the effect of perceptual load on entirely irrelevant external visual distractors, Forster and Lavie (2009) examined its effect on distractions generated internally. Across a series of experiments they found that perceptual load to a visual central task modulated the occurrence of task-unrelated thoughts, or, 'mind-wandering', with results suggesting that attentional capacity is shared between external perceptual processing and internal thoughts, with a filling of capacity similarly reducing the processing of both task

irrelevant external and internal information. This finding has since been replicated in an applied context, with a decreased occurrence of mind wandering whilst driving in a more perceptually demanding scenario (Geden et al., 2018). Additionally, reduced neuronal responses to external events during periods of mind wandering have been demonstrated during paradigms presented in both the visual (e.g. Barron et al., 2011; Kam et al., 2013; Smallwood et al., 2008), and the auditory domain (e.g. Braboszcz & Delorme, 2011; though see Kam et al., 2013).

Mind wandering shares a number of features with mental imagery (see Villena-González & Cosmelli, 2020 for a review) - whilst our attention is directed internally, we may experience vivid imagery in a variety of senses. Thus, this raises an interesting question regarding the extent to which a competition for resources may arise when mental imagery occurs at the same time as external perceptual processing, and whether such a competition occurs crossmodally, or is modality specific. For example, when driving a car, if our thoughts are occupied with picturing a friend we are meeting, or a place we would like to travel to, could this reduce our visual processing of the road ahead? And if instead we were occupied by an ‘earworm’, a song stuck in our heads, would any reduction happen to the same extent, if at all?

Initial evidence comes from an ERP study conducted by Villena-González et al. (2016). They implemented a paradigm whereby participants were instructed to either freely imagine anything they wanted provided it was solely visual, or to think using inner speech with no mental imagery, whilst checkerboards appeared on a computer screen in front of them. In an additional set of instructions, participants were required to either attend to these checkerboards, or to the contents of their thoughts. This study demonstrated that ERP measures of sensory processing (the visual P1 component) were

reduced when attention was directed internally compared to externally, and that this reduction was greater when participants were engaged in the visual mental imagery compared to the inner speech. Additionally, increased alpha power, associated with cortical attentional suppression (e.g. Foxe & Snyder, 2011) was found when participants were engaged in visual mental imagery compared with inner speech, showing that processing of the external visual stimulus was differently affected depending on the modality of the internal thoughts. In a second study, using the same paradigm but with auditory ‘beep’s rather than visual checkerboards as the external stimuli, Villena-González et al. (2018) found that the ERP measures of sensory processing (P1, auditory N1 component) did not differ depending on whether participants were required to direct their attention internally or externally, though the reduction in amplitude of gamma band activity was proposed to reflect reduced conscious attention to the external stimuli, given it has previously been linked to top-down attentional processes (Debener et al., 2003). Furthermore, beta activity was found to differ only during visual imagery. Modulation of this activity has been associated with audiovisual integration (e.g. Hipp et al., 2011; Roa Romero et al., 2015), therefore this result is suggestive of crossmodal interactions when the visual imagery is processed at the same time as the auditory external stimulus.

The research conducted by Villena-González and colleagues suggests that the effect of mental imagery on external perceptual processing differs depending on whether the imagery occurs in the same, or different modality to the external stimulus. However, given the evidence reviewed earlier in this chapter concerning the crossmodal effects of a filled perceptual capacity (e.g. Dalton & Lavie, 2004; Forster & Spence, 2018; Murphy & Dalton, 2016; Raveh & Lavie, 2015), it remains possible that the direction of attention to an internal sensory stimulus may also disrupt processing of

external stimuli crossmodally, and thus an effect may occur regardless of the modality of imagery.

Thesis aims and overview

As reviewed in this introduction, there are a number of areas of research in the area of multisensory attention that have yet to be addressed, and this thesis will endeavour to do so. With multisensory stimuli having been proposed to have a ‘special’ attentional status, a key question is whether or not attentional capture by these stimuli is dependent on the availability of perceptual capacity, as is the case for unisensory visual stimuli. Chapter 2 of this thesis will address whether multisensory stimuli are, as has been suggested, immune to the effects of perceptual load on multisensory facilitation of secondary target detection, and interference by multisensory distractors. Through a comparison of the effects of both facilitation and distraction, this chapter will also examine whether multisensory attentional capture depends on the allocation of resources regulated by top-down attention. Chapter 4 extends the research into the immunity of multisensory stimuli to the effects of perceptual load using a measure which, as noted above, has not yet been applied to multisensory stimuli in this context. Combining previously applied inattention blindness and deafness paradigms to measure detection sensitivity, this chapter will examine whether or not a multisensory stimulus is special in its ability to reach conscious awareness under perceptually demanding conditions.

Chapter 3 furthers the research conducted in Chapter 2 on distraction by multisensory stimuli, using ERP measures. This chapter addresses whether the interference from entirely irrelevant visual, auditory and multisensory distractors reflects competition for spatial attention, while also testing for a potential enhancement

of this competition from the multisensory distractors. Additionally, given the limited evidence of an auditory analogue of spatial attentional selection, through the use of unisensory auditory distractors we aimed to either support previous findings of the N2ac component, or establish for the first time an auditory P_D equivalent.

The research reviewed above shows that prior demonstrations of the effects of perceptual load are limited to unisensory forms, with the majority of this research being on visual perceptual load. Given that many tasks encountered in our daily lives occur in more than one modality at a time, Chapter 5 will explore the effects of a multisensory perceptual load, through both behavioural and ERP measures. The research aims here are twofold; firstly, we aim to determine whether a multisensory load task would reduce the demand on perceptual capacity due to the different sensory components integrating, or whether it would be particularly effective at biasing top down attention to a primary task (and thus, away from a secondary detection task). Secondly, we investigate whether the effects of multisensory load would selectively affect the response to multisensory stimuli in a secondary task, which may be expected if only one multisensory stimulus can be attentionally ‘boosted’ at a time.

Finally, Chapter 6 of this thesis examines a different type of multisensory interaction, exploring the impact of internal mental imagery on external sensory processing. Furthering the previous research into the interference in sensory processing of an external stimulus by thoughts occurring in either the same or a different modality to this stimulus, this chapter uses fMRI to explore the crossmodal effects of auditory and visual mental imagery on auditory and visual perceptual processing, and determine whether there are dissociable effects of mental imagery on external perception in the visual and auditory cortices. Using a broader range of naturalistic scenarios to imagine

than in previous research, particularly for the auditory imagery where participants are asked to think of a range of different sounds as opposed to only inner speech, we additionally look to explore how early on in the processing of external stimuli any potential attenuation may occur.

Chapter 2: Multisensory enhancement of attention depends on whether you are already paying attention

Chapter Summary

In this thesis, I am aiming to investigate the proposed ‘special’ attentional status of multisensory stimuli, with a key question being whether or not attentional capture by these stimuli is subject to modulation by perceptual load, as is the case for unisensory stimuli. A second key question is whether facilitatory effects of multisensory stimuli on attentional capture are dependent on the allocation of top-down attention. In the current chapter, I present both multisensory and unisensory stimuli in the context of two different measures of attentional capture: facilitation (Experiments 1-3) and distraction (Experiment 4), in order to determine whether multisensory enhancement of attentional capture occurs irrespective of a reliance on endogenous attention. All experiments also implement well-established methods to modulate perceptual load, in order to look for any potential immunity that multisensory stimuli may have to the effects of this. This chapter has been published in *Cognition* as: Lunn, J., Sjöblom, A., Ward, J., Soto-Faraco, S. & Forster, S. (2019). Multisensory enhancement of attention depends on whether you are already paying attention. *Cognition*, 187, 38-49.

Abstract

Multisensory stimuli are argued to capture attention more effectively than unisensory stimuli due to their ability to elicit a super-additive neuronal response. However, behavioural evidence for enhanced multisensory attentional capture is mixed. Furthermore, the notion of multisensory enhancement of attention conflicts with findings suggesting that multisensory integration may itself be dependent upon top-down attention. The present research resolves this discrepancy by examining how both endogenous attentional settings and the availability of attentional capacity modulate capture by multisensory stimuli. Across a series of four studies, two measures of attentional capture were used which vary in their reliance on endogenous attention: facilitation and distraction. Perceptual load was additionally manipulated to determine whether multisensory stimuli are still able to capture attention when attention is occupied by a demanding primary task. Multisensory stimuli presented as search targets were consistently detected faster than unisensory stimuli regardless of perceptual load, although they are nevertheless subject to load modulation. In contrast, task irrelevant multisensory stimuli did not cause greater distraction than unisensory stimuli, suggesting that the enhanced attentional status of multisensory stimuli may be mediated by the availability of endogenous attention. Implications for multisensory alerts in practical settings such as driving and aviation are discussed, namely that these may be advantageous during demanding tasks, but may be less suitable to signalling unexpected events.

Introduction

Daily life bombards us with an overwhelming amount of sensory input, including sights, sounds, tactile sensations, odours and tastes. Most of them are simply neglected, others instead, summon our attention. Why do certain sensory stimuli attract (or ‘capture’) our attention while others may not be noticed? The types of stimuli argued to capture attention in this way include ‘singletons’ which differ in some unique attribute (e.g. colour) from surrounding items (Theeuwes, 1992), abrupt onsets (Jonides & Yantis, 1988), moving stimuli (Franconeri & Simons, 2003), or events that have motivational relevance or value (Anderson et al., 2011; Purkis et al., 2011). One type of event that has been proposed to be particularly effective at capturing attention are those which produce correlated stimulation in more than one sensory modality at a time (e.g. Santangelo & Spence, 2007).

Multisensory stimuli are often processed faster or produce stronger responses than unisensory stimuli. According to many studies, this enhanced multisensory response is not merely due to the summed effects of concurrent information, as multisensory stimuli often elicit faster and more accurate responses than would be predicted by additive models of the two unisensory stimuli (Colonius & Diederich, 2004; Hughes et al., 1994; Laurienti et al., 2004; Molholm et al., 2002; Murray et al., 2005; Pannunzi et al., 2014; Senkowski et al., 2005; Talsma et al., 2007; though see Otto & Mamassian, 2012). This has led to the suggestion that multisensory stimuli may also be particularly effective in capturing attention (e.g. Santangelo & Spence, 2007). Whilst this may, under some conditions, be beneficial (i.e. when a multisensory stimulus is of behavioural relevance), it may, on the contrary, be disruptive in other conditions (i.e. by pulling attention away from our current goals). These results have often been taken to assume that some multisensory integration processes happen prior

to, or independent of, the allocation of attention. Contrary to this idea, in the present research we show that whilst multisensory stimuli are particularly effective in capturing attention, this effectiveness is modulated by perceptual load (high load reduces effectiveness in absolute terms, but increases effectiveness relative to that expected from its unisensory parts), and depends on whether or not the stimuli are part of the attentional set (i.e. it is found for targets but not distractors).

Previous Evidence for Attentional Capture by Multisensory Stimuli

Attentional capture can be measured through both its facilitation effects, whereby a target is identified more rapidly or more accurately when its features capture attention, as well as its distractor interference, whereby a salient but irrelevant distractor disrupts (e.g. slows down) performance from a main task because it summons attention automatically (e.g. Theeuwes, 1992). Facilitatory attentional capture has been found with multisensory stimuli, such as in the ‘pip and pop effect’. The ‘pip and pop effect’ refers to a phenomenon in visual search whereby the presence of an auditory ‘pip’ in time with a target colour change significantly speeds up search times (produces ‘pop out’) in an otherwise difficult (serial) search task (Van der Burg et al., 2008). It is less clear, however, whether attentional capture by multisensory stimuli can lead to increased distractor interference. For example, employing the widely used response competition flanker measure of distraction, Matusz et al. (2015) found no difference in the level of interference from multisensory audiovisual distractors versus unisensory auditory distractors.

Another paradigm used to test the ability of multisensory stimuli to capture attention is the spatial cuing task (e.g. Posner cueing task; Posner, 1980). In this task, spatial cues are presented shortly prior to imperative targets, either at, or away from, the

upcoming target location. These can either facilitate or interfere with target detection depending on whether or not they cue the correct target location. Using this task, two studies have provided evidence suggesting that multisensory cues (both audiovisual and audiotactile) can capture attention more effectively, and therefore produce stronger cueing effects, than unisensory cues (Santangelo et al., 2008; Santangelo & Spence, 2007). However, this multisensory superiority was only found if participants were also performing a demanding central Rapid Serial Visual Presentation (RSVP) task. Under such multi-tasking conditions, unisensory cues (but not multisensory cues) were rendered ineffective. The authors of the original study have since provided further evidence to suggest that the ineffectiveness of the unisensory cues during the simultaneous RSVP task was due to the transients from the central RSVP causing a faster disengagement of attention from the cued location prior to target presentation (Santangelo et al., 2011). It is as yet unclear why the central transients did not similarly disrupt multisensory cues, and hence whether the preservation of multisensory cuing effects during the RSVP task reflects enhanced attentional capture, or delayed disengagement. Nevertheless, these findings do appear to reflect some kind of ‘special’ attentional status for multisensory stimuli. Spence and colleagues have pointed to promising implications regarding the application of multisensory cues during demanding tasks in real life contexts and have found, for example, that a multisensory warning signal in a driving simulator appears to be particularly effective at capturing attention and eliciting faster braking responses in emergency situations (Ho et al., 2007)

The suggestion that multisensory stimuli could be particularly effective in capturing attention during demanding tasks, supported by the studies above, is challenged by work indicating that multisensory integration (of auditory and visual information) may be strongly limited when the stimuli occur away from the focus of

attention both in terms of behaviour (e.g. Alsius et al., 2005; Pápai & Soto-Faraco, 2017) as well as in brain responses (Morís Fernández et al., 2015; Senkowski et al., 2005; Talsma et al., 2007). Alsius et al. (2005), examined attentional influences on the McGurk effect (McGurk & Macdonald, 1976), a multisensory illusion whereby conflicting visual and auditory streams cause the observer to perceive a ‘fused’ event between the two (for example hearing the syllable ‘ba’ but seeing someone mouth ‘ga’ may be perceived as ‘da’). Using this illusion, they showed that when participants were involved in a dual task, they were less likely to hear the fused event. This implies that multisensory integration was disturbed by the additional task demands (see also Alsius et al., 2014, for related ERP evidence). One could argue that multisensory processes in complex, linguistic stimuli could be more prone to attention modulation than simple flash and beep events. Yet, Talsma and Woldorff (2005; 2007) demonstrated that non-linear event related potential (ERP) responses, indicating multisensory integration of simple audiovisual stimuli in humans, were observed only in conditions when the stimuli were presented at the location being attended to. Based on this research it seems important to ascertain under what task conditions increased capture occurs, both in terms of understanding the mechanisms of multisensory attention and to inform potential practical applications.

Load Theory and Multisensory Stimuli

Load Theory provides a useful theoretical framework predicting the contexts in which attentional capture is more likely to occur. Lavie (1995) proposed Load Theory in order to resolve conflicts between early- and late-selection hypotheses (see Lavie, 2010, for a review). The theory posits that attention works with a limited perceptual capacity, automatically processing stimuli until capacity is depleted. During tasks which involve

only low perceptual load, spare capacity remaining after processing relevant information spills over to allow processing of other, less relevant, stimuli. On the other hand, under high perceptual load conditions, all processing capacity must be fully devoted to the relevant task and therefore stimuli irrelevant to the primary task are typically not processed.

Load Theory has been supported by a large body of evidence using various different manipulations of load, and various measures of task-irrelevant processing. Perceptual load manipulations fall largely into two categories: One type of manipulation involves performing the same task in conditions with varying amounts of information – for example, searching for a target letter either when presented alone (low load) or among five other non-target letters (high load). The second type of perceptual load manipulation involves using the same stimuli in both conditions, but changing the task so that it becomes more or less perceptually demanding – for example, searching an RSVP stream for a target defined by either a single feature (low load) or a conjunction of features (high load). A key implication of this framework is that increasing the load of a primary task through these methods reduces behavioural interference from irrelevant distractors (e.g. Forster & Lavie, 2008; Lavie & Cox, 1997), decreases BOLD responses in the visual cortex for irrelevant peripheral stimulation (e.g. Schwartz et al., 2005), and reduces sensitivity to detect both auditory and visual peripheral stimuli presented in the context of a secondary task (Macdonald & Lavie, 2008; Raveh & Lavie, 2015). Within applied contexts, high visual perceptual load in a driving task has been shown to decrease accuracy in recalling when the sound of braking in a car crash occurred (Murphy & Greene, 2016), and to reduce awareness of both driving-relevant, and driving-irrelevant, visual and auditory stimuli (Murphy & Greene, 2015). It

therefore appears that the effects of perceptual load can occur crossmodally in real-life scenarios.

Santangelo and Spence (2007) raised the intriguing possibility that the potential to capture attention by multisensory stimuli may be immune to the effects of perceptual load. This was initially concluded because the effects of unisensory cues, but not multisensory ones, were abolished by a dual task condition. However, as noted above, a subsequent study by the same authors changed their interpretation of these results, concluding that the abolition of unisensory cuing effects was not in fact due to perceptual load, but rather due to the dual task condition involving the presentation of an additional stimulus in between cue and target (Santangelo et al., 2011). As such, the question of whether or not attentional capture by multisensory stimuli is immune to perceptual load, remains unanswered. Such an immunity to load has been found in the past for other classes of stimuli thought to be ‘special’ (particularly effective) in their capacity to grab attention, such as human faces (Lavie et al., 2003).

Santangelo and Spence’s proposal has exciting applied implications, such as the utility of multisensory warning-signals during perceptually demanding activities such as driving down a busy street or landing an aircraft. However, the potentially contradictory findings of crossmodal integration being dependent upon attention (Alsius et al., 2005; Pápai & Soto-Faraco, 2017; Senkowski et al., 2005) appear to suggest that, rather than being immune to load effects, multisensory stimuli may not be integrated when conditions demand high levels of attention. If irrelevant multisensory stimuli are not integrated under high load, they would presumably lose the ‘special’ quality that enables them to capture attention so effectively.

Could the Special Attentional Status of Multisensory Stimuli Itself Depend on Attention?

The evidence discussed above points to a paradoxical situation whereby multisensory stimuli appear to require attention before they acquire the quality that enables them to capture attention. One clue as to how this paradox might be resolved lies in the measures used to demonstrate multisensory attentional capture. As mentioned above, unisensory forms of attentional capture have been widely demonstrated using both facilitation and distraction measures. By contrast, the most convincing behavioural evidence for enhanced multisensory attentional capture involves facilitation effects. Although facilitation effects are widely used as measures of attentional capture, and also most relevant to the applied contexts discussed above, it should be noted that such effects involve stimuli which have already been allocated some top-down attention given that they are part of a search array (e.g. the ‘pip and pop’ effect). As such, it might be that this attentional allocation is sufficient to allow multisensory integration and hence heightened attentional capture. This account could explain the lack of evidence of heightened distractor interference from multisensory stimuli which are not part of the task set – such stimuli would not be allocated sufficient top-down attention to integrate, preventing their enhanced attentional status. The present work examines this possibility by using both facilitation and distraction measures to test for attentional capture by multisensory stimuli.

A second key question raised by the previous literature is whether multisensory attentional capture is, like unisensory attentional capture, dependent upon the availability of perceptual capacity. As discussed above, in unisensory contexts increasing the perceptual load of a primary task has been found to powerfully undermine processing of stimuli, whether these are irrelevant distractors (e.g. Forster &

Lavie, 2008) or search targets in a secondary task (e.g. Macdonald & Lavie, 2008). To address whether or not multisensory stimuli are, as has been suggested, immune to these load effects, the current research tested the effects of established manipulations of perceptual load on both multisensory facilitation of secondary target detection (generalising across paradigms and peripheral target salience in Experiments 1, 2 & 3), and interference by multisensory distractors (Experiment 4).

The strongest account of multisensory attentional capture - that multisensory stimuli can capture attention in a purely stimulus driven manner and are immune to any effects of perceptual load – would predict that multisensory stimuli occurring away from a primary task should produce both facilitation and distraction effects, irrespective of perceptual load of that primary task. If, on the other hand, multisensory enhancement of attentional capture is subject to some form of attentional modulation, this might manifest in two different ways (which are not mutually exclusive). If multisensory attentional capture depends on the allocation of resources regulated by top-down attention, then this would manifest only as facilitation effects and not distraction effects. If multisensory attentional capture depends on the availability of perceptual capacity, it would be eliminated altogether when perceptual load in the primary task is increased.

Experiment 1

To address the effects of perceptual load on facilitatory attentional capture by multisensory stimuli, Experiment 1 adapted an established perceptual load manipulation (e.g. Bahrami et al., 2007), in which participants search a central RSVP stream for either a single feature (colour, low load) or a conjunction (colour and shape conjunction, high load). Unlike dual versus single task comparisons (e.g. Santangelo & Spence, 2007), our manipulation of the level of load within a task allows us to isolate any influence of load from the effects of single-vs-dual task, on multisensory attentional capture. While performing the central task, participants were also asked to detect peripheral stimuli which could be either multisensory or unisensory. Facilitatory attentional capture in this paradigm would manifest as faster reaction times to multisensory versus unisensory peripheral targets. We should observe a multisensory facilitation at least for the low load task, and an effect of load for unisensory targets. The question is whether or not load will affect the responses to multisensory targets.

Materials and Methods

Participants

40 participants (26 female) aged between 18 and 35 years ($M = 23.20$, $SD = 3.68$) were recruited at the University of Sussex. All studies were approved by the University of Sussex Sciences & Technology Cross-Schools Research Ethics Committee. 20 participants completed Experiment 1a, and 20 participants completed Experiment 1b. A sample size calculation conducted using G*Power software (Faul et al., 2009) revealed that to detect an effect size of $\eta^2 = .19$ ($\alpha = .05$; $1-\beta = .80$), a sample of 18 participants was required for each Experiment 1a and 1b. The expected effect size

was taken from the main effect of cue type in Santangelo and Spence (2007) comparing multisensory and unisensory stimuli. All participants reported normal or corrected-to-normal vision and hearing. Both Bayesian and null hypothesis testing is reported given that the latter is more widely understood, but only the former provides a measure of evidence regarding whether the null or alternative hypothesis is supported by the data (Wagenmakers et al., 2017). Note that Bayesian analysis does not depend on the stopping rule and thus the measure of evidence is valid regardless of stopping rule (see Dienes, 2014; Rouder, 2014). All participants achieved over 75% average accuracy across the experiment.

Stimuli and Procedure

The experiment was programmed and presented using E-prime v2.0, on a 17-inch Dell flat screen, placed 50cm from the participants face, at eye level. Viewing distance was maintained using a chin rest. Loudspeakers, positioned left and right of the screen, were used to present sounds. Each trial began with a central fixation cross presented for 500ms, followed by a stream of nine coloured characters (each subtending $2.3^\circ \times 1.1^\circ$), presented centrally one at a time. Part of the task was to monitor a central stream of characters which were either an S or a 5, and could be coloured red, green, yellow, blue, purple or turquoise. In the high load condition, the target was either a green 5 or yellow S, whereas in low load the target was any red character. Participants reported detection of the target with a foot pedal. Targets appeared as either the 3rd or 6th stimulus in a trial. The timing of presentation was irregular, to increase demand. This was achieved by randomising presentation time of each character (167, 267 or 367ms) with a fixed interstimulus interval (ISI) of 233ms. All stimuli were presented on a light grey background.

In addition to the central task, participants were asked to monitor for peripheral targets which appeared on 50% of trials, presented to the left or right of the central stream. These were presented concurrently with a non-target central stimulus, and therefore did not interfere with responses to the central task. Participants were required to press the left button on a response box if the peripheral target was on the left, and the right button on a response box if it was on the right. Half of these peripheral targets were unisensory, and half were multisensory (i.e. each occurring on 25% of trials). In Experiment 1a, the unisensory target was a black circle of 1.7 ° diameter (visual only; 100ms), while in Experiment 1b it was a ‘beep’ sound (auditory only; 100ms, 1100Hz). In both experiments, the multisensory target was both the black circle and the ‘beep’ presented together. The unisensory peripheral targets (circle or sound) as well as the multisensory one were presented on either the left or right of the screen (in multisensory targets, the circle and ‘beep’ always occurred at the same side).

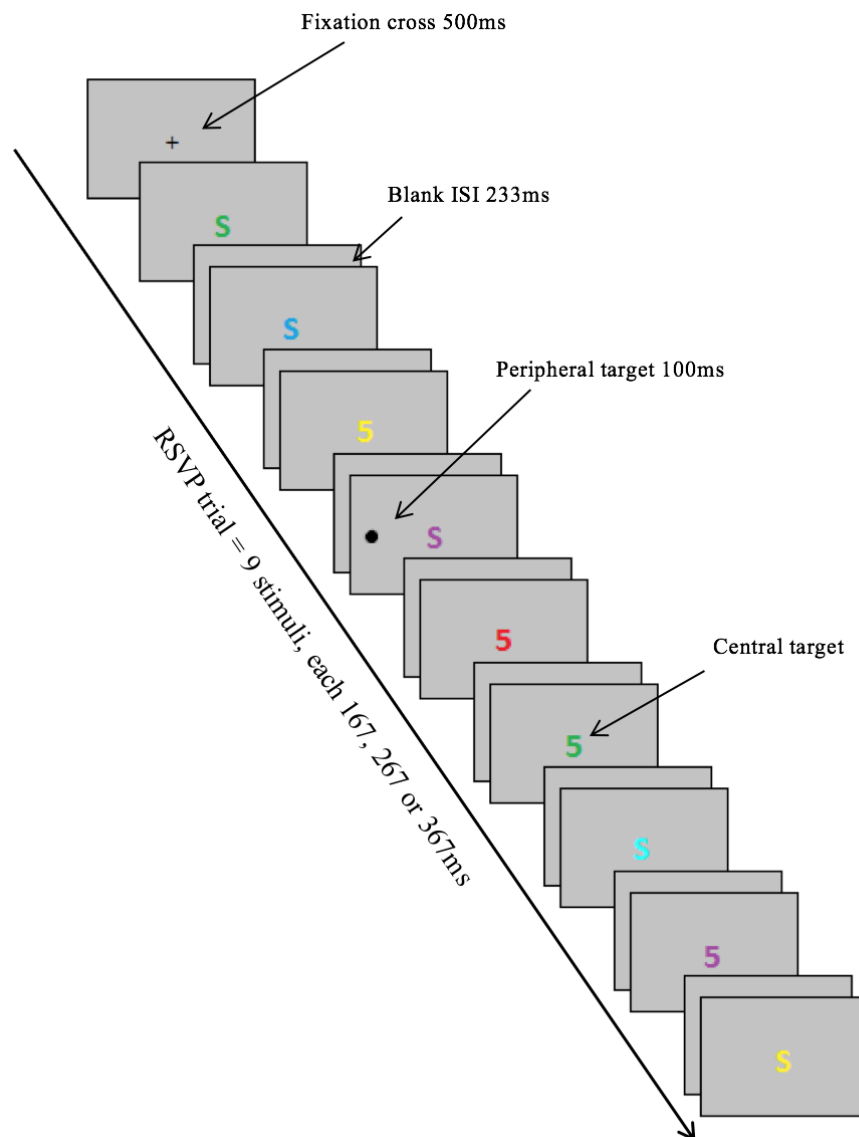


Figure 1. Example RSVP trial, with visual peripheral target

Participants completed two blocks of 144 trials for each load condition, in the order ABBA or BAAB, which was counterbalanced between participants.

Bayesian Analysis

For all tests, p -values are reported. Additionally, Bayes factors (B) are reported for all one degree of freedom tests and planned contrasts. Bayes factors (B) were used to assess the strength of the evidence for H1 relative to H0 (Wagenmakers et al., 2017). A B of above 3 is indicative of substantial evidence for H1, whereas a B of below 1/3 indicates substantial evidence for H0, and between these values indicates the data is insensitive (Dienes, 2014). Bayes factors were calculated using a half-normal distribution, as predictions were all directional, here referred to as $B_{H(0, x)}$, where x is the SD of the distribution. These SDs were based on the results found by Santangelo and Spence (2007) regarding the differences between the attentional capture effects of peripheral targets of different modalities, and Forster and Lavie (2008a) regarding load effects.

Results and Discussion

Data for all experiments can be downloaded from the Open Science Framework (osf.io/cvy8k).

Reaction time (RT)

Inter-participant average RTs to the central task (correct responses only) were significantly slower under high load (for Experiment 1a $M = 753$, $SD = 100$; for Experiment 1b $M = 749$, $SD = 105$) than under low load (for Experiment 1a $M = 541$, $SD = 81$; for Experiment 1b $M = 544$, $SD = 85$; $t(19) = 12.66$, $p < .001$, $B_{H(0,300)} = 5.77 \times 10^{33}$ for the difference in 1a; $t(19) = 16.11$, $p < .001$, $B_{H(0,300)} = 1.60 \times 10^{55}$ for the difference in 1b), reflecting the increased demands of the high load task.

Correct RTs to the peripheral targets were entered into a mixed ANOVA, with the within-subjects factors of load (low, high) and peripheral target modality (multisensory, unisensory), and the between subjects factor of experiment number (Experiment 1a and Experiment 1b). This revealed no main effect of experiment number, $p = .534$, and no interactions between experiment number and any of the within-subjects factors, $ps > .587$. In fact, an identical pattern of results was observed in both Experiments 1a and 1b: Two 2 x 2 within-subject ANOVAs with the factors of load (low, high) and peripheral target modality (multisensory, unisensory) revealed a main effect of load for both Experiment 1a, ($F(1,19) = 38.52, p < .001, \eta^2 = .67, B_{H(0,142)} = 2.40 \times 10^7$), and Experiment 1b ($F(1,19) = 39.63, p < .001, \eta^2 = .68, B_{H(0,142)} = 3.97 \times 10^7$). As can be seen in Figure 2, detection of the peripheral targets was slowed in both Experiments 1a and 1b in the high load condition relative to the low load condition, suggesting load modulation.

There was also a main effect of peripheral target modality, both in Experiment 1a ($F(1,19) = 68.90, p < .001, \eta^2 = .78, B_{H(0,21)} = 3.08 \times 10^{13}$), and Experiment 1b ($F(1,19) = 18.05, p < .001, \eta^2 = .49, B_{H(0,27)} = 1808.52$), with faster detection of multisensory targets than both unisensory visual or unisensory auditory targets. However, the critical test was the interaction. There was no significant interaction between load and target modality (Experiment 1a $p = .58$; Experiment 1b $p = .44$). Rather, detection of both multisensory and unisensory targets alike was significantly modulated by load, in both experiments (Experiment 1a $t(19) = 5.47, p < .001, B_{H(0,140)} = 3.64 \times 10^5$ for multisensory stimuli, $t(19) = 5.32, p < .001, B_{H(0,140)} = 1.85 \times 10^5$ for visual; Experiment 1b $t(19) = 5.29, p < .001, B_{H(0,140)} = 1.33 \times 10^5$ for multisensory stimuli; $t(19) = 4.92, p < .001, B_{H(0,140)} = 2.51 \times 10^4$ for auditory). On the other hand, we note that the detection speed advantage for multisensory stimuli was observed to a

similar degree in each of the load conditions, in both experiments (Experiment 1a $t(19) = 8.12, p < .001, B_{H(0,20)} = 8.28 \times 10^9$ under high load, $t(19) = 5.68, p < .001, B_{H(0,20)} = 1.34 \times 10^4$ under low; Experiment 1b $t(19) = 3.45, p = .001, B_{H(0,27)} = 98$ under high load, $t(19) = 4.49, p < .001, B_{H(0,27)} = 5582.35$ under low). Hence, multisensory stimuli did not appear immune to load effects, although their advantage over unisensory stimuli remained across low and high load.

Error

Percentage error rates in the central task were significantly higher under high load (for Experiment 1a $M = 14.80, SD = 14.88$; for Experiment 1b $M = 10.20, SD = 12.01$) than under low load (for Experiment 1a $M = 5.30, SD = 6.32$; for Experiment 1b $M = 5.70, SD = 10.31$; $t(19) = 3.52, p = .001, B_{H(0,10)} = 166.95$ for the difference in 1a; $t(19) = 2.18, p = .021, B_{H(0,10)} = 3.91$ for the difference in 1b).

Error rates in detection of the peripheral targets were entered into a mixed ANOVA, with the within-subjects factors of load (low, high) and peripheral target modality (multisensory, unisensory), and the between subjects factor of experiment number (Experiment 1a and Experiment 1b). This revealed no main effect of experiment number, $p = .810$, and no interactions between experiment number and any of the within-subjects factors, $ps > .164$. In fact, an identical pattern of results was observed in both Experiments 1a and 1b: Two 2 x 2 within-subject ANOVAs with the factors of load (low, high) and peripheral target modality (multisensory, unisensory) on error rates to peripheral targets (Table 1) revealed no significant effect of load, or interaction between load and peripheral target modality, for Experiment 1a or 1b (1a $ps > .085$; 1b $ps > .100$). In Experiment 1a there was also no main effect of peripheral stimulus modality found ($p = .506, B_{H(0,10)} = 0.20$), however in Experiment 1b

percentage error rates for multisensory stimuli were lower than those for auditory ($F(1,19) = 13.71, p = .002, \eta^2 = .42, B_{H(0,10)} = 18.46$). These results show that the RT effects were not due to a speed accuracy trade-off. Error rates were generally very low, thus the advantageous nature of multisensory stimuli is reflected mostly in RTs.

Overall, this experiment demonstrated two key findings. First, our results are consistent with existing evidence of enhancement of attentional capture by multisensory stimuli which are part of the top-down attentional set. Furthermore, consistent with suggestions regarding the applied utility of multisensory cues during demanding tasks, the multisensory advantage over unisensory stimuli remained regardless of load. However, contrary to previous suggestions, multisensory stimuli did not appear entirely immune to load effects, in that increasing perceptual load in a central task slowed detection of peripheral multisensory targets as much as unisensory ones. As such, processing of multisensory and unisensory stimuli alike appears modulated by the availability of attentional capacity.

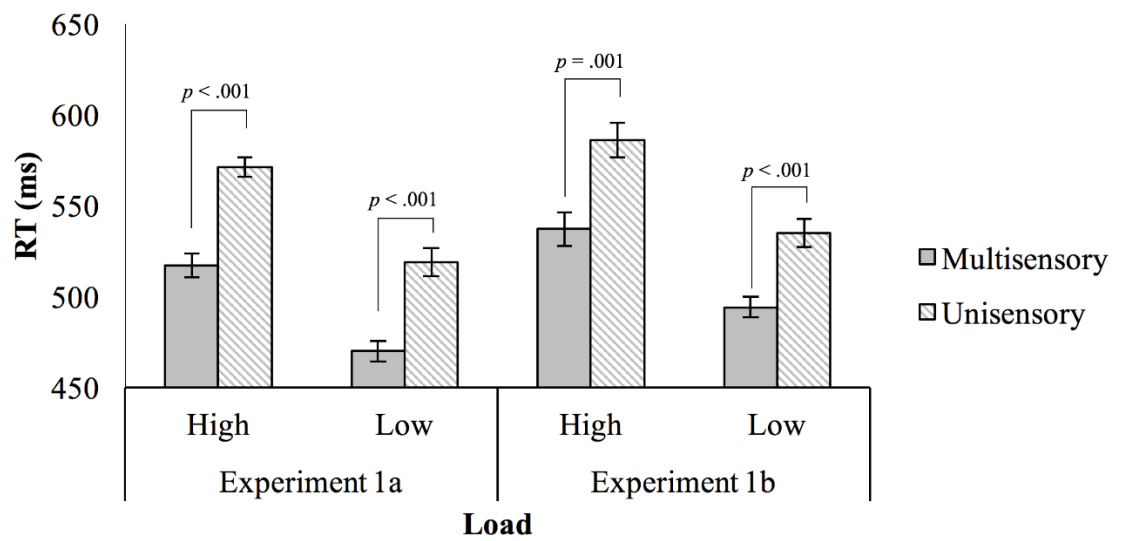


Figure 2. RT (ms) for detection of multisensory and visual only peripheral targets, as a function of load in Experiment 1a, and multisensory and auditory only peripheral targets as a function of load in experiment 1b, error bars show ± 1 SEM with Cousineau-Morey correction (Cousineau, 2005; Morey, 2008)

		Multisensory	Visual Only	Auditory Only
Experiment 1a	Low Load	6.50 (5.92)	5.90 (6.07)	
	High Load	7.85 (6.33)	9.90 (8.21)	
Experiment 1b	Low Load	5.35 (5.30)		8.00 (5.80)
	High Load	8.05 (7.35)		10.45 (10.66)
Experiment 2	Low Load	4.25 (6.98)	5.38 (7.21)	10.04 (9.25)
	High Load	3.88 (4.30)	6.42 (4.75)	9.46 (8.27)
Experiment 3	Low Load	2.57 (3.65)	5.71 (6.19)	6.32 (6.64)
	High Load	2.14 (3.35)	6.71 (6.66)	7.07 (6.08)

Table 1. Mean percentage error rates (SD in parentheses) as a function of load and target type, across experiments 1-3

Experiment 2

The results of Experiment 1 appear initially consistent with claims that multisensory stimuli can capture attention. These claims derive from evidence that multisensory stimuli elicit super-additive responses. However, data from Experiment 1 could not be tested for non-linear effects, given that the between-subjects design did not allow the application of the usual modelling benchmarks (e.g., race model). In the following experiment, we sought to test whether the response to a multisensory stimulus was greater than that which would be predicted by the summed probability of the two unisensory stimuli by testing violations of the race model, which would suggest neural integration of the two sensory stimuli (Miller, 1982, 1986). In order to be able to calculate race model, auditory and visual unisensory targets were tested within the same experiment, rather than between tasks as in Experiment 1. In addition, Experiment 2 sought to replicate and generalize the findings of Experiment 1 regarding perceptual load effects on multisensory stimuli to another well-established visual search load manipulation in the central task (e.g. Forster & Lavie, 2008).

Materials and Methods

Participants

26 participants (22 female) aged between 18 and 35 years ($M = 20.31$, $SD = 2.62$) were recruited at the University of Sussex. Two participants were excluded for failing to comply with the instructions. All participants reported normal or corrected-to-normal vision and hearing. The apriori stopping rule for this experiment, and all subsequent experiments, was based on Bayes Factors for the main effect of load and peripheral target type, and all planned comparisons, on reaction time data reaching

sensitivity (see Rouder, 2014)¹. All participants achieved over 75% average accuracy across the experiment.

Stimuli and procedure

The experiment was programmed and presented using E-Prime v2.0, on a 17inch screen. A viewing distance of 57cm was maintained using a chin rest. Loudspeakers positioned on the left and the right side of the screen were used to deliver sounds. Each trial began with a central fixation dot presented for 500ms, followed by a 100ms stimulus display. The stimulus display consisted of six letters (each subtending $0.7^\circ \times 0.8^\circ$) evenly arranged in an imaginary circle (2.0° radius). On 50% of trials, one of the letters was the target letter which participants were required to search for (X).

Participants were required to indicate detection of the target letter by pressing the space bar. In the high load condition, the non-target letters were pseudo-randomly selected from a set of angular letters (H, K, M, V, W, Z, N), whereas in the low load condition the non target letters were all small, placeholder O's (diameter 0.2°). All stimuli were presented on a black background, and all letters were white. The sizing of the stimuli, and the display, was based on previous use of this visual search task (Forster & Lavie, 2007, 2008a, 2008b, 2009, 2014).

On 22.5% of the trials, a peripheral target was presented to either the left or the right of the circular array of letters. On these trials, participants were required to indicate which side of the screen the target was presented on, by pressing one of two keys. These targets could be unisensory visual, unisensory auditory, or multisensory

¹ This stopping rule was adopted in line with a general change of practice in the lab, in order to determine that any null differences reflect a true no difference between conditions, and are not due to a lack of sensitivity within the data.

with equal probability. Peripheral targets were the same as those used in experiments 1a and 1b, and presented for 100ms, with onset at the same time as the central task (see Fig 3). They could not appear in the first three trials of each block. All targets were presented to every participant, with the load, central target position, central target identity, peripheral target side, and peripheral target type, fully randomised. Each trial could contain a peripheral target or a central target but not both, or no targets at all. In the latter case (27.5% of trials), no response was required from participants. This prevented the participants from inferring that if there is no peripheral target, that there must be a central one.

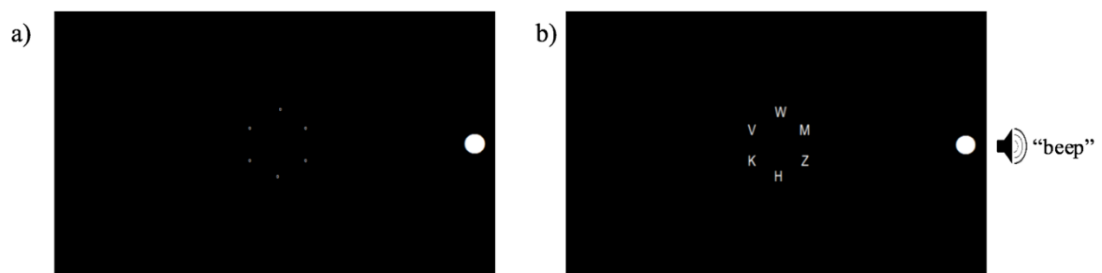


Figure 3. Example stimulus displays: a) low load central letter array with small placeholder letters, with visual peripheral target b) high load central letter array, with multisensory peripheral target

Participants completed three slowed down example trials, followed by 12 practice trials, for both high and low load. They then completed four blocks of 80 trials for each load condition, in the order ABBAABBA or BAABBAAB, which was counterbalanced between participants. Participants were instructed to respond as quickly as possible whilst still being accurate (2000ms response deadline).

Bayesian analysis

Bayes factors were calculated using a half-normal distribution, with SDs based on Experiment 1 of this paper.

Results and Discussion

Reaction time (RT)

Inter-participant average RTs to the central task (correct responses only) were significantly slower under high load ($M = 634$, $SD = 117$) than under low load ($M = 496$, $SD = 82$), $t(23) = 8.18$, $p < .001$, $B_{H(0,212)} = 4.7 \times 10^9$, indicating that the high load task was more demanding.

Correct responses to peripheral targets were entered into a 2 x 3 within-subject ANOVA with the factors of load (low, high) and peripheral stimulus modality (multisensory, unisensory visual, unisensory auditory; Figure 5). As in Experiment 1, the results for peripheral targets revealed main effects of load, $F(1,23) = 23.94$, $p < .001$, $\eta^2 = .51$, $B_{H(0,46)} = 2.36 \times 10^9$, and stimulus modality, $F(2, 46) = 106.45$, $p < .001$, $\eta^2 = .82$, the latter reflecting faster RTs to multisensory peripheral targets compared to either visual only ($B_{H(0,52)} = 6.77 \times 10^6$) or auditory only ($B_{H(0,44)} = 8.68 \times 10^{53}$).

In contrast to the previous experiment, these main effects were qualified by a significant interaction between load and peripheral stimulus modality, $F(2,46) = 13.00$, $p < .001$, $\eta^2 = .36$). This reflected that responses to multisensory targets were modulated by load to a lesser extent than responses to visual targets ($t(23) = 3.90$, $p < .001$, $B_{H(0,120)} = 580.84$), but to an equivalent extent to auditory targets ($p = .940$, $B_{H(0,146)} = .07$). As in Experiment 1, detection of peripheral targets was slower under high versus low load conditions regardless of sensory modality (multisensory ($t(23) = 3.54$, $p < .001$, $B_{H(0,82)}$

= 176.73), visual only ($t(23) = 8.88, p < .001, B_{H(0,52)} = 9.83 \times 10^{15}$ and, auditory only ($t(23) = 1.88, p = .036, B_{H(0,52)} = 3.01$). A detection speed advantage for multisensory targets over both types of unisensory target was also observed under conditions of both high load ($t(23) = 10.64, p < .001, B_{H(0,40)} = 1.05 \times 10^{22}$; $t(23) = 6.24, p < .001, B_{H(0,54)} = 5.10 \times 10^7$, for multisensory compared with auditory and visual, respectively) and low load ($t(23) = 12.28, p < .001, B_{H(0,50)} = 2.08 \times 10^{30}$; $t(23) = 2.56, p = .009, B_{H(0,50)} = 7.83$, for multisensory compared with auditory and visual, respectively).

In order to further determine whether multisensory stimuli provide a detection speed advantage consistent with integration, we used the race model (Miller's inequality, Miller, 1982). The race model allows us to investigate whether the reaction times in the multisensory condition exceed the statistical facilitation predicted by probability summation based on two independent unisensory signals. In this model, a theoretical cumulative density function (CDF) is calculated based on the reaction time CDFs of each of the two unimodal stimulus types - F_x and F_y - and the redundant-stimulus, or multisensory, condition, F_z . The race model inequality

$$F_z(t) \leq F_x(t) + F_y(t), t > 0,$$

is examined for every value of t . Where the empirical CDF towards multisensory stimuli is greater than the theoretical CDFs based on the two unisensory components (tested using paired sample t-tests), the reaction time advantage can be assumed to be caused by integrative effects. Analyses were carried out using the RMITest software, which applies the algorithm in Ulrich, Miller and Schröter (2007).

The results showed that under low load, whilst the reaction time towards multisensory stimuli tends to be faster than the race model bound for the lowest (fastest) percentiles of the reaction time distribution, this does not reach statistical significance (p

> .05). However, under high load, reaction time to multisensory stimuli is significantly faster than the race model bound for three of the fastest percentiles (Figure 4). This supports the assumption that the detection time advantage under high load might result from crossmodal integration.

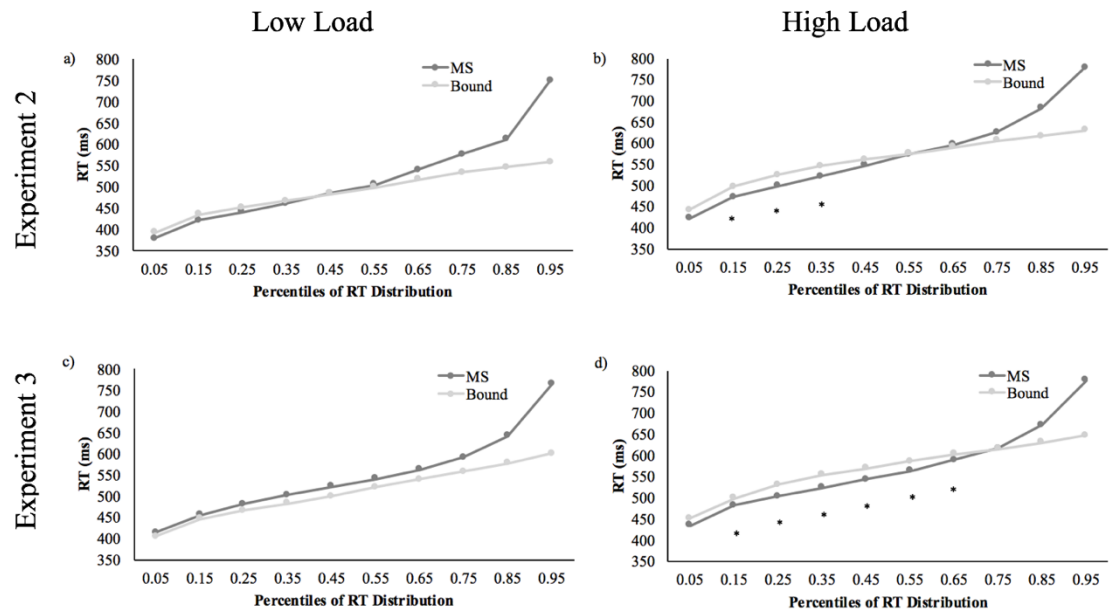


Figure 4. Cumulative probability distribution of reaction times for detection of multisensory (MS) peripheral stimuli, with race model bound for the two unisensory peripheral stimuli predicted by RMITest for a) Experiment 2 low central perceptual load, b) Experiment 2 high central perceptual load, c) Experiment 3 low central perceptual load, d) Experiment 3 high central perceptual load, asterisks refer to where race model inequality was significantly violated, based on Ulrich et al., (2007) algorithm

Error

Percentage error rates in the central task were significantly higher under high load ($M = 11.75$, $SD = 10.60$) than under low load ($M = 6.42$, $SD = 3.27$), $t(23) = 2.41$, $p = .012$, $B_{H(0,7)} = 8.85$.

Percentage error rates in peripheral stimuli detection were then entered into a 2 x 3 within-subject ANOVA with the factors of load (low, high) and peripheral stimulus modality (multisensory, unisensory visual, unisensory auditory), revealing a main effect of peripheral stimulus modality, $F(1.47, 33.77) = 15.32$, $p < .001$, $\eta^2 = .40$; Table 1. Mauchly's test indicated that the assumption of sphericity had been violated ($c^2(2) = .64$, $p = .007$, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($e = .73$). However, there was no main effect of load ($p = .979$, $B_{H(0,7)} = 0.13$), nor a significant interaction between load and peripheral stimulus modality, ($p = .668$).

Under conditions of high perceptual load, error rates for multisensory stimuli ($M = 3.88$, $SD = 4.30$) were significantly lower than those for visual stimuli ($M = 6.42$, $SD = 4.75$), $t(23) = 2.88$, $p = .004$, $B_{H(0,3)} = 26.00$, and for auditory stimuli ($M = 9.46$, $SD = 8.27$), $t(23) = 3.15$, $p = .002$, $B_{H(0,3)} = 40.41$. This effect was also seen under conditions of low perceptual load for auditory stimuli; error rates for multisensory stimuli ($M = 4.25$, $SD = 6.98$) were significantly lower than those for auditory stimuli ($M = 10.04$, $SD = 9.25$), $t(23) = 3.39$, $p < .001$, $B_{H(0,3)} = 74.84$, however there was no significant difference in error rates for detection of multisensory and visual peripheral targets ($M = 5.38$, $SD = 7.21$), ($p = .167$, $B_{H(0,3)} = .93$) under low load

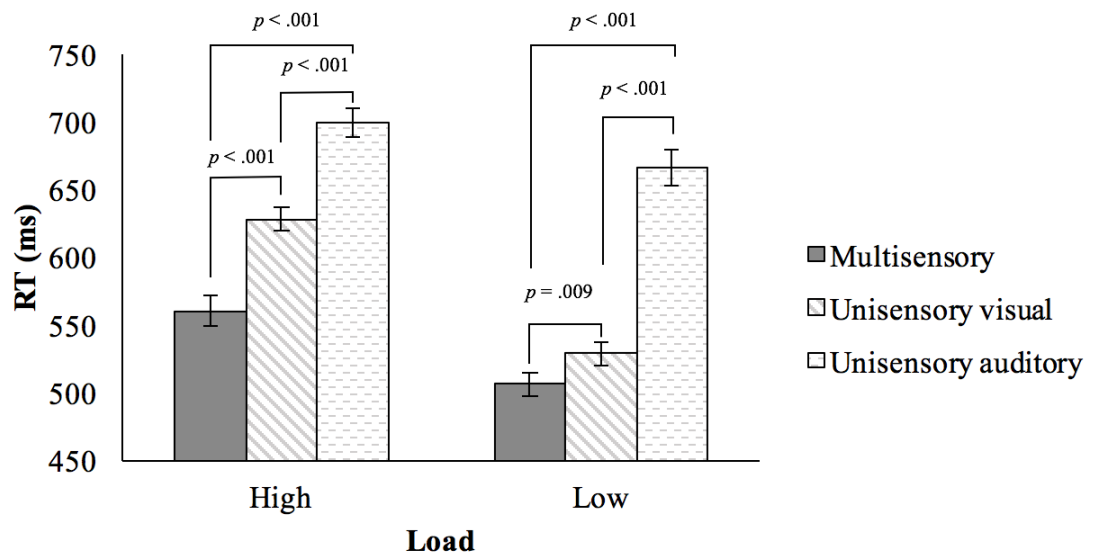


Figure 5. Reaction time (ms) of detection of peripheral targets in different modalities, as a function of load, in Experiment 2. Error bars show ± 1 SEM with Cousineau-Morey correction (Cousineau, 2005; Morey, 2008)

In summary, Experiment 2 replicates the two key findings of Experiment 1: multisensory stimuli are affected by load modulations, and multisensory stimuli show a detection advantage above unisensory events which is preserved across load manipulations. Again, these results are consistent with an enhancement of attentional capture by stimuli which have been allocated some degree of top-down attention. Unlike Experiment 1, Experiment 2 also found that the reaction time advantage for multisensory stimuli was more pronounced under the high load condition versus the low load condition, with RTs demonstrating that multisensory stimuli were modulated by load to a lesser extent than visual stimuli, and the race model only producing evidence consistent with integration in the high load condition. This raises the intriguing

possibility that, while not entirely immune to any effect of perceptual load, multisensory stimuli might be somewhat more resistant to these effects than unisensory events.

Experiment 3

Experiments 1 and 2 suggest an advantage for multisensory stimuli over unisensory stimuli. We note, however, that the unisensory stimuli used in these experiments were of relatively low salience, which may have encouraged the multisensory advantage: Following the principle of inverse effectiveness (e.g. Meredith & Stein, 1986), initially proposed for single neuron responses, the weaker the responses to the individual unisensory stimuli, the more likely super-additive responses are to occur if they are presented together as a multisensory event. A remaining question is therefore whether facilitation of attentional capture by multisensory stimuli is limited to low salience stimuli. This appears particularly important given applied suggestions regarding the use of multisensory alerts: in a real world scenario, a unisensory stimulus that is hard to detect would not be reasonably used as an alert or warning signal. Here we addressed whether multisensory stimuli would be capable of facilitating detection above and beyond stimuli highly salient in one unisensory domain, instead of the low salience events used in Experiments 1 and 2. To test this, Experiment 3 repeated the paradigm of Experiment 2 using peripheral target stimuli that are larger, more colourful, meaningful and familiar, so that the visual unisensory stimuli would be highly salient relevant to the central task.

Materials and Methods

Participants

28 participants (22 female) aged between 18 and 33 years ($M = 21.04$, $SD = 2.92$) were recruited at the University of Sussex. Participants were recruited until all Bayes Factors for the main effects of load and peripheral target type on reaction time

data reached sensitivity. All participants reported normal or corrected-to-normal vision and hearing. Participants either gained course credits, or were paid, to take part. All participants achieved over 75% average accuracy across the experiment.

Stimuli and procedure

The stimuli and procedure were identical to Experiment 2, with the exception of the identity of the peripheral target stimuli (see Figure 7). Visual peripheral targets consisted of a photograph of an animal, randomly selected from six possible images (dog, cat, pig, horse, cow, sheep). These were presented in full colour with a black background, subtending 5.0° - 7.5° vertically, by 6.0° - 7.0° horizontally, between 2.5° and 3.0° edge-to-edge from the nearest circle letter. The auditory peripheral targets consisted of the sound each of the six animals makes, played from one of the speakers at the side of the screen (600-1120ms). The multisensory targets were both the animal image and sound, presented on the same side.

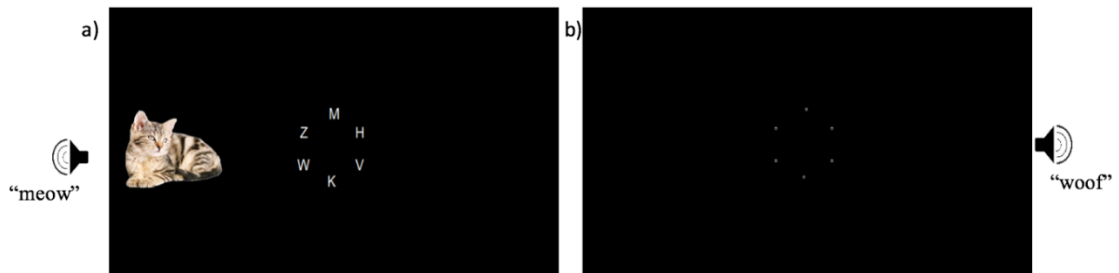


Figure 7. Example stimulus displays: a) high load central letter array, multisensory peripheral target, b) low load central array, auditory peripheral target

Participants completed three slowed down example trials, followed by 12 practice trials, for both high and low load. They then completed four blocks of 81 trials

for each load, in the order ABBAABBA or BAABBAAB, which was counterbalanced between participants. Participants were instructed to respond as quickly as possible whilst still being accurate. They had 2000ms to make a response, a short beep indicated where this had been incorrect.

Bayesian analysis

Bayes factors were calculated using a half-normal distribution, with SDs based on Experiment 2 of this paper.

Results and Discussion

Reaction time (RT)

RTs to the central task (correct responses only) were significantly slower under high load ($M = 635.60$, $SD = 69.91$) than under low load ($M = 506.40$, $SD = 61.28$), $t(27) = 12.72$, $p < .001$, $B_{H(0,138)} = 1.22 \times 10^{34}$, indicating that the high load task was more demanding.

Correct responses to peripheral targets were entered into a 2 x 3 within-subject ANOVA with the factors of load (low, high) and peripheral stimulus modality (multisensory, unisensory visual, unisensory auditory; Figure 6). As in previous experiments, this revealed main effects of load, $F(1,27) = 30.59$, $p < .001$, $\eta^2 = .53$, $B_{H(0,62)} = 6.47 \times 10^9$, and peripheral stimulus modality, $F(1.50, 40.57) = 317.22$, $p < .001$, $\eta^2 = .92$ (Mauchly's test indicated that the assumption of sphericity had been violated ($\epsilon^2(2) = .67$, $p = .005$, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity, $e = .75$) with RTs to multisensory peripheral targets being faster than to either visual only ($t(27) = 4.72$, $p < .001$, $B_{H(0,46)} = 1.55 \times 10^4$) or auditory only ($t(27) = 24.25$, $p < .001$, $B_{H(0,148)} = 2.69 \times 10^{126}$). As in Experiment 2, this

was qualified by a significant interaction between load and peripheral stimulus modality, $F(2,54) = 5.45$, $p = .007$, $\eta^2 = .17$. Responses to multisensory targets were modulated by load to a lesser extent than responses to visual targets ($t(27) = 2.96$, $p = .003$, $B_{H(0,45)} = 30.50$), but not less than auditory targets ($p = .552$, $B_{H(0,45)} = 0.26$).

As in both previous experiments, detection of all three types of peripheral target was slower under high versus low load (multisensory ($t(27) = 4.02$, $p < .001$, $B_{H(0,52)} = 829.07$), visual only ($t(27) = 5.69$, $p < .001$, $B_{H(0,100)} = 2.07 \times 10^6$) and auditory only ($t(27) = 2.80$, $p = .005$, $B_{H(0,32)} = 21.39$). The detection speed advantage for multisensory targets over auditory only was observed under both high ($t(27) = 15.18$, $p < .001$, $B_{H(0,138)} = 7.32 \times 10^{48}$) and low ($t(27) = 19.62$, $p < .001$, $B_{H(0,160)} = 2.54 \times 10^{82}$) load conditions. The advantage over visual only targets was significant under high load ($t(27) = 5.78$, $p < .001$, $B_{H(0,68)} = 3.23 \times 10^6$), however no sensitive evidence was obtained under low load ($p = .082$, $B_{H(0,22)} = 1.54$).

RMITest software (Ulrich et al., 2007) was again employed to test for violation of the race model inequality. Similar to Experiment 2, the results showed that under low load there were no significant violations of the race model, whereas under high load reaction time to multisensory stimuli was significantly faster than that which would be predicted by the race model across most of the fastest percentiles, indicating again that integration is occurring and resulting in the faster detection times (Figure 4).

Error

Percentage error rates in the central task were significantly higher under high load ($M = 10.38$, $SD = 8.55$) than under low load ($M = 7.00$, $SD = 4.12$), $t(27) = 2.67$, $p = .006$, $B_{H(0,6)} = 22.31$.

A 2 x 3 within-subject ANOVA with the factors of load (low, high) and peripheral stimulus modality (multisensory, unisensory visual, unisensory auditory) on percentage error rates of peripheral stimuli detection was conducted (Table 1). As in Experiment 2 there was a main effect of peripheral stimulus modality, $F(1.64, 44.33) = 13.10$, $p < .001$, $\eta^2 = .33$. Mauchly's test indicated that the assumption of sphericity had been violated ($\epsilon^2(2) = .78$, $p = .041$, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\epsilon = .82$). There was no main effect of load found ($p = .680$, $B_{H(0,6)} = 0.22$), nor a significant interaction between load and peripheral stimulus modality, ($p = .702$).

Under conditions of high perceptual load, error rates for multisensory stimuli ($M = 2.14$, $SD = 3.35$) were significantly lower than those for visual stimuli ($M = 6.71$, $SD = 6.66$), $t(27) = 3.79$, $p < .001$, $B_{H(0,3)} = 345.15$, and for auditory stimuli ($M = 7.07$, $SD = 6.08$), $t(27) = 4.22$, $p < .001$, $B_{H(0,6)} = 1982.36$. These effects were also seen under conditions of low perceptual load; error rates for multisensory stimuli ($M = 2.57$, $SD = 3.65$) were significantly lower than those for visual stimuli ($M = 5.17$, $SD = 6.19$), $t(27) = 2.78$, $p = .005$, $B_{H(0,3)} = 8.81$, and for auditory stimuli ($M = 6.32$, $SD = 6.64$), $t(27) = 2.73$, $p = .006$, $B_{H(0,6)} = 15.61$.

Thus, this result further indicates an advantage for detection of multisensory stimuli. Even when engaged in a perceptually demanding central task, participants were able to detect the spatial location of a multisensory target with more accuracy than either visual or auditory alone.

In summary, the results of Experiment 3 demonstrate that multisensory stimuli can enhance facilitatory attentional capture even for stimuli that are already highly salient. Consistent with the previous experiments, the results do not support the

strongest claim of multisensory immunity to effects of perceptual load insofar as the detection of multisensory targets was slowed down under high load. However, as in Experiment 2, both multisensory and auditory targets were modulated by load to a lesser extent than visual unisensory targets. Furthermore, also as in Experiment 2, the race model analysis only showed sensitive evidence of an integration-facilitated detection advantage in the high load condition. Taken together these findings support the notion that multisensory integration can lead to benefits in detecting searched for stimuli, which may be particularly apparent during more demanding tasks.

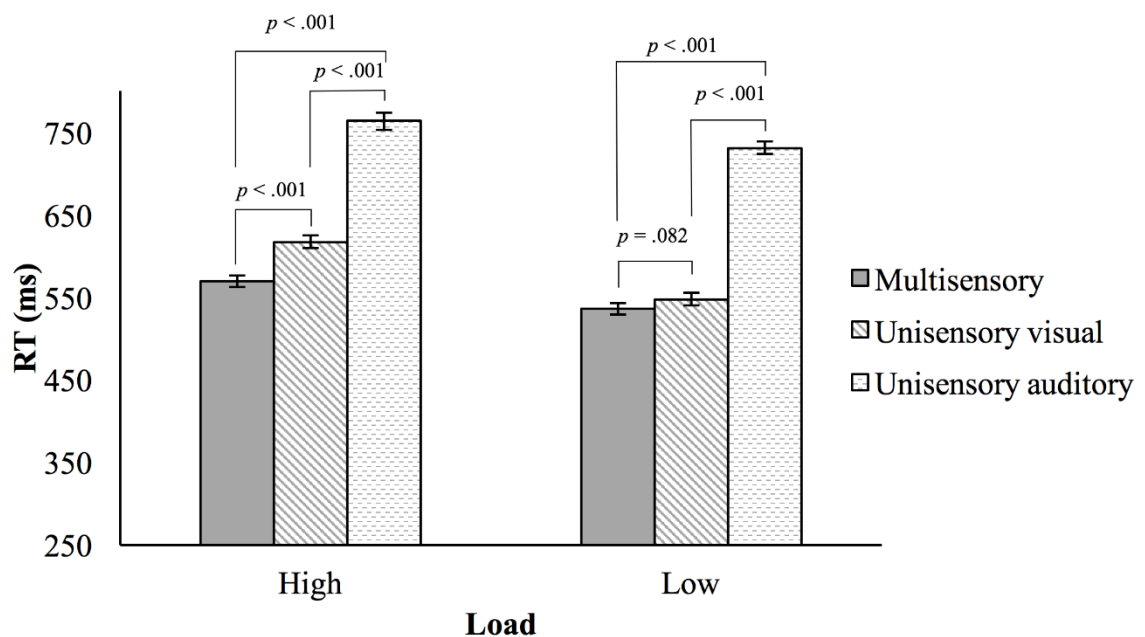


Figure 6. Reaction time (ms) of detection of peripheral targets in different modalities, as a function of load, in Experiment 3. Error bars show +/- 1 SEM with Cousineau-Morey correction (Cousineau, 2005; Morey, 2008)

Experiment 4

Experiments 1-3 demonstrate facilitatory attentional capture by multisensory stimuli, in terms of both faster (Experiments 1-3) and more accurate (Experiments 2-3) detection of peripheral targets (relative to unisensory stimuli). This is consistent with the notion of multisensory stimuli having a special attentional status in terms of facilitated detection, although they do not appear to be entirely immune to the effects of perceptual load. Experiment 4 was designed to test whether attentional capture by multisensory stimuli would extend beyond facilitation effects, which necessarily involve the allocation of some top-down attention to the locations in which the multisensory stimuli appear. To test whether these effects extend to task irrelevant distractor stimuli, we adapted the protocol used in Experiment 3 to an ‘Irrelevant Distractor Task’ (Forster & Lavie, 2008a) which has been previously established in the unisensory visual domain. The paradigm was similar to Experiment 3 with the exception that, as in the Irrelevant Distractor Task, participants were instructed to ignore the peripheral stimuli rather than respond to them. Now, capture by the peripheral targets should be inferred from their capacity to slow-down responses to the central task events (i.e. distractor interference). To maintain the low frequency of the distractors, which is necessary to observe a strong irrelevant distractor effect (Forster & Lavie, 2008b), whilst also maintaining an adequate number of trials in each condition,

multisensory distractors were compared here with unisensory visual distractors (as used in the original version of the Irrelevant Distractor Task)².

Materials and Methods

Participants

52 participants (39 female) aged between 18 and 27 years ($M = 20.26$, $SD = 1.94$) were recruited at the University of Sussex. Seven participants were excluded for failing to reach an average of over 75% accuracy across the experiment. Participants were recruited until Bayes Factors for the main effect of load, and all planned distractor cost comparisons on reaction time, reached sensitivity. Participants either gained course credits, or were paid, to take part. All participants reported normal or corrected-to-normal vision and hearing.

Stimuli and procedure

Stimuli and procedure are identical to Experiment 3, with the exception that participants were instructed to ignore anything in the periphery which may distract them from their task, that no auditory-only distractors were presented, and that there were two potential central targets (as in e.g. Forster & Lavie, 2008a). Participants were instructed to search for either an X or an N in both high and low load conditions. On 16% of the trials, a distractor was presented, half being visual only and half being multisensory.

² Note that the fact that multisensory distractor stimuli do not, by their nature, require a response precluded a race model analysis in this experiment. In all of the present paper's previous experiments, visual stimuli produced greater capture effects than auditory stimuli and hence appeared the most competitive control condition. Had we found evidence of any multisensory enhancement of irrelevant distraction we would have proceeded to conduct a second experiment using unisensory auditory distractors – however, in the absence of any such effect this further experiment was not necessary (as any true multisensory benefit would be found in comparison to both visual and auditory distractors).

They could not appear in the first three trials of each block. All distractors were presented to every participant, with the load, target position, target identity, distractor side, distractor type and distractor identity, fully randomised.

Participants completed three slowed down example trials, followed by 12 practice trials, for both high and low load. They then completed four blocks of 80 trials for each load, in the order ABBAABBA or BAABBAAB, which was counterbalanced between participants. Participants were instructed to respond as quickly as possible whilst still being accurate, and told to ignore anything else which may be presented to them other than the circle of letters. They had 2000ms to make a response, a short beep indicated where this had been incorrect.

Bayesian analysis

Bayes factors were calculated using a half-normal distribution, with SDs based on Forster and Lavie's (2008a) irrelevant distractor study using the same paradigm.

Results and Discussion

Mean RTs to correct responses and percentage error rates, as a function of distractor condition and load, are displayed in Table 2.

Reaction time (RT)

A 2 x 3 within-subject ANOVA with the factors of load (low, high) and distractor (multisensory, unisensory, no distractor), revealed a main effect of load, $F(1, 45) = 261.89, p < .001, \eta^2 = .86, B_{H(0,176)} = 4.90 \times 10^{171}$. RTs were slower in the high load than the low load condition, reflecting the increased demands of the high load task. There was no main effect of distractor ($p = .194$), however there was a significant

interaction between load and distractor, $F(1.63, 71.67) = 4.83$, $p = .016$, $\eta^2 = .10$.

Mauchly's test indicated that the assumption of sphericity had been violated ($\epsilon^2(2) = .77$, $p = .004$, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\epsilon = .81$). As can be seen in Figure 8, this interaction reflected that both multisensory and unisensory distractors slowed down RTs, relative to the no distractor condition, in the low load condition only.

	Distractor Condition		
	Multisensory	Visual	No Distractor
Low Load			
RT(ms)	535 (13)	530 (10)	512 (9)
% Error	9	9	9
High Load			
RT(ms)	792 (19)	783 (18)	795 (18)
% Error	25	23	21

Table 2. Mean RTs (SE in parentheses) and error rates (%) as a function of load and distractor type

Planned comparisons revealed sensitive evidence for interference from both distractor types under low load: RT was significantly slower in the presence of a multisensory distractor than when no distractor was present, $t(44) = 3.36$, $p < .001$, $B_{H(0,60)} = 59.64$, and in the presence of a unisensory distractor than when no distractor was present, $t(44) = 4.25$, $p < .001$, $B_{H(0,60)} = 1185.32$. By contrast, under high load, comparison of RTs in the presence of either multisensory or visual only distractors compared with no distractor revealed sensitive null effects ($p = .623$, $B_{H(0,60)} = .12$ and $p = .932$, $B_{H(0,60)} = .05$ respectively). Critically, RTs in the presence of a multisensory

distractor did not differ from RTs in the presence of a visual only distractor, for either high load ($p = .932$, $B_{H(0,60)} = .40$) or low load ($p = .228$, $B_{H(0,60)} = .27$).

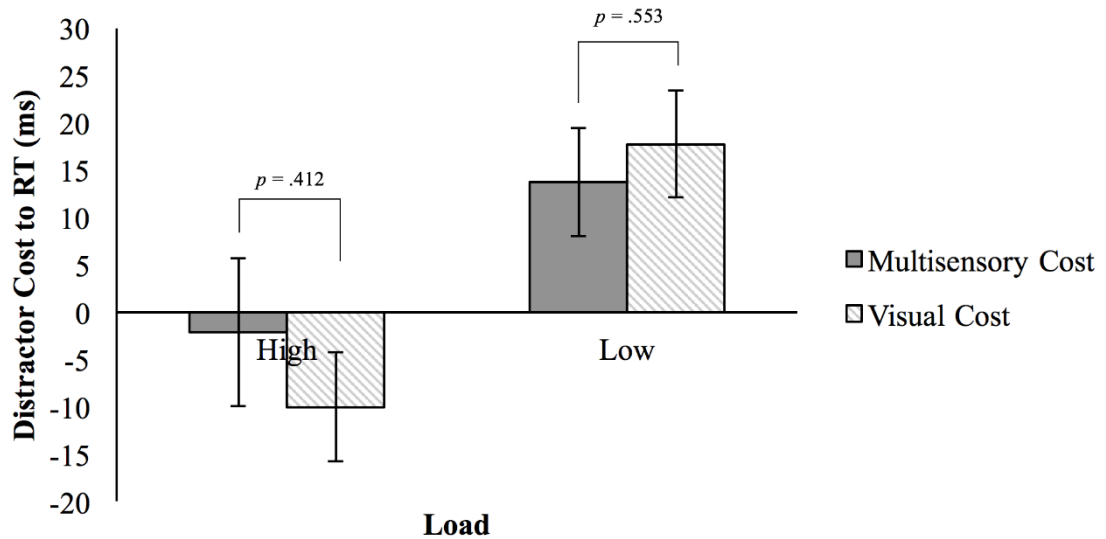


Figure 8. Cost to reaction time (ms) for detection of central target, due to the presence of a multisensory distract or visual distractor, as a function of load, error bars show +/- 1 SEM with Cousineau-Morey correction (Cousineau, 2005; Morey, 2008)

Error

A 2 x 3 within-subject ANOVA with the factors of load (low, high) and distractor type (multisensory, unisensory visual, no distractor), revealed a main effect of load, $F(1,44) = 78.96$, $p < .001$, $\eta^2 = .64$, $B_{H(0,4)} = 1.64 \times 10^{34}$. Error rates were lower in the low load than the high load condition. There was no main effect of distractor type found ($p = .084$), or interaction between load and distractor type ($p = .163$).

The present results for visual distractors replicate previous findings using this paradigm (Forster & Lavie, 2008a, 2014, 2016), as the effect of distractor presence on reaction time to the visual search task, present under low load conditions, was eliminated under high perceptual load. However, this pattern was seen not only for visual only distractors (the canonical effect) but also when the distractors were multisensory, demonstrating that their effects are not immune to perceptual load. Furthermore, unlike Experiments 1-3, this Experiment produced a striking absence of any evidence for enhanced attentional capture by multisensory stimuli above and beyond unisensory stimuli in any condition of load – this being a sensitive null result under low load, falling just short of sensitivity under high load.

General Discussion

The present study sought to determine to what extent multisensory stimuli may be particularly effective at capturing attention, under both high and low perceptual load conditions. We have used, for the first time, established and controlled manipulations of perceptual load comparable to the ones traditionally used in unisensory perceptual load studies. The first key finding from this study is that across three experiments (Experiments 1-3), involving both high and low salience distractors, we demonstrate clear evidence of facilitatory attentional capture by multisensory events, in terms of both faster and more accurate detection compared to unisensory stimuli. This replicates a well-known multisensory advantage, previously reported in many studies and different paradigms (e.g. Colonius & Diederich, 2004; Frassinetti et al., 2002; Pannunzi et al., 2014; Pérez-Bellido et al., 2012). In contrast, however, we did not find any evidence of greater distractor interference by multisensory stimuli (Experiment 4). Note that Experiments 3 and 4 involved the same high salience peripheral stimuli - the key difference being whether participants were instructed to attend (and respond to) the peripheral stimuli (Experiment 3), or to ignore them (Experiment 4) – yet multisensory stimuli produced enhanced attentional capture in the former, but not latter, case. This is consistent with suggestions that multisensory integration is compromised when the stimuli involved are not attended (e.g. Alsius et al., 2005; Talsma et al., 2007). In other words, multisensory integration, and hence enhanced attentional capture by multisensory events, may only occur for things which we are already looking out for.

The second key finding to emerge from the present study is that across all four experiments, and across two different manipulations of perceptual load, multisensory stimuli were not strictly ‘immune’ to perceptual load effects as has been previously proposed (Santangelo & Spence, 2007). Using a controlled, standard, manipulation of

perceptual load, our Experiments 1-3 show that, similar to unisensory peripheral events, RT to multisensory peripheral targets was slowed down when the central task was high in perceptual load, and in Experiment 4 multisensory distractor costs were reduced under these conditions just like distractors costs associated with unisensory (visual) distractors. Hence, the processing of multisensory stimuli was modulated by load.

On the other hand, our findings are broadly compatible with the suggestion of Santangelo and Spence that multisensory stimuli might be particularly useful (e.g. as alerts) during high load tasks. Even under the most perceptually demanding conditions of our experiment, there was a detection time advantage towards multisensory stimuli. In fact, the multisensory capture effects observed in Experiments 1-3 were particularly pronounced in perceptually demanding situations, with significant violations of the race model found only under high perceptual load. While this could potentially imply greater resistance to perceptual load effects, we cannot rule out the alternative possibility that this could simply be a floor effect – that multisensory facilitation cannot decrease reaction time beyond a certain point which can already be achieved by unisensory stimuli under low load. In either case, from an applied perspective, multisensory stimuli may present greater advantages during demanding tasks. From a theoretical viewpoint, for now we conclude simply that multisensory stimuli do not appear to belong to the ‘special’ class of stimuli which are fully immune to the effects of perceptual load (e.g. human faces; Lavie et al., 2003).

The results of our research have not only theoretical implications, discussed above, but also practical implications for real-world scenarios. When driving a car, or focussing in a lecture, our results imply that an irrelevant multisensory event may be no more distracting than an already distracting unisensory one. On the other hand, if

warning signals or alerts come from a location the driver is monitoring already (e.g. a particular place on the dashboard), having this as a multisensory signal could mean a faster reaction time to detect it, compared to a unisensory one of equivalent strength. In addition, according to our results, detection may still be slower when driving through a busy town (high perceptual load) than down an empty lane (low perceptual load) for either unisensory or multisensory warning signals, although the multisensory advantage mentioned above would still be present in both circumstances.

A limitation of the present research is that our perceptual load task was always unisensory. An interesting question to consider in further research is whether multisensory capture would still be observed even if the load task itself, here the central stream monitoring, was multisensory. Research exploring this possibility is currently underway. A further fruitful direction for future research would be to identify the degree of task-relevance that is sufficient to allow multisensory enhancement of attentional capture – for example, is directing attention to the location of an item sufficient, or is it necessary to adopt an attentional setting for this item? This could be tested by adapting a task such as the singleton attentional capture task (e.g. Theeuwes, 1992), in which salient distractors appear as non-targets within the search array. Our findings could also be extended by testing whether our null findings concerning multisensory enhancement of irrelevant distraction could be replicated within other measures of distraction, for example temporal measures such as the attentional blink (Raymond et al., 1992).

In conclusion, the present research points to a nuanced bidirectional relationship between multisensory integration and attention. On one hand, our results support the possibility that multisensory integration can, in certain contexts, enhance attention. On the other hand, our findings support suggestions that some degree of endogenous

attention must be in place before integration (and hence any resulting attentional enhancement) may occur. When it does occur, multisensory enhancement of attention is further modulated by the availability of perceptual capacity, but may nevertheless be usefully exploited in applied contexts during demanding and undemanding conditions alike. As such our findings build on recent theoretical perspectives (e.g. Hartcher-O'Brien et al., 2017; ten Oever et al., 2016) by revealing a paradoxical interplay between integration and attention: multisensory processing may enhance attention, but only if you are already paying attention.

Chapter 3: Irrelevant sights and sounds require spatial suppression: ERP evidence

Chapter Summary

In Chapter 2, I established that multisensory stimuli presented as search targets offer an advantage over unisensory stimuli, in terms of accuracy and, primarily, speed of detection. On the other hand, multisensory stimuli presented as irrelevant distractors elicited no more distractor interference than unisensory visual distractors, suggesting that the enhancement of attentional capture afforded by these stimuli is dependent on the allocation of endogenous attention. However, it remains possible that the mechanisms underlying distraction by multisensory stimuli do differ from unisensory stimuli, but that behavioural measures are not sensitive enough to detect this. In the current chapter, I adapt the irrelevant distractor paradigm to be suitable to measure ERP components associated with spatial attentional selection and suppression, thus granting the ability to look for differences in spatial attentional processing. Additionally, this chapter also tests the capacity of salient yet entirely irrelevant unisensory visual and auditory distractors to compete for spatial attention. Where additional analyses will be included in the article submission as supplementary materials, these have been provided at the end of the chapter.

Abstract

Both real-world experience and behavioural laboratory research suggest that entirely irrelevant ‘distractor’ stimuli can interfere with a primary task. However, it is as yet unknown whether such interference reflects competition for spatial attention – indeed, prominent theories of attention predict that this should not be the case. Whilst electrophysiological indices of spatial capture and spatial suppression have been well-investigated, experiments have primarily utilised stimuli which share a degree of task-relevance, and are presented in the visual domain. The present research tests the ability of salient yet entirely irrelevant visual and auditory distractors to compete for spatial attention, while also testing for potentially enhanced competition from multisensory distractors. Participants completed a letter search task, while ignoring lateralized visual (e.g., a dog), auditory (e.g., a dog’s bark), or multisensory (e.g., dog image + sound) distractors. Results showed that visual and multisensory distractors elicited a P_D component indicative of active lateralized suppression. We also establish for the first time an auditory analogue of the P_D component, the P_{AD} , elicited by auditory and multisensory distractors. This ERP evidence of spatial attentional processing was observed in the absence of behavioural distractor interference, suggesting effective early suppression. Interestingly, there was no evidence to suggest enhanced ability of multisensory distractors to compete for attentional selection, despite previous proposals of a ‘special’ saliency status for such items. Our findings speak to two theoretical debates on the ability of task-irrelevant stimuli to compete for spatial attention, and the role of task-relevance in the special attentional status of multisensory stimuli.

Introduction

In our daily lives we often find that we are distracted by irrelevant sights and sounds, and this irrelevant distraction can lead to a variety of negative consequences. For example, highly distractible individuals are more at risk of serious accidents (Larson et al., 1997), and distraction by task-irrelevant stimuli has been found to account for over 10% of vehicle crashes that resulted in driver hospitalisation (McEvoy et al., 2007). Despite its apparent ubiquity in daily life, the phenomenon of distraction by entirely irrelevant events was for some time under-represented by laboratory studies of attention (see Forster, 2013). Motivated by a desire to parallel real life distraction, Forster and Lavie (2008a) developed the irrelevant distractor task. This task measures the degree of reaction time slowing to a central letter search task, associated with the presentation of highly salient distractors (cartoon images). These distractors were not only irrelevant to the main search task participants were required to undertake, but also appeared at task irrelevant, off-centre, locations. The finding that the irrelevant distractors slow down reaction times to a central task has been replicated many times (e.g. Cunningham & Egeth, 2018; Forster et al., 2014; Forster & Lavie, 2008a, 2011, 2014; He & Chen, 2010; Lunn et al., 2019) including in a large sample (Forster & Lavie, 2014), and has also been shown to correlate with clinical symptoms of distraction in adults with Attention Deficit Hyperactivity Disorder (ADHD) and those with trait level symptomology (Forster & Lavie, 2016). As of yet, the irrelevant distractor task has mainly been applied to the visual domain. However, a recent paper has reported a distraction effect employing audiovisual distractors (Lunn et al., 2019), with equivalent results.

While it is now well-established that entirely irrelevant distractors can slow us down and interfere with our current tasks, it is not known precisely what mechanism

underlies this interference. One possibility is that the reaction time slowing reflects the additional time taken for covert orienting to the distractor location and then re-orienting back to the target (i.e. spatial attentional capture). However, it has long been debated whether spatial attentional capture by entirely irrelevant stimuli is possible. Imagine searching in your sock drawer for a single pair of navy-blue socks. You would likely expect that an entirely irrelevant but salient stimulus such as a meowing ginger cat on top of the dresser would capture your attention, yet well-evidenced theories actually predict that this might not be the case. Goal-driven theories of attention argue that capture can occur towards stimuli presented in irrelevant spatial locations, but only if they have a degree of task relevance, such as those that match the attentional set of a search task (e.g. Bacon & Egeth, 1994; Folk & Remington, 1998). Therefore, attention could be captured by a navy-blue mug on top of the dresser, due to an ‘attentional setting’ for navy-blue coloured items (Folk et al., 1992), but not by a stimulus sharing no features with the target item, such as the cat. On the other hand, some prominent stimulus-driven theories of attention propose that salient stimuli are able to automatically capture attention regardless of our current goals (e.g. Theeuwes, 1992, 2010), but only if they are inside of our ‘attentional window’, an area of visual space in which task-relevant stimuli are expected to be found. Thus, this theory would predict that attention could be captured by an irrelevant item provided it was within the sock drawer, due to this being the current attentional window, but not when it was on top of the dresser. Taken together, these theories imply that entirely irrelevant distractors, appearing outside of the attentional window, are unable to catch our attention.

A challenge for demonstrating spatial capture by entirely irrelevant stimuli is that behavioural measures of spatial attention, such as spatial cuing, typically involve stimuli occurring in, or perceptually grouped around, a task relevant location (e.g. Folk

et al., 1992). Hence, in order to measure spatial capture by irrelevant stimuli without introducing a degree of task relevance, electrophysiological indices of these processes are particularly useful as indirect measures. There are two lateralised Event Related Potential (ERP) components primarily used in this area of research. First, the N2pc is held to reflect attentional selection of items in visual space (Kiss et al., 2008). This component is reflected at posterior electrodes, characterised by a negative deflection in the ERP waveform contralateral to a stimulus presented at an attended location, and elicited 200-350ms post stimulus onset. An analogous component has also been documented for selection of information within the auditory domain, occurring at anterior electrodes - the N2ac (Gamble & Luck, 2011). Second, the P_D component is the opposite of the N2pc - a contralateral *positivity*, occurring at the same electrode sites from as early as 100-200ms (e.g. Fortier-Gauthier et al., 2012) to as late as 300-400ms (e.g. Sawaki et al., 2012), which is believed to index spatially localised suppression of a distractor stimulus in response to a spatial 'attend-to-me' signal (Hickey et al., 2008; Sawaki & Luck, 2010). The size of the P_D has been found to correlate with behavioural distractor interference (Gaspelin & Luck, 2018), and also predicts whether or not a distractor elicits an N2pc. McDonald et al. (2013) found that in visual search trials where responses to a target were slow, a significant N2pc to the distractor was found followed by a marginally significant P_D (indicating the two may coexist), but on fast response trials no distractor N2pc was found, with only the P_D being elicited instead. Thus, a distractor can elicit an 'attend-to-me' signal and be actively suppressed by a participant, despite not producing a behavioural cost. No auditory analogue of the P_D has, as of yet, been established.

Research using both the N2pc and the P_D component has typically utilised stimuli which share a degree of task relevance, for example appearing in a potential

target location (as in singleton attentional capture or spatial cuing tasks, (e.g. Burra & Kerzel, 2014; Gaspar & McDonald, 2014; Gaspelin et al., 2015; Sawaki & Luck, 2010), or sharing similar features with the target (e.g. Burra & Kerzel, 2014; Gaspar & McDonald, 2014; Hickey et al., 2006; Sawaki & Luck, 2013; Weaver et al., 2017). Hence, it is as yet unclear whether entirely irrelevant stimuli would still elicit an ‘attend-to-me’ signal to capture our attention, which must be suppressed.

As noted above, the dominant focus of research into both irrelevant distraction and attentional capture has been in the visual domain, but real-world sources of distraction often involve other senses, or even a combination of more than one sense. Indeed, multisensory stimuli have been proposed to be especially effective in capturing attention (e.g. Santangelo & Spence, 2007), and it has even been suggested that signals from two different sensory modalities may integrate pre-attentively (e.g. Bertelson et al., 2000; though see Pápai & Soto-Faraco, 2017). Imagine again the cat jumping up onto the dresser whilst you searched in your sock drawer. If the cat was meowing at the same time as it appeared into your peripheral vision, would it be a stronger competitor for your attention? If this is the case, then we may expect a more efficient (faster and/or stronger) spatial attentional capture. Note that in a previous study we have found no evidence of increased distraction by multisensory stimuli when using behavioural measures alone (Lunn et al., 2019). This result initially appeared in line with previous research showing that multisensory integration is compromised when the stimuli involved are un-attended (e.g. Alsius et al., 2005; Pápai & Soto-Faraco, 2017; Talsma et al., 2007). However, behavioural distractor interference may not be sensitive to detecting a fast-acting attentional capture by multisensory distractors. Indeed, if multisensory distractors do capture attention faster, then this might actually result in less interference. Faster capture accompanied by delayed disengagement could also cancel

each other out, resulting in no effect of multisensory presentation on behavioural distraction. Such effects, however, could be captured with time-resolved neuroimaging measures such as ERPs. Hence, we adapted the present experiment from the low perceptual load condition in a multisensory distractor task used in our previous study (Lunn et al., 2019). This approach allowed us to test for capture of attention by irrelevant visual, auditory and multisensory stimuli, and index whether or not a faster spatial capture by, or a larger spatial suppression of, the multisensory distractor is present. Only low perceptual load was utilised in this experiment, as distraction by irrelevant stimuli tends to be eliminated when high load exhausts attentional capacity (Forster & Lavie, 2008a).

In summary, the primary goal of the present research was to provide a direct test of spatial capture by entirely irrelevant visual, auditory and multisensory stimuli. A contralateral negativity to these distractors (N2pc/N2ac) would signify the selection of these distractors in the visual or auditory space, and a contralateral positivity (P_D) would indicate that they do elicit a spatially localised ‘attend-to-me’ signal which must be actively suppressed. A contralateral positivity to auditory distractors would also be the first evidence of an auditory analogue to the P_D . Our second goal was to compare effects across visual, auditory and multisensory distractors. In particular, if multisensory stimuli are more effective at capturing attention then we may expect a faster time course, or increased amplitude, of these components for multisensory compared to either unisensory distractor stimulus.

Methods

Participants

Eighteen participants (15 female) aged between 18 and 24 years ($M = 20.88$, $SD = 1.82$) were recruited at the University of Sussex to participate in this experiment. Participants either gained course credits, or were paid, to take part. Data from three participants was excluded and replaced, two due to excessive movement artefacts and one due to technical issues with the electrophysiological recording. The study was approved by the Sciences and Technology Research Ethics Committee (C-REC) at the University of Sussex. All participants reported normal or corrected-to-normal vision and hearing, and no known skin, neck or head problems. Sample size calculations were conducted prior to data collection, using G*Power software (Faul et al., 2009), revealing that to detect an effect size of $\eta^2p = .37$ ($\alpha = .05$, $1-\beta = .99$), a sample size of 16 was required. This effect size was chosen from the main effect of electrode laterality (contralateral vs ipsilateral) in Hickey, Di Lollo and McDonald's (2008) Experiment 1 for the P_D , and would also sufficiently power the behavioural distraction effect ($d = 4.17$ based on Forster and Lavie's (2016) low load condition), and measurement of the N2pc ($\eta^2p = .55$ based on the main effect of laterality in Hickey, McDonald, and Theeuwes' (2006) Experiment 1) and N2ac ($\eta^2p = .50$ based on the overall N2ac in Gamble and Luck (2011)).

Stimuli and Procedure

The experiment was programmed and presented using E-Prime v2.0, on a 19-inch CRT monitor (resolution 1600 x 1200) at a refresh rate of 85Hz and at a viewing

distance of approximately 57cm. Auditory stimuli were delivered over Trust Leto 2.0 speakers positioned on the left and right side of the screen.

The task was adapted from the multisensory irrelevant distractor paradigm as used in Lunn et al. (2019) (Figure 1). As in this task, trials began with a central fixation point (white, radius 0.1°) presented over black background for 100ms, followed by a 500ms stimulus display. Participants were required to identify a target letter (X or N; white; subtending $0.4^\circ \times 0.5^\circ$) presented on screen, responding with one of two key presses to indicate which letter they had seen. Participants were instructed to ignore anything that was presented elsewhere on the screen. Distractor stimuli were presented on the left and right side of the screen, or from the left or right speaker, with each of three distractor types (visual, auditory, multisensory) being presented on 8% of trials. Visual distractors were identical to those previously used in our earlier study, consisting of a photograph of an animal, randomly selected from six possible images (dog, cat, pig, horse, cow, sheep). These were presented in full colour with a black background, subtending 5.0° to 5.6° vertically, by 4.5° to 6.3° horizontally. Auditory distractors consisted of characteristic animal sounds (same animals as the images) presented from one of the speakers (600ms-1000ms in length). Multisensory distractors were the image presented with the corresponding animal sound presented from the speaker on the same side.

Changes were made to the behavioural task established in Lunn et al. (2019), to allow for measurement of the N2pc/N2ac and P_D while controlling for low level effects. In the present experiment the target letters were presented at one of 6 possible locations along the vertical meridian of the screen (either 1.0° , 2.5° , or 4.0° above and below the central fixation point) rather than arranged in the shape of a circle, so that targets were

not presented laterally. Furthermore, on each trial, 6 images arranged in an imaginary circle (radius 9°) were displayed in the periphery. In the no distractor trials all 6 images were identical scrambled images of an animal (scrambled in 4×4 pixel blocks; selected randomly from dog, cat, cow, horse, pig, or sheep) subtending 4.4° vertically, by 6.2° horizontally. At the same time, the scrambled noise of this animal (produced in MATLAB R2018b by shuffling a matrix of the audio data while maintaining the same sampling rate) was presented from both the left and right speaker. Visual distractor images, when present, were displayed at the 3 o'clock or 9 o'clock position instead of the scrambled image of the same animal. Auditory distractor sounds consisted of the characteristic noise of the animal instead of the scrambled sound from one of the loudspeakers, whilst the scrambled noise of the same animal played from the other speaker. Multisensory distractors consisted of a combination image and sound, which was both spatially and semantically congruent. Scrambled and non-scrambled images of the same animal were approximately equiluminant (average 21.04 based on the CIELAB colour space, measured using the SHINE_colour toolbox for MATLAB; Dal Ben, 2019). The scrambled images were added to reduce low-level perceptual asymmetry between the two sides of the display. As it was particularly important to avoid eye movements in the present task, the central fixation point remained on screen throughout the entire block, to avoid an additional offset/onset, and the ITI was jittered (1300-1800ms) in order to make habituation to the displays less likely.

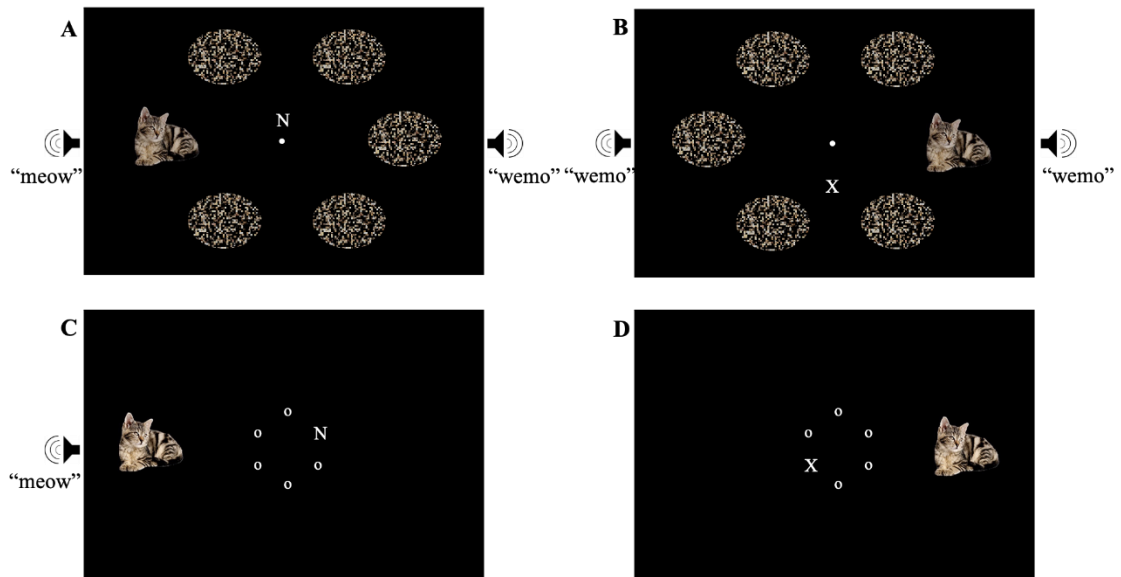


Figure 1. Example stimulus displays from the present experiment (top row) compared with displays from Lunn et al. (2019; bottom row). A. multisensory cat distractor with scrambled cat images and scrambled cat sound from the speaker on the opposite side, B. visual cat distractor with scrambled cat images and audio from both speakers, C. multisensory peripheral cat distractor, D. visual cat distractor. As can be seen the present experiment includes the addition of visual and auditory scrambles to control for low level effects, and the central task was changed from a low load X/N letter identification in a central circle, to an X/N letter identification along the vertical midline, to avoid lateralised targets which would elicit an N2pc themselves.

Participants completed three, slowed down, example trials (i.e. stimuli remained on screen until response), followed by 12 practice trials. They then completed 12 blocks of 100 trials for a total of 1200 trials, of which 96 contained auditory distractors, 96 visual distractors and 96 multisensory distractors. Distractors were presented randomly throughout blocks, excluding the first three trials which were warm-up trials and always had no distractor.

ERP Recording and Analysis

Electrical brain activity was continuously digitized using a 64 channel ANT Neuro amplifier at a 1000 Hz sampling rate. Horizontal EOG was also recorded bipolarly using electrodes placed at the outer canthi of both eyes, and vertical EOG from the inferior and superior orbit of the left eye. Data processing was conducted using EEGLab (Delorme & Makeig, 2004) and ERPLab (Lopez-Calderon & Luck, 2014). Bad channels were interpolated using EEGLab's spherical spline interpolation function (<1% of total channels, none of which were those used in the main analyses). Following referencing to the average of the left and right mastoids, the following filters were applied; 0.1Hz (12 db/oct; zero phase) high-pass, 30Hz (24 db/oct; zero phase) low-pass, 50Hz notch filter (to remove line noise) and an 85Hz notch filter (to remove line noise specifically caused by the speakers, visual inspection of the data confirmed that the notch filter did not create or remove components). Epochs were baseline corrected according to a 100ms pre-stimulus presentation window and neural activity was examined for 500ms post-stimulus presentation. Automatic offline artefact rejection was performed by removing epochs contaminated by eye blinks (peak-to-peak threshold $\pm 75\mu\text{V}$) and eye movements (step-like artifacts threshold $\pm 30\mu\text{V}$), at VEOG and HEOG, respectively. Where there were technical issues with the recording from these electrodes, Fp1 was used and epochs were manually inspected ($n = 2$ in each case). Given the importance of removing all HEOG activity when measuring the N2pc or the P_D, we note that all statistical analyses remain significant with the exclusion of the two participants for whom the recording was affected. In all remaining participants, we assess whether residual HEOG activity remained after artifact rejection during our critical time-window (100-300), by comparing mean HEOG amplitude on trials where the distractor was presented in the left versus right hemifield, separately for all three

distractor types. No systematic differences were present ($ts < 1.38$, $ps > .189$). Epochs contaminated by drifts, blocked electrodes or muscle-related potentials ($\pm 200\mu V$ at all other electrodes) were also rejected. No correct distractor trials had RTs of less than 100ms, so none were excluded on this basis. All trials in which participants made incorrect responses were also excluded, leaving on average 68.90% (SD = 15.99) of distractor trials to be analysed. PO7/PO8 electrodes where the N2pc and P_D are typically maximal were chosen a priori to be examined for these components for both visual and multisensory distractor conditions. N2ac was found at a cluster of anterior electrode sites (Gamble & Luck, 2011) which we examined for the presence of this component, or an auditory analogue of the P_D, for both auditory and multisensory distractor conditions.

The N2pc and P_D components have been found to vary substantially in their onset latency, with the N2pc typically occurring approximately 200-350ms after stimulus onset (Kiss, Velzen & Eimer, 2008), and the P_D from as early as 100-200ms (e.g. Fortier-Gauthier, Moffat, Dell'Acqua, McDonald, & Jolicœur, 2012), to as late as 300-400ms (e.g. Sawaki, Geng, & Luck, 2012). Due to this, as well as the potential for bias in mean amplitude measurements depending on the choice of measurement window (Luck, 2014), we confirmed the amplitude analyses using a non-parametric permutation method developed by Sawaki et al. (2012). See supplementary materials for more information.

The first 3 trials from each block, and any in which RT was less than 100ms, were excluded from behavioural analysis of the central task.

Results

Behavioural Results

A one-way within-subjects ANOVA with the factor of distractor type (multisensory, unisensory auditory, unisensory visual, no distractor) on RT in correct trials, revealed no main effect of distractor ($p = .144$, $\eta^2 = .10$). Unexpectedly, and in contrast to our earlier findings, no evidence of behavioural distraction was found: RTs were no slower in the presence of any of the three distractor types compared with no distractor trials ($t < 1$, $B_{H(0,23)} = .34$ for multisensory distractors; $t < 1$, $B_{H(0,23)} = .09$ for auditory distractors; $p = .067$, $B_{H(0,17)} = 1.68$ for visual distractors), nor were multisensory distractors any more distracting than either of the two unisensory types ($p = .108$, $B_{H(0,23)} = .84$ and $t < 1$, $B_{H(0,23)} = .12$ for multisensory compared with auditory and visual, respectively). We note that the present experiment is substantially longer than previous studies employing this paradigm, which typically employ 4 blocks for each load condition (e.g. Forster & Lavie, 2008a, 2008b; Lunn et al., 2019). To rule out the possibility that the distractor effect was present at the start of the experiment but reduced across its duration, we also ran these analyses on the first 4 blocks only. All results remained the same, with no main effect of distractor ($p = .310$, $\eta^2 = .07$), and no differences in RT in the presence of any of the three distractor types ($ps < .102$).

A one-way within-subjects ANOVA with the factor distractor type (multisensory, unisensory auditory, unisensory visual, no distractor) on percentage error rate, revealed a main effect of distractor type ($F(2.11, 35.81) = 3.62$, $p = .035$, $\eta^2 = .18$). Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(3) = .34$, $p = .004$, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .70$) (Table 1). Error rates were higher when a multisensory

or auditory distractor was present ($ps < .001$), but not when a visual distractor was present ($p = .088$), compared to no distractor.

	Distractor Condition			
	Multisensory	Auditory	Visual	No Distractor
RT(ms)	551	545	556	548
	(26)	(27)	(28)	(25)
% Error	7.58	6.67	8.06	9.22

Table 1. Mean RTs (SE in parentheses) and error rates (%) as a function of load and distractor type

Electrophysiological Results from Distractor Trials

Visual and Multisensory Distractors

Figure 2 presents the grand averaged waveforms from lateral occipital scalp sites (PO7/8) for multisensory and unisensory distractors at contra- and ipsilateral electrodes, as well as the contralateral-ipsilateral difference waves. As can be seen, there is no evidence for presence of an N2pc, yet there is clear positivity at contralateral relative to ipsilateral electrode sites, spanning approximately 100-300ms (see also Figure 3). An initial mean amplitude analysis from 100-300ms revealed a main effect of electrode laterality ($F(1, 17) = 6.84, p = .018, \eta^2 = .29$), with contralateral waveforms more positive than ipsilateral waveforms, indicating the presence of a reliable P_D component. However, there was no significant interaction between laterality and distractor type ($F < 1$), indicating that the P_D did not differ in size between multisensory and unisensory visual distractors.

Visual inspection of the difference wave suggests that the P_D component for multisensory and visual distractors has two distinct peaks. To test whether the active suppression occurs earlier for the multisensory distractors compared with the visual distractors, we examined the peak latency of both the early and late peaks, determined by the maximum positive amplitude in the contra-minus-ipsilateral difference wave before and after 150ms. Neither peak occurred significantly earlier for the multisensory distractors ($M = 124.80$, $SD = 15.23$ for early distractors; $M = 220.70$, $SD = 43.04$ for late distractors) than the visual distractors ($M = 127.60$, $SD = 16.39$ for early distractors; $M = 225.60$, $SD = 27.40$ for late distractors), $ts < 1$, indicating that the spatial suppression occurred no earlier for the irrelevant multisensory stimuli. We also confirmed that both the early and late peaks were of equal mean amplitude between visual and multisensory distractors, and each showed a significant contralateral positivity, see Supplementary Materials for the full analyses.

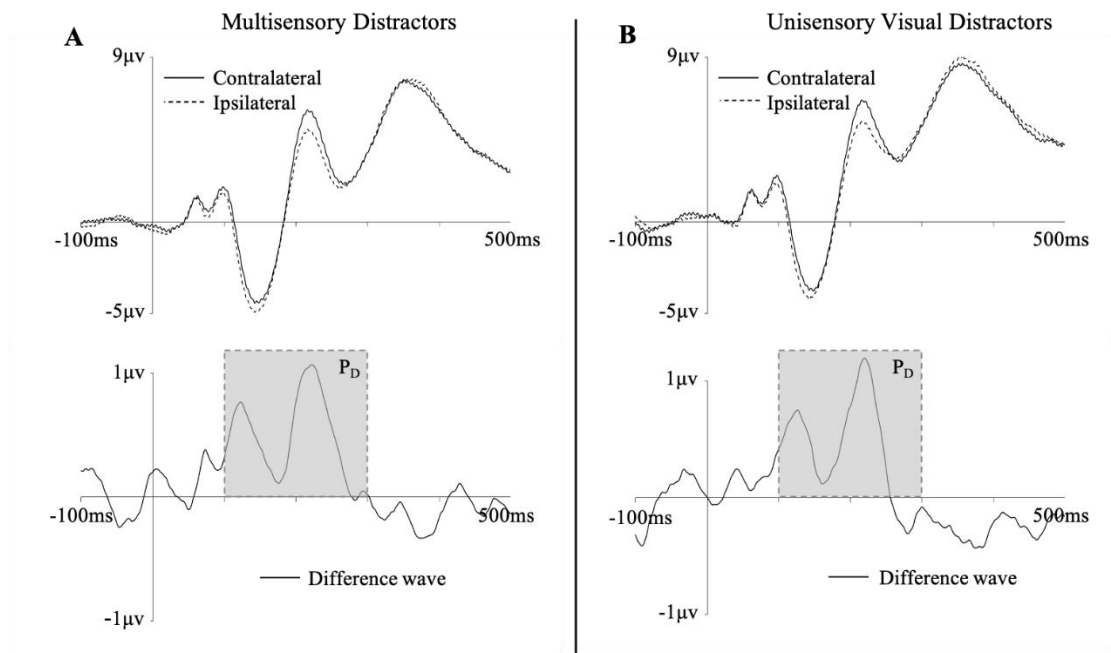


Figure 2. Grand averaged contralateral, ipsilateral and contra-minus-ipsilateral ERP waveforms for A. multisensory distractors and B. unisensory visual distractors, at electrode sites PO7/PO8

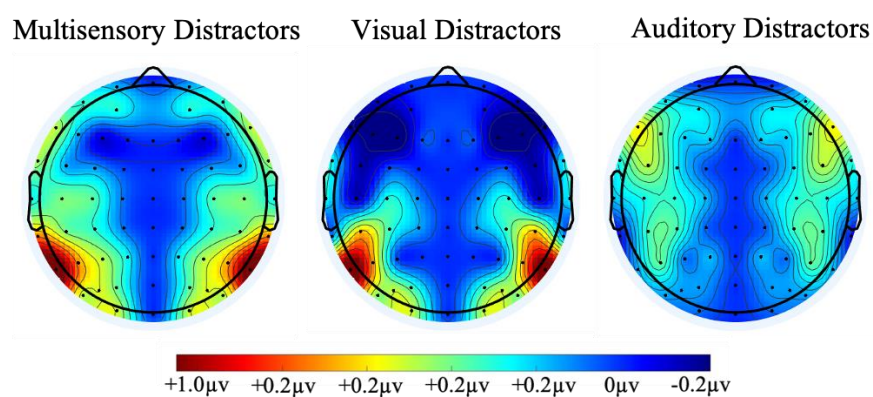


Figure 3. Topographic scalp maps of the P_D and P_{AD} , based on mean voltage 100-300ms for the contra-minus-ipsilateral difference wave

Auditory and Multisensory Distractors

Figure 4 presents the grand averaged waveforms, across a cluster of anterior scalp sites (F3/4, F7/8, C3/4 and T7/8), for multisensory and unisensory auditory distractors at contra and ipsilateral electrodes, as well as the difference waves. As can be seen, there is no evidence for the presence of an N2ac at this anterior electrode cluster, yet there is clear positivity at contralateral relative to ipsilateral electrode sites, spanning approximately 100-300ms (see also Figure 3). An initial mean amplitude analysis from 100-300ms revealed a main effect of electrode laterality ($F(1, 17) = 7.48, p = .014, \eta^2 = .31$), with contralateral waveforms more positive than ipsilateral waveforms. However, there was no significant interaction between laterality and distractor type ($F < 1$). This suggests that there was an auditory equivalent of the P_D – indicative of spatial suppression of auditory distractors – for both conditions, and that this was of equal size between multisensory and unisensory auditory distractors.

To test whether this spatial suppression occurred earlier for the multisensory distractors than the auditory-only distractors, we compared the peak latency of the contralateral negativity at the anterior electrode cluster, determined by the maximum positive amplitude in the contra-minus-ipsilateral difference waves between 100 and 300ms. This showed that the auditory P_D did not occur significantly earlier for the multisensory distractors ($M = 185.60, SD = 54.89$) than the auditory distractors ($M = 182.90, SD = 50.45$), $t < 1$, indicating that the active suppression occurred no faster for a multisensory distractor.

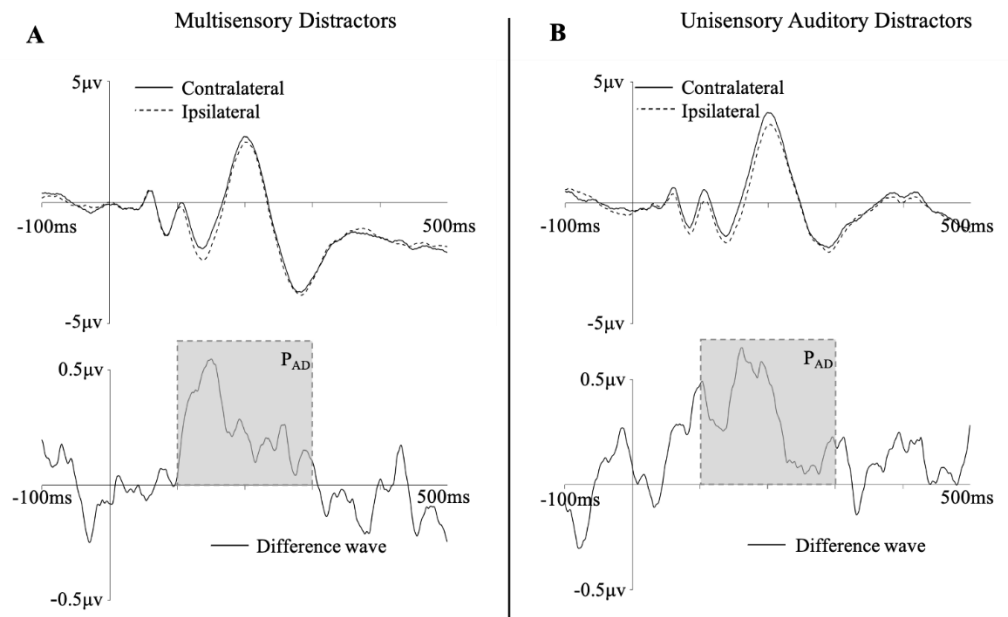


Figure 4. Grand-averaged contralateral, ipsilateral and contra-minus-ipsilateral ERP waveforms for A. multisensory distractors and B. unisensory auditory distractors, collapsed across electrode sites F3/4, F7/8, C3/4 and T7/8

General Discussion

The present study has three key findings. First, we demonstrate for that first time that entirely irrelevant visual distractors, appearing outside of the attentional window and sharing no features with the task stimuli, reliably elicited a P_D component. The P_D is argued to reflect suppression of a distractor stimulus in response to a spatial ‘attend-to-me’ signal (Hickey et al., 2008; Sawaki & Luck, 2010). Second, we establish for the first time an auditory analogue to the P_D component, henceforth referred to as the P_{AD} . Finally, we did not reveal any evidence for differences in spatial suppression of multisensory versus unisensory stimuli, in terms of either the time-course or amplitude of the P_D / P_{AD} elicited by these stimulus categories.

As noted in our introduction, it has long been debated whether task-relevance, through feature settings or location relevance, is necessary for attentional capture (e.g. see Folk et al., 1994; Theeuwes, 1992, 2010). The ability of our distractors to elicit the P_D and P_{AD} components imply that stimuli can compete for spatial attention, requiring active suppression to avoid distraction, even when they are entirely irrelevant to any top-down goals associated with the task, and presented in an irrelevant location outside of the attentional window. Given the established sensitivity of the irrelevant distractor task to index the clinically inflated levels of distractibility seen in ADHD (Forster et al., 2014; Forster & Lavie, 2016), the suppression mechanism underlying the P_D appears a potential candidate for the disruption underlying ADHD-related distractibility.

Our distractors did not elicit an $N2pc$ or $N2ac$ component, which might initially appear to imply a limitation in the ability of irrelevant distractors to fully capture spatial attention. While we cannot rule out this possibility, we note that this finding must be taken in the context that, unlike prior studies using this task, we did not detect a

behavioural distractor effect. As such, our ERP findings and behavioural findings concur in suggesting that, in this version of the task, the suppression mechanism reflected by the P_D component was successful in overcoming full spatial attentional capture (as would be reflected by an $N2pc$ and behavioural distraction effect). The lack of a behavioural effect is surprising and we cannot fully attribute it to lack of sensitivity, given the high power to detect a well-replicated irrelevant distractor effect (99% based on both the effect size from the largest sample size, $N=100$ Forster & Lavie, 2016 for the irrelevant visual distractor task, and from our most similar study by Lunn et al. 2019). We speculate that the behavioural distraction effect may have been undermined by changes necessitated in order to measure the ERP components of interest. Specifically, the inclusion of scrambled versions of each distractor was necessary to rule out an account of any lateralised effect in terms of early sensory differences, with balancing of displays to equate for low level factors being typical of research where either the $N2pc$ or P_D components are measured (e.g. Berggren & Eimer, 2018; Gaspar & McDonald, 2014; Hickey et al., 2008; Sawaki & Luck, 2013); However, this change meant that the distractor lost its status as a unique abrupt onset, hence undermining its salience (see Figure 1). A remaining question therefore is whether the robust behavioural effects seen in the original version of the task are accompanied by an $N2pc$. For now, we conclude that our results imply that, at the very least, entirely irrelevant distractors can elicit a spatially localised ‘attend-to-me’ signal that competes for attention and requires suppression.

The second key contribution of the present paper is to establish a new ERP component: the P_{AD} , an auditory analogue of the P_D component. This extends work by Gamble and Luck (2011), who demonstrated an auditory analogue of the $N2pc$ – the $N2ac$ – characterised by a comparable contralateral negativity to the target, but

occurring at a lower amplitude and across a more prolonged time window, at a cluster of anterior electrodes. Similarly, we observed a lower in amplitude but more sustained contralateral positivity to the distractor at the same electrode cluster. This is not only, to the best of our knowledge, the first evidence of a P_{AD} , but also suggests that even an irrelevant stimulus presented in a different sensory modality to the target requires active suppression to avoid distraction.

Future research may focus on the extent to which the P_{AD} reflects a similar (or the same) suppression mechanism as the P_D . For example, is the size of the P_{AD} also associated with individual differences in attentional control (Gaspar et al., 2016; Sawaki et al., 2017), and is the P_{AD} only elicited where participants successfully orient their gaze to the target before or without fixating on the distractor (Weaver et al., 2017)?

Finally, our results speak to an ongoing debate regarding the proposed special status of multisensory stimuli for attentional capture (e.g. Van der Burg et al., 2008). If stimuli presented concurrently in two sensory modalities must integrate in order to enhance attention, then determining the conditions under which successful integration occurs is essential. Whilst there are a number of studies demonstrating that this integration requires endogenous attention, there are others suggesting that integration is automatic (see Soto-Faraco et al., 2019; Spence & Frings, 2019; ten Oever et al., 2016 for reviews). Our design allowed us to compare the time course of spatial suppression for multisensory (audiovisual) versus unisensory (auditory or visual) distractors, as well as testing for potential interaction between these components. Whilst a multisensory distractor elicited both a P_D and a P_{AD} , these did not occur at a different latency, or at an increased amplitude, to those elicited by the individual unisensory components. This suggests that the two sensory modalities are suppressed through two independent

processes, with no multisensory integration which would require an enhanced, or non-linear, active spatial suppression. We note that our design did not allow for a comparison between the multisensory and the sum of unisensory ERPs, however previous research that has done this has demonstrated a lack of non-linearity, indicative of multisensory integration, for an unattended multisensory stimulus (Talsma et al., 2007; see also Pápai & Soto-Faraco, 2017 for a similar argument based on behavioural methods). This finding also supports our previous research where a to-be-ignored multisensory stimulus did not result in greater distraction than a unisensory stimulus (Lunn et al., 2019). Additionally, the ERP study conducted by Talsma et al., (2007) examined the effect of endogenous attention on the sensory components (e.g. P1 and N1), thus our study extends this previous research by looking at those that reflect spatial attention processes. As such, our results imply that entirely irrelevant multisensory stimuli do not hold any special attentional status.

In conclusion, we demonstrate that both visual and auditory distractors require active spatial suppression even when completely irrelevant, and that we avoid distraction by multisensory stimuli through independent suppression mechanisms acting on the two sensory components. Our findings speak to two theoretical debates on the ability of task-irrelevant stimuli to compete for spatial attention, and the role of task-relevance in the special attentional status of multisensory stimuli, and point to the P_D and P_{AD} as potential indices of the disruption underlying clinical symptoms of distractibility.

Supplementary Materials

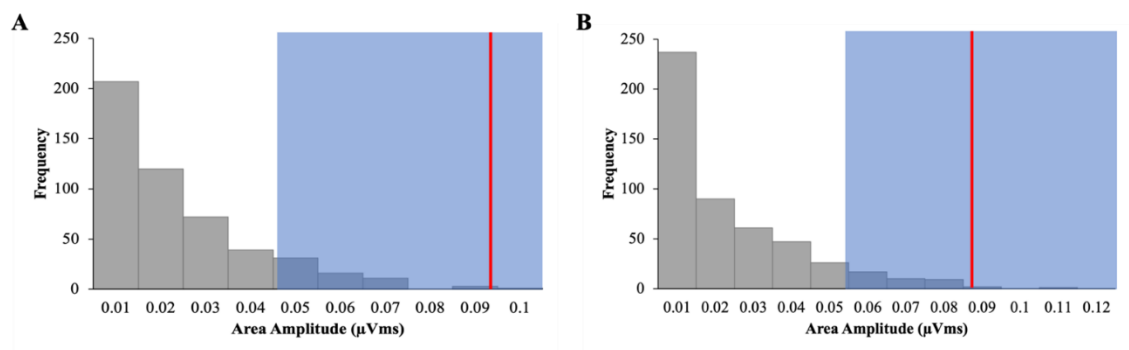
Non-Parametric Permutation Analyses

Due to the potential for bias in mean amplitude measurements depending on the choice of measurement window (Luck, 2014), we confirmed the amplitude analyses using a non-parametric permutation method developed by Sawaki et al. (2012). In this approach, the distribution of area amplitude that could be expected from noise alone is estimated from random permutations of the data, and no assumption is required with regards to normality, equal variance, etc (Gaspelin & Luck, 2018; Sawaki et al., 2012, 2017; Sawaki & Luck, 2013).

To perform the permutation test, the side of the distractor was randomly recoded for every trial and participant, and ERPs were re-averaged to a grand average. Positive area value was then measured between 100-300ms, at PO7/PO8 for the visual and multisensory distractors, and at the averaged anterior electrode cluster for the auditory and multisensory distractors. This was repeated 500 times, resulting in a null distribution. If the observed area amplitude from the original data is greater than 95% of the values in the null distribution, then the P_D can be considered significant. To compare the P_D and P_{DA} elicited by multisensory versus unisensory distractors, the same method was used, but it was the modality of the distractor that was randomly recoded, with the laterality kept the same.

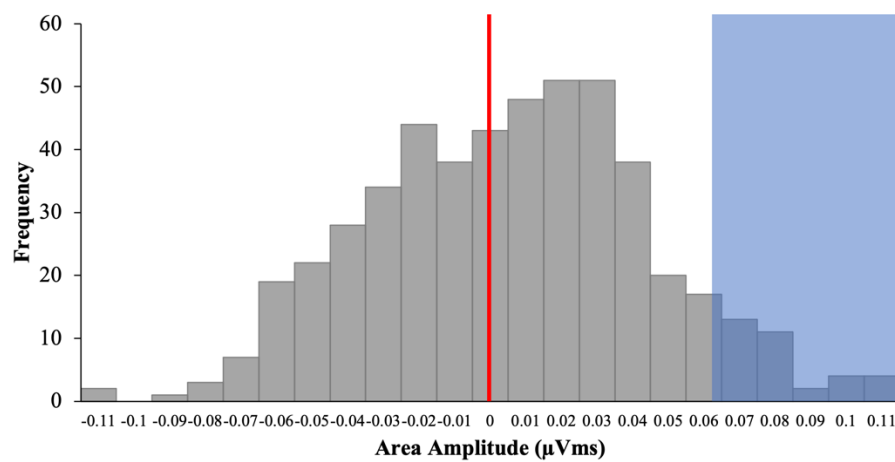
Visual and Multisensory Distractors

The null distributions of the multisensory and visual distractor conditions are shown in Supplementary Figure 1. For both distractor types the observed area (shown as a red vertical line) is greater than the 95th percentile of the null distribution (anywhere within the blue area would indicate this), and therefore can be taken as a significant P_D .



SF1. Observed value (red line) is significant if it falls within the top 5% of values of the permutation distribution (blue area). A. Null distribution of multisensory distractors and observed significant value, B. Null distribution of visual distractors and observed significant value

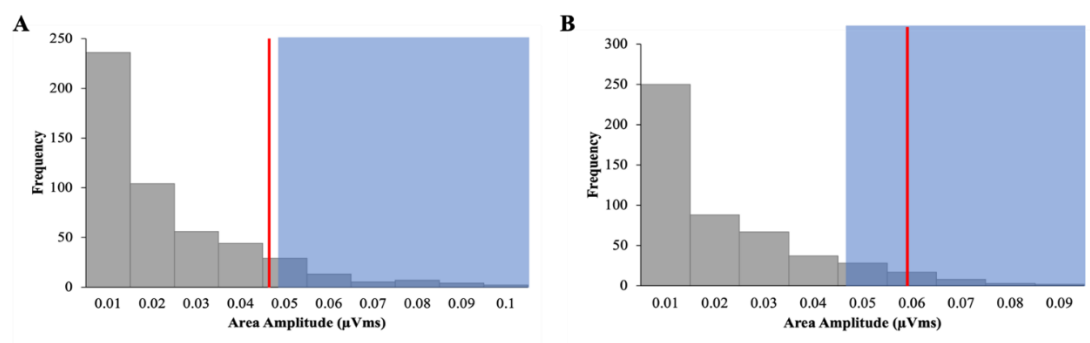
Null distributions of the interaction effect (i.e. the difference wave of the multisensory and visual difference waves) were also calculated, and the observed area (red line) is less than the 95th percentile of the distribution, and is confirmed to be non-significant (Supplementary Figure 2).



SF2. Null distribution of interaction effect between distractor type (multisensory and visual), and laterality (contra, ipsi), and non-significant observed value

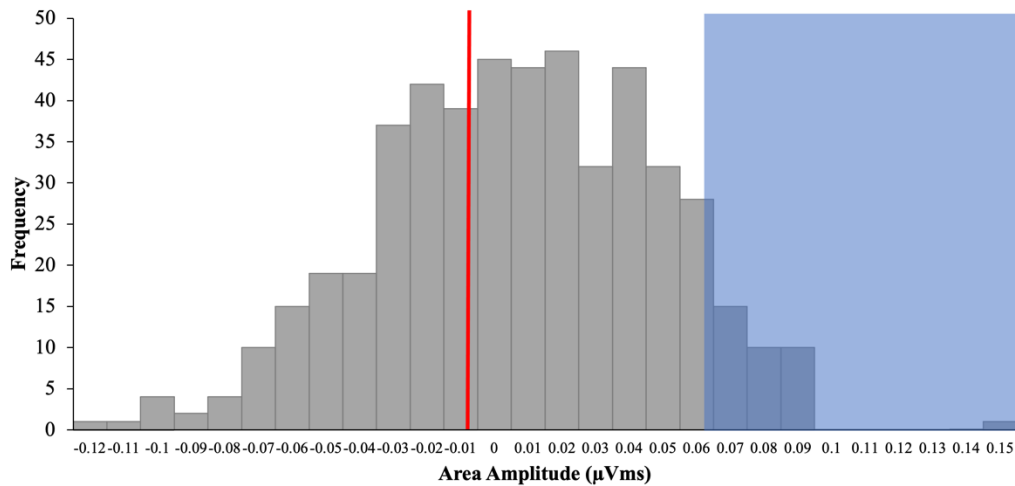
Auditory and Multisensory Distractors

The null distributions of the multisensory and auditory distractor conditions are shown in Supplementary Figure 3. For the auditory distractor, the observed area (shown as a red vertical line) is greater than the 95th percentile of the null distribution (anywhere within the blue area would indicate this), and therefore can be taken as a significant P_{AD} . The observed area for the multisensory distractor falls just short of the 95th percentile.



SF3. A. Null distribution of multisensory distractors and observed marginally significant value, B. Null distribution of visual distractors and observed significant value

Null distributions of the interaction effect (i.e. the difference wave of the multisensory and auditory difference waves) were also calculated, and the observed area (red line) is less than the 95th percentile of the distribution, and is confirmed to be non-significant (Supplementary Figure 4).



SF4. Null distribution of interaction effect between distractor type (multisensory and auditory), and laterality (contra, ipsi), and non-significant observed value

Early and Late Pd

We examined the P_D amplitude for both the early and the late peaks, to determine whether or not the mean amplitude of each one is equal across visual and multisensory distractors. Mean amplitude at PO7/PO8 was entered into a 2 x 2 x 2 within-subjects ANOVA with the factors of distractor type (multisensory, unisensory visual), laterality (contra, ipsi) and P_D peak (early: 100-150ms, late: 200-250ms). This revealed a main effect of electrode laterality ($F(1, 17) = 15.50, p = .001, \eta^2 = .48$), with contralateral waveforms being more positive than ipsilateral waveforms. There was also

a main effect of time ($F(1, 17) = 18.44, p < .001, \eta^2 = .52$), with waveforms at the late P_D peak being more positive than those at the early P_D peak. Critically, there were no significant interactions between any of the factors ($ps > .120, \eta^2s < .14$), indicating that the extent to which the contralateral waveform is more positive than the ipsilateral waveform is equal for both multisensory and unisensory visual salient distractors, at both the early and late P_D peaks.

Chapter 4: Load-induced inattentional deafblindness

Chapter Summary

In Chapter 2 I established that, as is the case for unisensory stimuli, multisensory stimuli are subject to modulation by perceptual load, in terms of both speed of detection and distractor interference. In this chapter, I extend this finding by applying established load-induced inattentional blindness and deafness paradigms to measure detection sensitivity of a multisensory stimuli when the unisensory counterparts themselves are hard to detect. In this way, I examine whether multisensory stimuli maintain their ability to reach conscious awareness under perceptually demanding conditions. Where additional analysis will be included in the article submission as supplementary materials, these have been provided at the end of the chapter.

Abstract

In the absence of attention, people often miss important signals, which can have catastrophic consequences. The perceptual demands (or load) of a current task have been demonstrated to be a powerful determinant of such failures of attention and awareness, with a demanding primary task shown to increase susceptibility to inattentional blindness, deafness, or numbness. Given prior proposals regarding the special status of multisensory stimuli, including potential immunity to perceptual load effects, the present experiment tested whether awareness of multisensory stimuli would be similarly modulated by load, or whether such stimuli might maintain their ability to reach conscious awareness even under perceptually demanding task conditions. To this end, we adapted previous load-induced blindness and deafness paradigms to include multisensory stimuli: Participants performed a letter search task with high or low load, while also monitoring for peripheral targets which could be auditory, visual, or audiovisual. We replicated prior findings of load-induced blindness and deafness. Critically, a similar reduction in detection sensitivity under high load was observed for multisensory (audiovisual) and unisensory (visual or auditory) stimuli. These findings provide the first demonstration of load induced ‘inattentional deafblindness’, and have important real-world implications for predicting the situations in which even multisensory signals might be missed.

Introduction

A paradox of attention is that, while we may often wish to focus exclusively on a particular task, it is also critically important that we remain able to notice other important signals. For example, while reading a book we might wish to block out the sound of our neighbour's loud music, but we would not wish to miss an important signal such as the doorbell or the beeping of a smoke alarm. Unfortunately, it is well documented that people often do miss such signals (Dalton & Fraenkel, 2012; Macdonald & Lavie, 2008; Mack & Rock, 1998), sometimes with catastrophic consequences such as automobile or aviation accidents. For example, Adam Air flight 574 crashed in 2007 after the crew, preoccupied by troubleshooting a malfunction of a navigation system, did not notice the plane's turn and descent despite an alarm sounding. Indeed, inattention has been found to be a leading cause of fatal road crashes (Sundfør et al., 2019; Wundersitz, 2019). Predicting when such failures of awareness are most likely to occur, and which types of stimuli are most likely to be missed, are hence key theoretical questions for the field of attention research with important applied implications.

One well-established and powerful determinant of awareness is the perceptual load of the current task. Load Theory (Lavie, 1995) argues that attention works with a limited perceptual capacity, automatically processing stimuli until this capacity is depleted. Under conditions of high perceptual load (in terms of either a large amount of information to be processed or complex perceptual discrimination), all processing capacity must be fully devoted to the task at hand, and therefore processing of other, concurrent stimuli is reduced. On the other hand, under low perceptual load conditions, spare capacity remaining after processing the most relevant information 'spills over' to allow processing of other stimuli. These predictions have been supported by numerous

behavioural and neuroimaging studies, including those on awareness, which demonstrate that high perceptual load can render us inattentionally blind (Macdonald & Lavie, 2008), deaf (Raveh & Lavie, 2015) and numb (Murphy & Dalton, 2016). As such, a key implication emerging from the Load Theory literature is that whilst engaged in a more perceptually demanding task such as driving, we are more vulnerable to failures to detect warning lights or alarms.

In terms of predicting which type of stimuli are more or less likely to be missed, one type of stimulus which may be ‘special’ in their ability to reach conscious awareness are those which produce correlated stimulation in more than one sensory modality at a time. Multisensory stimuli are often processed faster or produce stronger responses than unisensory stimuli: According to many studies, this enhanced multisensory response is not merely due to the summed effects of concurrent information, as multisensory stimuli often elicit faster responses than would be predicted by additive models of the two unisensory stimuli (Colonius & Diederich, 2004; Hughes et al., 1994; Laurienti et al., 2004; Lunn et al., 2019; Molholm et al., 2002; Murray et al., 2005; Pannunzi et al., 2014; Senkowski et al., 2005, though see Otto & Mamassian, 2012). Whilst a lack of conscious awareness of a variety of different stimuli has been demonstrated when attention is focused elsewhere, as of yet this research has only examined inattention towards events occurring in one modality. An interesting theoretical question, therefore, is whether multisensory stimuli would also be subject to such failures of detection.

Intriguingly, Santangelo and Spence (2007) proposed that multisensory stimuli may be immune to the effects of perceptual load (although see Santangelo et al., 2011 for an updated interpretation of the findings originally presented in support of this

claim). Such immunity would confer a particular advantage for multisensory alerts, in that they would be effective even during perceptually demanding tasks (e.g. troubleshooting a navigation system or weaving in and out of traffic), and would support previous proposals regarding their success in these situations (Ho et al., 2007). Our recent research (Lunn et al., 2019), however, has shown that multisensory stimuli are not always entirely immune to perceptual load effects: While facilitatory multisensory attentional capture (in terms of speeded responses by multisensory versus unisensory stimuli) was observed under both load conditions, responses to multisensory stimuli were slowed under high load conditions to the same degree as unisensory stimuli.

It might be argued, however, that this prior study was not optimally sensitive to detect multisensory immunity to load effects. In the series of reaction time experiments run by Lunn et al., (2019), both the multisensory and unisensory stimuli were relatively easy to detect (e.g. in Experiment 3 they were pictures of animals with their corresponding sound), as reflected in high detection accuracy (~90-98% across experiments). Stimulus intensity is known to be an important factor in multisensory integration, with the largest multisensory enhancements and interactions being elicited when the unisensory constituents are weakest (Senkowski et al., 2011; Stevenson et al., 2012; Stevenson & James, 2009), a phenomenon termed the principle of inverse effectiveness (IE). As such, it remains possible that the relatively high salience of the stimuli used by Lunn et al impeded the integration of the multisensory stimuli and hence undermined their ‘special’ attentional quality. It hence remains possible that with less salient, harder to detect stimuli such as those used in established load induced inattention blindness and deafness paradigms (Macdonald & Lavie, 2008; Raveh &

Lavie, 2015), the advantageous nature of multisensory interactions may be further enhanced, and an immunity to perceptual load demonstrated for awareness.

The present experiment will adapt the aforementioned load induced blindness and deafness paradigms to compare perceptual load effects on the detection sensitivity of a multisensory versus unisensory (visual and auditory) stimuli. The key aim here is to determine whether awareness of a multisensory stimulus is modulated by perceptual load. If, based on the principle of inverse effectiveness, multisensory integration of these hard to detect stimuli allows it to break through into conscious awareness irrespective of perceptual capacity limitations, we may find similar detection sensitivity in both high and low perceptual load conditions for the multisensory critical stimuli (CS). On the other hand, if multisensory stimuli are not immune to the effects of perceptual load, we would expect detection sensitivity for both multisensory and unisensory stimuli to be reduced under conditions of high perceptual load.

Materials and Methods

Participants

74 participants aged between 18 and 35 years were recruited at the University of Sussex to take part in this experiment. All participants reported normal or corrected-to-normal vision and hearing. The apriori stopping rule for this experiment was based on Bayes Factors for the main effect of load, all planned comparisons, and the multisensory advantage in each load condition, on d' reaching sensitivity (see Rouder, 2014) after applying the same exclusion criteria as used in both Macdonald and Lavie (2008) and Raveh and Lavie (2015). These criteria led to the exclusion of two participants for failing to reach 50% accuracy in the central task, 32 for failing to reach 30% accuracy in detection of one or more CS type (17 in high load, 15 in low), five for failing to comply with task instructions, and three due to technological issues. Given that the exclusion criteria were part of our apriori stopping rule, and for consistency with previous research, we report here the data from the 32 participants (age $M = 20.53$, $SD = 2.17$, 29 female) who survived these criteria. However, in the light of the unexpectedly high number of participants who failed to meet the 30% accuracy criterion for CS detection ³, in the Supplementary Materials we report our analysis based on the 46 participants (age $M = 20.41$, $SD = 2.07$, 43 female) who met a more lenient inclusion criterion of 10% accuracy – critically, we note that the key contrast of load modulation of d' for the MS stimulus remained sensitive regardless of which exclusion criterion is used. We note

³ We speculate that this was due to both the increased difficulty of the unisensory CS detection relative to prior work (which was necessary in order to avoid ceiling effects for the MS stimulus), as well as the sample being comprised of undergraduates at the end of term where their attention may be more focused on gaining enough course credits for a module, rather than their performance (for a similar argument, see Wilbiks et al., 2020)

that exclusions on the basis of CS detection rate did not disproportionately affect one of the load conditions, which would bias our load modulation.

Stimuli and procedure

The experiment was programmed and presented using E-Prime v2.0, on a 16inch screen. A viewing distance of 57cm was maintained using a chin rest. Loudspeakers positioned on the left and the right side of the screen were used to deliver sounds. Each trial began with a central fixation dot presented for 500ms, followed by a 100ms stimulus display. The stimulus display consisted of six letters evenly arranged in an imaginary circle (2.0° radius) that was centred at fixation. The target letter, either an X or an N (subtending $0.5^\circ \times 0.7^\circ$) appeared at random, but with equal likelihood, at one of the six letter locations, and participants were required to indicate which of the two letters was seen by pressing one of two keys. In the high load condition, the non-target letters were pseudo-randomly selected from a set of angular letters (H, K, M, V, W, Z) and presented in the same size as the target letter, whereas in the low load condition the non-target letters were all small, placeholder O's (diameter 0.2°). All stimuli were presented on a black background, and all letters were white.

Combining the methods of MacDonald and Lavie (2008) and Raveh and Lavie (2015), low salience visual and auditory stimuli (or both in the multisensory condition) were presented to either side of the screen as the CS, and visual and auditory masks were used, making detection of these stimuli more difficult. White noise was played for 600ms during each trial, starting at the onset of the stimulus display, and a black mesh pattern used as a visual mask was presented on either side of the screen (subtending $11.6^\circ \times 12^\circ$; 2° edge to edge from the central circle of letters) immediately following the

stimulus display, and remained on screen for 500ms. On the unisensory auditory CS trials, a 1025Hz pure tone was presented for 100ms from one of the speakers. On the unisensory visual CS trials, a white rectangle of 0.2° vertically by 0.3° horizontally was presented to either the left or the right of the circular array of letters for 100ms. The multisensory CS was both the circle and the beep presented at the same time, and on the same side. CS type was the same across a block, and appeared on 17% of trials, but could not appear in the first three trials of each block. Participants were required to indicate detection of the CS by pressing the space bar before making their target letter search response to rule out an alternative account of any load effect being due to memory failure induced by the high load condition (see MacDonald & Lavie, 2008 Experiment 3; Raveh & Lavie, 2015, Experiment 2). Participants were instructed to make the letter search response as quickly and accurately as possible if they did not detect the CS in a given trial.

Participants first completed three, slowed down, example trials (i.e. the circle of letters remained on screen until response), and then 12 practice trials, of the central task. The practice trials were repeated until participants achieved at least 75% accuracy for the central task. Participants then completed a further 12 practice trials, containing each type of CS, where they were instructed to ignore the central task and focus only on detecting the peripheral stimuli. This was repeated until participants achieved 100% accuracy for detection of the three CS types. They then completed two experimental blocks of 70 trials for each CS type, all of the same perceptual load, with the order counterbalanced across participants. Within a block, the central target position, central target identity, and CS side was fully randomised. Participants were instructed to respond as quickly as possible whilst still being accurate (2000ms response deadline).

Bayesian analysis

Bayes Factors (B) were used to assess the strength of evidence of for H1 relative to H0 (Wagenmaker et al., 2015). A B of above 3 is indicative of substantial evidence for H1, whereas a B of below 1/3 indicates substantial evidence for H0, and between these values indicates the data is insensitive (Dienes, 2014). The strength of the evidence for either hypothesis is interpreted according to the classification scheme established by Jeffreys (1961). For all t-tests, Bayes Factors were calculated using a half-normal distribution for directional predictions, here referred to as $B_{H(0, x)}$, and a normal distribution for non-directional predictions, here referred to as $B_{(0, x)}$ where x is the SD of the distribution. These SDs were based on the results found by Macdonald & Lavie (2008; Experiment 3) and Raveh & Lavie (2015; Experiment 2), for load induced inattention blindness and deafness, respectively. For the effect of load on the multisensory CS, priors were selected based on whichever was larger from either of the two papers, for a more stringent test. Bayesian ANOVAs were run using the open source statistical software JASP (JASP Team, 2020). For all tests, p values are additionally reported for information, but all inferences are drawn only from Bayes Factors (for similar practice, see Skelton et al., 2017). Where the assumption of sphericity has been violated, p values based on Greenhouse-Geisser corrected degrees of freedom are reported, along with the epsilon value (ϵ).

Results

Trials in which the central letter search task response was incorrect were excluded from all RT and detection sensitivity analyses. RT was not greater than 1500ms for any trials, therefore no exclusions were made on this basis.

Central letter search task

Bayesian independent t-tests on mean reaction time (RT) and percentage error rate revealed decisive evidence that mean RT to the central search task was faster, and percentage error rate lower, under conditions of low perceptual load (RT, $M = 632$, $SD = 120.90$; error rate, $M = 8.69$, $SD = 6.87$) than high perceptual load (RT, $M = 841$, $SD = 111.42$; error rate, $M = 30.50$, $SD = 10.63$), $B_{H(0, 261)} = 94370.41$, $p < .001$; and $B_{H(0, 26)} = 3.60 \times 10^9$, $p < .001$ for RT and error rates, respectively.

CS Detection

Mean percentage detection rate, false alarm rate, mean d' , and response bias (β) were calculated as a function of perceptual load (Table 1). 2x3 mixed Bayesian ANOVAs were conducted on these measures, with the factors of load (high, low), and CS modality (unisensory visual, unisensory auditory, multisensory audiovisual). Substantial evidence was found for a main effect of perceptual load on d' , $B = 3.35$, $p = .020$, being reduced under high load as compared with low. Decisive evidence for a main effect of CS modality on d' was also found, $B = 130.98$, $p < .001$, with d' being highest for the multisensory CS, which is unsurprising given that they were experiencing both auditory and visual stimuli at the same time. However, there was no interaction between load and modality of the CS, with substantial evidence in support of

H_0 , $B = .17$, $p = .838$. Critically, follow up tests confirmed that sensitive, substantial evidence was found for the modulation of d' by load for all three CS types ($t(30) = 2.07$, $p = .023$, $B_{H(1.43)} = 3.18$ for multisensory CS; $t(30) = 2.27$, $p = .015$, $B_{H(0, 1.43)} = 4.23$ for auditory CS; $t(30) = 1.90$, $p = .034$, $B_{H(0, 1.12)} = 3.23$ for visual CS). As can be seen in Table 1, this modulation was of a similar magnitude for both auditory, visual and multisensory stimuli.

Strong evidence for a main effect of perceptual load on correct detection rate was revealed, $B = 19.62$, $p = .001$, with detection rate being lower under high load than low, and decisive evidence for a main effect of CS modality on correct detection rate $B = 37044.70$, $p < .001$, $\varepsilon = .76$. There was no sensitive interaction found between load and CS modality, $B = 1.68$, $p = .058$. Follow up tests confirmed that correct detection rate for both unisensory CS types was modulated by load ($t(30) = 3.39$, $p < .001$, $B_{H(0, 30)} = 99.35$ for auditory CS; $t(30) = 2.92$, $p = .003$, $B_{H(0, 31)} = 26.35$ for visual CS). There was no sensitive evidence for the effect of load on detection of the multisensory CS, though the trend is in the same direction, $t(30) = 1.06$, $p = .149$, $B_{H(0, 31)} = .45$.

No main effect of load was found on response bias (β), with anecdotal evidence in support of the null hypothesis, $B = .49$, $p = .291$. There was substantial evidence supporting a main effect of CS modality, $B = 6.22$, $p = .007$, with b to the multisensory CS being lower than the auditory CS ($t(32) = 2.96$, $p = .006$, $B_{(0, 3.8)} = 29.99$), and no interaction between this and load, $B = .40$, $p = .291$. The response bias to all CS did not differ as a function of load ($ps > .125$, $B_{(0, 3.8)} < 1.12$).

Load	CS Type	Detection Rate (%)	False Alarm Rate (%)	d'	β
High	Multisensory	86	11	2.61	2.56
	Auditory	58	4	2.01	6.15
	Visual	57	8	1.91	5.71
Low	Multisensory	91	7	3.24	2.79
	Auditory	79	7	2.60	4.77
	Visual	79	10	2.69	3.19

Table 1. Percentage detection and false alarm rates, d', and response bias (β) as a function of perceptual load and CS modality

As in previous research using a similar paradigm (Macdonald & Lavie 2008; Raveh & Lavie, 2015; Murphy & Dalton, 2016), the data was reanalysed with the incorrect central search task responses included. The pattern of results for the effect of load on the multisensory CS were unchanged. Though sensitivity was not quite reached for the main effect of load on d' ($B = 1.67, p = .049$), follow up tests confirmed that there was substantial evidence in favour of d' for the multisensory CS being modulated by load ($t(30) = 2.12, p = .021, B_{H(0, 1.43)} = 3.42$ for the multisensory CS. Evidence fell short of sensitivity for the auditory and visual CS types ($t(30) = 1.57, p = .063, B_{H(0, 1.43)} = 1.10$ and $t(30) = 1.49, p = .073, B_{H(0, 1.12)} = 1.68$, for the auditory and visual CS, respectively). Again, strong evidence for a main effect of load on correct detection rate was found, $B = 22.87, p = .001$. Decisive evidence for a main effect of CS modality on correct detection rate was found ($B = 283.15, p < .001, \varepsilon = .84$) as well as anecdotal evidence for this effect on d', $B = 1.92, p = .037, \varepsilon = .84$. No interaction between CS modality and load was found for either measure ($B = .19, p = .709$ and $B = .43, p = .248$, for d' and correct detection rate, respectively). As previously, there was very

strong evidence in favour of correct detection rate for both the auditory ($t(30) = 3.24$, $p = .001$, $B_{H(0, 30)} = 56.72$) and the visual ($t(30) = 2.72$, $p = .005$, $B_{H(0, 31)} = 14.57$) CS being modulated by load, but evidence for the modulation of the multisensory CS did not reach sensitivity ($t(30) = 1.23$, $p = .114$, $B_{H(0, 31)} = .55$).

Discussion

The present experiment provides the first demonstration of a load induced modulation of awareness occurring in two sensory modalities at the same time. Hence, alongside replicating prior findings of load-induced deafness and load-induced blindness (MacDonald & Lavie, 2008; Raveh & Lavie, 2015), we establish the new phenomenon of load-induced simultaneous deafblindness. As such our findings suggest that awareness of multisensory stimuli is not immune to the effects of perceptual load. Despite conditions favouring multisensory enhancement due to the weaker unisensory constituents, following the principle of inverse effectiveness (Senkowski et al., 2011; Stevenson et al., 2012; Stevenson & James, 2009), it appears that detection of multisensory stimuli is subject to perceptual capacity limits, and hence disrupted by a high perceptual load task to a similar degree to unisensory stimuli.

Our findings extend the literature on inattention blindness. Previous research has shown that an array of stimuli may be missed when our attention is engaged elsewhere, from an out of place gorilla (Simons & Chabris, 1999), to the strong smell of coffee in a room (Forster & Spence, 2018), but to date all research has, to our knowledge, concerned events occurring in one modality. Our research shows that it is possible to experience a simultaneous failure of awareness for stimuli that appear in more than one sense, and furthermore suggests that the Load Theory framework can be extended to these stimuli, with an increase in such failures under more perceptually demanding conditions.

Our finding of load modulation of detection sensitivity even for multisensory stimuli is consistent with our previous findings regarding load modulation of facilitatory multisensory attentional capture (Lunn et al., 2019), and further supports the idea that

multisensory stimuli do not have a ‘special’ status that makes them entirely immune to the effects of load. Our previous work demonstrated that perceptual load similarly reduced the speed with which salient multisensory and unisensory cues can be detected – our present work extends these findings by demonstrating that perceptual load also modulates awareness of low salience multisensory stimuli.

Given the need to use identical stimuli for the multisensory CS and its unisensory counterparts, the multisensory CS was by its nature more salient (reflected in higher baseline detection sensitivity in the low load condition). As such, the overall detection sensitivity was necessarily higher for the multisensory CS versus unisensory CS (given that there was more signal present in the former condition versus the latter). However, we note that this difference in salience appears unlikely to have masked any differential load impact on multisensory versus unisensory stimuli for several reasons: First, as noted above our prior work did not find any differences in perceptual load effects on detection speed for highly salient multisensory versus unisensory stimuli. Furthermore, detection sensitivity for the multisensory CS in the present study is similar to that found in low load for the unisensory stimuli in previous studies (range 2.70-3.70, mean = 3.33 across the previous research, compared to 3.30 in the present experiment, see Table 2). A comparison of the effect sizes across our stimulus types and those in prior studies suggests that load effects are relatively consistent regardless of this variation in baseline detection sensitivity (Figure 1), with the effect size for the load modulation of detection sensitivity for the multisensory CS being within the range of the CIs from almost all (8/9) of the previously demonstrated effects.

Investigation	Experiment	CS Modality	High Load d'	Low Load d'
MacDonald & Lavie (2008)	Expt 1	Visual	1.30	3.70
	Expt 2	Visual	1.84	3.53
	Expt 3*	Visual	1.98	3.10
	Expt 4	Visual	1.61	3.38
	Expt 5	Visual	2.43	2.70
Raveh & Lavie (2015)	Expt 1	Auditory	1.34	3.36
	Expt 2*	Auditory	1.79	3.22
	Expt 3	Auditory	1.38	3.31
	Expt 4	Auditory	2.66	3.69
Current Investigation	Expt 1	Multisensory	2.61	3.24
	Expt 1	Visual	2.01	2.60
	Expt 1	Auditory	1.91	2.69

Table 2. d' as a function of perceptual load and CS modality for previous investigations of load induced inattentional blindness and deafness, and the present experiment.

Asterisks (*) denote experiments closest in design to the current investigation

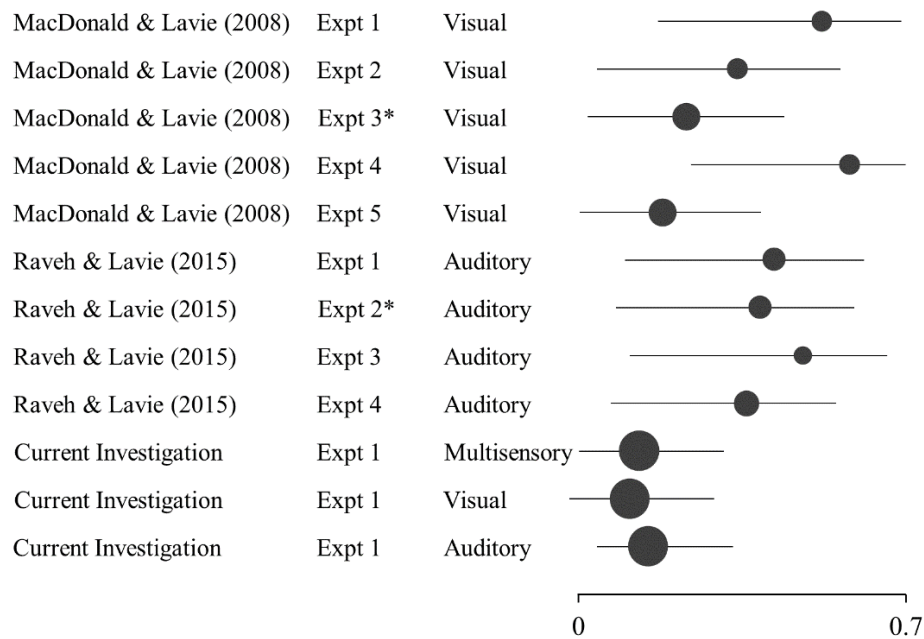


Figure 1. Forest plot depicting effect size (η^2_p) and 90% confidence intervals of load modulation of d' , for previous literature on load induced attentional blindness and deafness and the present experiment. Marker scaled according to sample size, asterisk denotes experiments most similar in procedure to the current investigation, effect sizes calculated using reported F values and degrees of freedom

The results of our research have not only theoretical implications, but also practical implications for real-world scenarios. Taken together with our prior investigation (Lunn et al., 2019), our findings suggest that making a warning signal or alert multisensory might increase its salience and hence confer some advantages (in terms of faster and more accurate detection). However, it is important to note that even multisensory stimuli are vulnerable to load-induced deafblindness, and as such might not be noticed during a perceptually demanding task, such as driving down a busy road.

In conclusion, this experiment provides the first demonstration of load induced deafblindness - demonstrating that both load induced inattentional blindness and inattentional deafness can occur concurrently to a multisensory stimulus. As such, our findings have theoretical implications both for proposals regarding the special status of multisensory stimuli, as well as applied implications for predicting the situations in which even multisensory signals might be missed.

Supplementary Materials

The following analyses are on data using of a 10% minimum detection rate for each type of CS. Trials in which the central letter search task response was incorrect were excluded from all RT and detection sensitivity analyses. RT was not greater than 1500ms for any trials, therefore no exclusions were made on this basis.

Central letter search task

Two independent t-tests on mean reaction time (RT) and percentage error rate revealed sensitive evidence that mean RT to the central search task was faster, and percentage error rate lower, under conditions of low perceptual load (RT, $M = 645$, $SD = 119.10$; error rate, $M = 9.54$, $SD = 6.80$) than high perceptual load (RT, $M = 844$, $SD = 114.70$; error rate, $M = 32.46$, $SD = 10.39$), $t(44) = 5.78$, $p < .001$, $B_{H(0, 261)} = 3.63 \times 10^6$ and $t(44) = 8.92$, $p < .001$, $B_{H(0, 26)} = 2.57 \times 10^{16}$ for RT and error rates, respectively.

CS Detection

Mean percentage detection rate, false alarm rate, mean d' , and response bias (β) were calculated as a function of perceptual load (Table S1). 2x3 mixed Bayesian ANOVAs were conducted on these measures, with the factors of load (high, low), and CS modality (unisensory visual, unisensory auditory, multisensory audiovisual). Anecdotal evidence was found for a main effect of perceptual load on d' , $B = 1.68$, $p = .050$, being reduced under high load as compared with low. Decisive evidence for a main effect of CS modality on d' was also found, $B = 631.55$, $p < .001$, with d' being highest for the multisensory CS, which is unsurprising given that they were experiencing both auditory and visual stimuli at the same time. However, there was no

interaction between load and modality of the CS, with strong evidence in support of H_0 , $B = .19$, $p = .552$. Follow up tests confirmed sensitive evidence in favour of load modulation of the multisensory CS, $t(44) = 2.20$, $p = .017$, $B_{H(0, 1.43)} = 4.34$ falling short of sensitivity for the visual and auditory CS, ($t(44) = 1.33$, $p = .096$, $B_{H(0, 1.43)} = .87$ for auditory CS; $t(44) = 1.62$, $p = .056$, $B_{H(0, 1.12)} = 1.85$ for visual CS).

Anecdotal evidence for a main effect of perceptual load on correct detection rate was revealed, $B = 2.13$, $p = .034$. As expected, decisive evidence for a main effect of CS modality was also found, $B = 66004.79$, $p < .001$, $\varepsilon = .86$, but critically again there was no interaction between load and CS modality, with substantial evidence in support of the null ($B = .15$, $p = .674$). Follow up tests confirmed sensitive evidence in favour of load modulation of the auditory CS, $t(44) = 2.08$, $p = .022$, $B_{H(0, 30)} = 3.74$ for auditory CS. The effect of load on correct detection of the visual and multisensory CS fell short of sensitivity, ($t(44) = 1.85$, $p = .035$, $B_{H(0, 31)} = 2.35$ for visual CS, $t(44) = 1.42$, $p = .082$, $B_{H(0, 31)} = 1.08$ for multisensory CS).

Again, substantial evidence against an effect of load was found on response bias (β), $B = .33$, $p = .722$. There was no sensitive evidence in support of a main effect of CS modality, $B = .96$, $p = .041$. Anecdotal evidence in support of a null interaction between CS modality and load was found, $B = .35$, $p = .249$. The response bias to all CS did not differ as a function of load ($ps > .314$, $B_{(0, 3.8)}S < .53$).

Load	CS Type	Detection Rate (%)	False Alarm Rate (%)	d'	β
High	Multisensory	72	11	2.18	3.00
	Auditory	48	5	1.68	5.53
	Visual	52	9	1.71	5.33
Low	Multisensory	82	7	2.89	3.90
	Auditory	64	8	2.08	4.84
	Visual	67	9	2.27	3.97

Table S1: Percentage detection and false alarm rates, d' , and response bias (β), as a function of perceptual load and CS modality

Chapter 5: Establishing the effects of multisensory perceptual load

Chapter Summary

In Chapter 2, I demonstrated that multisensory stimuli offer an advantage over unisensory stimuli when presented as secondary search targets in the context of both low and high unisensory visual perceptual load. Chapter 5 aims to explore the effects of a multisensory perceptual load, employing both behavioural and ERP measures, to determine whether multisensory stimuli may differ in their ability to fill perceptual capacity, compared with unisensory visual (Experiments 1-3) and unisensory auditory (Experiment 3) high load tasks. Additionally, this Chapter will test whether multisensory perceptual load eliminates the facilitatory effects of the multisensory secondary search targets, which would indicate that only one multisensory stimulus can be attentionally ‘boosted’ at a time. Where additional analysis will be included in the article submission as supplementary materials, these have been provided at the end of the chapter.

Abstract

It is now well established that high perceptual load of a primary task can powerfully disrupt processing of task irrelevant stimuli. However, while many daily life tasks involve more than one sensory modality at a time, prior demonstrations of perceptual load effects have been limited to unisensory forms of perceptual load, with the overwhelming focus of this research being on visual perceptual load. In the light of claims of a special attentional status for multisensory stimuli, the present research examined the effects of multisensory perceptual load. In three experiments, participants performed a primary unisensory or multisensory RSVP task with high perceptual load, while also performing a secondary task of detecting and making speeding responses to peripheral multisensory (audiovisual) and unisensory (auditory or visual) targets. Using a combination of behavioural measures of attentional capture with ERP indicators of multisensory integration, we consistently found that multisensory load – despite eliciting a non-linear neuronal response - did not differ from a high unisensory load in its impact on secondary task processing. Secondly, there was no evidence of non-linear ERP responses to the multisensory secondary task stimuli, despite a clear behavioural advantage in terms of faster detection even under multisensory load. Our results imply that any special status of multisensory attentional stimuli does not alter the powerful established effects of perceptual load on attention. We also highlight that behavioural multisensory advantages should not be assumed to always reflect neuronal superadditivity, yet may nevertheless be useful in applied contexts even during multisensory tasks.

Introduction

Every day, we are bombarded with an overwhelming amount of sensory input, including sights, sounds, tactile sensations, odours, and tastes. Decades of research have considered both which types of sensory stimuli are most likely to attract our attention, and the contexts under which this is more likely to occur. To address the former point, the types of stimuli argued to capture attention effectively include ‘singletons’ which differ in some unique attribute (e.g. colour) from surrounding items (Theeuwes, 1992), abrupt onsets (Jonides & Yantis, 1988), moving stimuli (Franconeri & Simons, 2003), or events that have motivational relevance or value (Anderson et al., 2011; Purkis et al., 2011). Load Theory, proposed by Lavie (1995) in order to resolve conflicts between early- and late-selection hypotheses (see Lavie, 2010, for a review), provides a useful theoretical framework for predicting when this capture may happen. The theory posits that attention works with a limited perceptual capacity, automatically processing stimuli until capacity is depleted. During tasks which involve only low perceptual load, spare capacity remaining after processing relevant information spills over to allow processing of other, less relevant, stimuli (and, hence, potential capture of attention by salient task-irrelevant stimuli). On the other hand, under high perceptual load conditions, all processing capacity must be fully devoted to the relevant task and therefore stimuli irrelevant to the primary task are typically not processed.

Load Theory has been supported by a large body of evidence using various different manipulations of load, and various measures of task-irrelevant processing. Increasing the load of a primary task has been shown to reduce behavioural interference from irrelevant distractors (e.g. Forster & Lavie, 2008; Lavie & Cox, 1997), decrease BOLD responses in the visual cortex for irrelevant peripheral stimulation (e.g. Schwartz et al., 2005), and reduce sensitivity to detect both auditory and visual peripheral stimuli

presented in the context of a secondary task (Macdonald & Lavie, 2008; Raveh & Lavie, 2015). The overwhelming majority of demonstrations in support of Load Theory have involved effects of load in a visual primary task on processing of visual stimuli, however there have also been a small number of demonstrations of crossmodal effects of visual perceptual load on other sensory modalities, for example visual load-induced inattention to auditory (Macdonald & Lavie, 2011; Raveh & Lavie, 2015), tactile (Murphy & Dalton, 2016), and olfactory stimuli (Forster & Spence, 2018). Recent studies have also examined the effect of auditory perceptual load. One extensive investigation using a range of perceptual load manipulations in an auditory primary task found no modulation of an auditory distractor effect (Murphy et al., 2013). On the other hand, Fairnie et al., (2016) more recently looked at the effect of auditory perceptual load on awareness, and found that detection of a critical stimulus was reduced when the number of sounds in an audio-spatial search task was increased, consistent with the perceptual load literature in the visual domain (see Murphy et al., 2017 for a review).

One type of primary perceptual load task that has yet to be considered is that which occurs in more than one sensory modality at a time. In our daily lives it is rare to encounter a purely unisensory task. For example, common everyday experiences such as having a conversation involve both auditory processing of speech and visual processing of facial expressions and mouth movements. However, it is as yet unknown how the perceptual load of such multisensory tasks affects processing of other, peripheral, stimuli. This question is particularly intriguing in the light of evidence that multisensory stimuli may have special attentional properties, argued to result from enhanced neuronal responses to multisensory stimuli, first demonstrated in the superior colliculus of cats (e.g. Meredith & Stein, 1983). Parallel results in humans in terms of non-linear ERP responses to multisensory stimuli have also been demonstrated (e.g.

Giard & Peronnet, 1999; Molholm et al., 2002; Murray et al., 2005; Senkowski et al., 2005; Talsma et al., 2007; Talsma & Woldorff, 2005). Behavioural evidence for the special attentional status of multisensory stimuli has also been observed, in terms of enhanced spatial cuing (see Spence & Santangelo, 2009 for a review) and visual search performance (e.g. Van der Burg et al., 2008). As with the single cell and ERP data, there is also behavioural evidence of enhanced responses to multisensory stimuli above and beyond that which would be predicted by race models of the two unisensory stimuli (Hughes et al., 1994; Laurienti et al., 2004; Lunn et al., 2019; Molholm et al., 2002; Murray et al., 2005; Pannunzi et al., 2014; though see Otto & Mamassian, 2012). Violations of the race model suggest that participants are not merely responding to the first stimulus detected, but that the integration of the two sensory stimuli facilitates response times. However, it is as yet unknown how these apparent special attentional properties of multisensory stimuli impact upon their ability to occupy and exhaust perceptual capacity.

One conceivable possibility is that multisensory stimuli could make a particularly engaging form of perceptual load: A superadditive response to a primary perceptual load task might conceivably boost the biasing of top down attention, thus reducing processing of task-irrelevant stimuli even more effectively than unisensory perceptual load. On the other hand, the facilitated processing afforded by multisensory stimuli could also plausibly have the opposite effect, lowering the perceptual load of the primary task and hence impeding any load effects on task-irrelevant processing.

A second, related, question is how a task with high multisensory perceptual load might impact the processing of other multisensory stimuli, in terms of potentially disrupting their integration and hence undermining their ‘specialness’. There has been

mixed evidence regarding unisensory perceptual load effects on processing of multisensory stimuli. Although it was initially suggested the multisensory stimuli were fully immune to perceptual load effects (Santangelo & Spence, 2007) more recent evidence reveals that two classic visual perceptual load manipulations - a central RSVP stream (Rapid Serial Visual Presentation; e.g. as in Bahrami et al., 2007) and the classic letter search task as used by Lavie and Cox (1997) - were equally effective in modulating responses to multisensory versus unisensory secondary targets (Lunn et al., 2019). Nevertheless, facilitatory effects of multisensory versus unisensory presentation were observed during both the high and low load conditions, with violations of the race model found under high load. Hence, the special attentional status of multisensory stimuli did not appear to be disrupted by a high unisensory perceptual load. However, the possibility that multisensory load might specifically disrupt task-irrelevant multisensory processing is raised by suggestions that audiovisual integration can occur for only one item at a time (Van der Burg et al., 2013a). The possibility that multisensory tasks might remove any enhanced effectiveness of multisensory stimuli in capturing attention has not only theoretical relevance but also important applied implications, given that multisensory stimuli have been suggested to be good candidates for alerts and warning signals. For these to be useful in situations such as whilst driving or piloting an aircraft, it is also vital that an attentional capture advantage persists under multisensory conditions.

The present research will hence test for the first time how a high load multisensory primary task affects attentional capture by multisensory and unisensory stimuli. Here we adapted a paradigm previously used by Lunn et al. (2019) to show facilitatory multisensory attentional capture by peripheral secondary task targets during a high versus low unisensory perceptual load RSVP task, creating both unisensory and

multisensory versions of the high perceptual load condition. Across experiments we combined behavioural measures of facilitatory attentional capture with ERP indicators of multisensory integration, allowing us to additionally examine the degree to which any behavioural multisensory advantage in either biasing attention to the primary task or capturing attention to peripheral targets might reflect superadditivity. If multisensory load is particularly effective in biasing top down attention to the primary task, we would expect slowed detection of both unisensory and multisensory peripheral targets during multisensory versus unisensory load. Conversely, if multisensory integration of the central RSVP places reduced demands on perceptual capacity, then we would expect to see that the peripheral targets are detected faster under multisensory versus unisensory load. In either scenario, we would expect a non-linear ERP response to multisensory versus unisensory perceptual load tasks, confirming integration of the primary task stimuli, and a non-linear ERP response to multisensory versus peripheral targets where enhanced behavioural multisensory attentional capture is observed. Finally, if only one multisensory stimulus can be integrated (and therefore attentionally ‘boosted’) at a time, then we might additionally expect to see selective effects of a multisensory primary task on the response to multisensory peripheral targets, which could manifest as a decrease in the behavioural multisensory advantage and lack of significant race model violations for peripheral targets under multisensory load, as compared with unisensory load, accompanied by the abolition of any non-linear ERP response.

Behavioural Experiment 1

To address the effects of multisensory perceptual load on attentional capture by concurrent peripheral stimuli, Experiment 1 adapted an established perceptual load manipulation (e.g. as used in Lunn et al., 2019) in which participants search a central RSVP stream for either a visual conjunction (colour and shape, high perceptual load), or a multisensory conjunction (colour and sound). While performing the primary task, participants were also asked to detect intermittent peripheral targets which could be either multisensory (audiovisual) or unisensory (visual). In this initial experiment, only unisensory visual targets were compared with multisensory, as the experiment upon which it is based found no difference between unisensory visual and unisensory auditory targets (Lunn et al., 2019, Experiments 1a and 1b). This paradigm allows us to address how engaging the multisensory load is by looking for either an enhancement or reduction of reaction times to the peripheral stimuli, as well as whether there is a particular effect when these stimuli are themselves also multisensory.

Methods

Participants

Power analysis conducted using G*Power software (Faul et al., 2009) revealed that a sample size of 20 would be highly powered to detect effect sizes equivalent to the effect revealed by Lunn et al. (2019, Experiment 1a) of an identical visual high load RSVP stream compared with visual low load on peripheral target response times ($\eta^2 = .67$; $\alpha = .001$; $1-\beta = .95$). Twenty participants (16 female) aged between 19 and 27 years ($M = 21.21$, $SD = 2.41$) were recruited at the University of Sussex. Data from one participant was excluded from analysis for failing to comply with the instructions. All

participants reported normal or corrected-to-normal vision and hearing. Both Bayesian and null hypothesis testing is reported given that the latter is more widely understood, but only the former provides a measure of evidence regarding whether the null or alternative hypothesis is supported by the data (Wagenmakers et al., 2017).

Stimuli and Procedure

The experiment was programmed and presented using E-prime v2.0, on a 17-inch Dell flat screen, placed 50cm from the participant's face, at eye level. Viewing distance was maintained using a chin rest. Loudspeakers positioned left and right of the screen were used to present sounds. Figure 1 presents an example trial sequence. Each trial began with a central fixation cross presented for 500ms, followed by a stream of nine coloured characters (each subtending $2.3^\circ \times 1.1^\circ$), presented centrally one at a time, with an interstimulus interval of 233ms. The central characters were either an S or a 5, and could be coloured red, green, yellow, blue, purple or turquoise. In all conditions they were each accompanied by either an 'aah' or a 'buzz' sound presented over both speakers. Participants were instructed to monitor the central RSVP stream for a target, defined by a conjunction of features. In the high unisensory visual load condition, the target was either a green 5 or yellow S, whereas in multisensory load the target was either a blue character with an 'ahh' sound, or purple character with a 'buzz' sound. Participants reported detection of the target with a foot pedal. Targets appeared as either the 3rd or 6th stimulus in a trial. The timing of presentation was irregular, to increase demand. This was achieved by randomising presentation time of each character (167, 267 or 367ms) with a fixed interstimulus interval (ISI) of 233ms. All stimuli were presented on a light grey background.

In addition to the central task, participants were asked to monitor for peripheral targets which appeared on 50% of trials presented to the left or right of the central stream. These were presented concurrently with a non-target central stimulus, and therefore did not interfere with responses to the central task. Participants were required to press the left button on a response box if the peripheral target was on the left, and the right button on a response box if it was on the right. Half of these peripheral targets were unisensory visual, and half were multisensory (i.e. each occurring on 25% of trials). The unisensory visual target was a black circle of 1.7° diameter (100ms), and the multisensory target was the black circle accompanied by a 'beep' sound (100ms, 1100Hz). The targets were presented on either the left or right of the screen, and the multisensory targets always had the circle and 'beep' on congruent sides.

Participants completed two blocks of 144 trials for each load condition, in the order ABBA or BAAB, which was counterbalanced between participants.

Participants completed two blocks of 144 trials for each load condition, in the order ABBA or BAAB, which was counterbalanced between participants.

Bayesian Analysis

Bayes factors (B) are reported for all one degree of freedom tests and planned contrasts, to assess the strength of the evidence for H1 relative to H0 (Wagenmakers et al., 2017). A B of above 3 is indicative of substantial evidence for H1, whereas a B of below 1/3 indicates substantial evidence for H1, and between these values indicates the data is insensitive (Dienes, 2014). Bayes factors were calculated using a half-normal distribution for directional predictions, here referred to as $B_{H(0, x)}$, and a normal distribution for non-directional predictions, here referred to as $B_{(0, x)}$ where x is the SD of the distribution. These SDs were based on the results found in Lunn et al. (2019) Experiment 1a.

Results

Central Task

Reaction Time. Reaction times (RTs) to the central task (correct responses only) were significantly slower under multisensory load ($M = 841$, $SD = 128$) than under unisensory high load ($M = 748$, $SD = 80$), $t(18) = 3.78$, $p = .001$, $B_{(0, 210)} = 132.04$.

Error. Percentage error rates to the central task were significantly higher under multisensory load ($M = 23$, $SD = 12.56$) than under unisensory high load ($M = 12.95$, $SD = 10.49$), $t(18) = 3.98$, $p < .001$, $B_{H(0, 10)} = 407.17$.

Peripheral Task

Reaction Time. Correct RTs to the peripheral targets were entered into a 2 x 2 within-subject ANOVA with the factors of load (multisensory, high unisensory visual) and peripheral target modality (multisensory, unisensory visual). As seen in Figure 2,

there was a main effect of peripheral target modality, $F(1,18) = 131.00$, $p < .001$, $\eta^2 = .88$, $B_{H(0,52)} = 2.63 \times 10^{27}$. Replicating our prior findings (Lunn et al., 2019), this reflected faster detection of multisensory versus unisensory peripheral targets. However, there was no main effect of load ($p = .10$, $\eta^2 = .14$, $B_{(0,49)} = .90$) – overall RTs to peripheral targets did not differ dependent on whether the central task was high unisensory load or multisensory load ($t(18) = 1.64$, $p = .119$, $B_{(0,47)} = .69$ for multisensory targets; $t(18) = 1.68$, $p = .110$, $B_{(0,52)} = .94$ for unisensory visual targets) – and no significant interaction between load and target modality ($p = .32$, $\eta^2 = .05$, $B_{(0,49)} = .25$).

In order to assess the degree to which effects of the two high load conditions in the present experiment (i.e. unisensory and multisensory) differed from those of a closely matched low load condition as used in Lunn et al. (2019), we conducted an additional analysis comparing our present results to those of the prior study (see Figure 2). A mixed ANOVA between the two experiments, with the within subjects factor of peripheral target type (multisensory, unisensory visual) and between subjects factor of central load type (multisensory, low unisensory visual load from Lunn et al., 2019) revealed a main effect of peripheral target type, $F(1,37) = 108.55$, $p < .001$, $\eta^2_p = .75$, with RTs to multisensory peripheral targets being faster than those to visual peripheral targets. The main effect of load on peripheral target RT did not quite reach significance, $F(1,37) = 3.78$, $p = .059$, $\eta^2_p = .09$, but planned follow up t-tests indicate statistically significant and sensitive evidence that RTs to both types of peripheral target were slower under the multisensory load of the present study compared with the low unisensory visual load condition, indicating load modulation akin to that which would be expected with a unisensory visual high and low load comparison ($t(37) = 1.93$, $p = .031$, $B_{H(0,47)} = 3.58$ and $t(37) = 1.90$, $p = .033$, $B_{H(0,52)} = 3.61$ for multisensory and

unisensory visual peripheral targets, respectively). There was no significant interaction between load and target type ($p = .283$, $\eta^2_p = .03$).

A similar ANOVA but with the between subjects factor of central load type comparing the visual high load from the present experiment with the visual low load from Lunn et al. (2019) also revealed a main effect of peripheral target type, $F(1,37) = 102.84$, $p < .001$, $\eta^2_p = .74$, with RTs to multisensory targets being faster than to unisensory visual targets. There was a significant main effect of load, $F(1,37) = 5.47$, $p < .001$, $\eta^2_p = .13$, but no interaction between this and peripheral target type ($p = .106$, $\eta^2_p = .07$), with RTs to both multisensory and unisensory visual targets being significantly faster under the low load condition than high ($t(37) = 2.27$, $p = .015$, $B_{H(0,47)} = 6.82$ and $t(37) = 2.34$, $p = .012$, $B_{H(0,52)} = 7.59$ for multisensory and unisensory visual peripheral targets, respectively). No reaction time differences were observed between either the multisensory load or high visual load conditions from the present paper, and the high visual load condition from Lunn et al., 2019, for either of the peripheral target types ($t < 1$, $B_{(0,47)} = .45$ and $t < 1$, $B_{(0,47)} = .48$ for multisensory peripheral targets, comparing high unisensory visual load from Lunn et al., 2019 with multisensory and high unisensory visual load from the present experiment, respectively; $t < 1$, $B_{(0,52)} = .43$, and $t < 1$, $B_{(0,52)} = .58$, for the same comparisons on unisensory visual peripheral targets).

Error. A 2 x 2 within-subject ANOVA with the factors of central task load (multisensory, high unisensory visual) and peripheral target modality (multisensory, unisensory visual), revealed a main effect of peripheral target modality, $F(1,18) = 10.39$, $p = .005$, $\eta^2 = .37$, $B_{H(0,4)} = 65.54$, with lower error in detection of multisensory peripheral targets than unisensory visual (Table 1). There was no main effect of load (p

$= .192$, $\eta^2 = .09$, $B_{(0,3)} = 1.12$), nor a significant interaction between load of the central task and peripheral target modality ($p = .194$, $\eta^2 = .09$, $B_{(0,3)} = .02$).

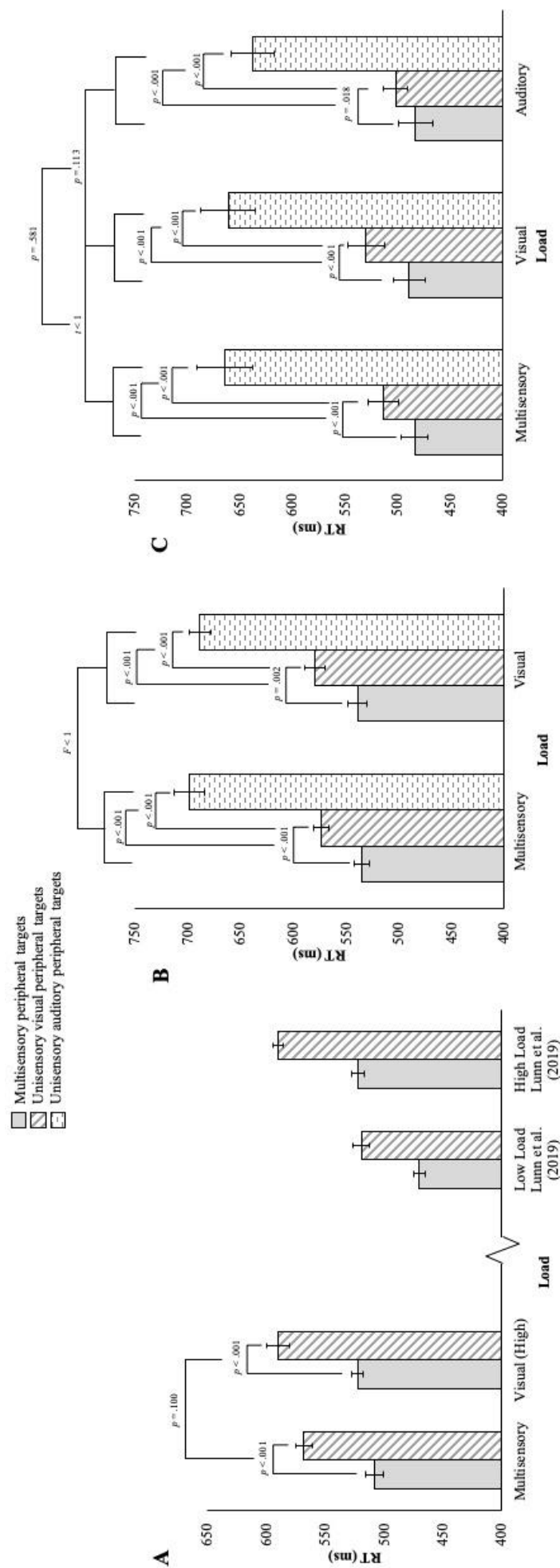


Figure 2. Mean reaction time (ms) to peripheral multisensory and unisensory targets, as a function of central load type, for A. Behavioural Experiment 1 (high and low visual load from Lunn et al. (2019) Experiment 1a is presented for comparison with MS load); B. Behavioural Experiment 2; C. Experiment 3. Error bars show +/- 1 SEM with Cousineau-Morey correction (Cousineau, 2005; Morey, 2008).

		Multisensory PTs	Visual PTs	Auditory PTs
Expt 1	Multisensory Load	11.68	18.00	
		(9.82)	(13.07)	
	Visual Load	9.00	12.21	
		(7.33)	(9.04)	
Expt 2	Multisensory Load	7.00	10.29	18.21
		(6.25)	(16.49)	(11.38)
	Visual Load	6.07	8.14	14.14
		(6.40)	(7.66)	(8.69)
Expt 3	Multisensory Load	14.00	13.75	26.38
		(14.53)	(13.18)	(4.15)
	Visual Load	9.13	11.69	23.69
		(8.41)	(10.20)	(13.01)
	Auditory Load	8.38	10.69	21.63
		(7.33)	(8.88)	(15.30)

Table 1. Mean percentage error rates (SD in parentheses) as a function of load and peripheral target (PT) type, across experiments 1-3.

Discussion

This experiment replicated our prior finding of multisensory stimuli enhancement of attentional capture by stimuli which are part of the top-down attentional set (Lunn et al., 2019), in terms of speeded responses to multisensory versus unisensory peripheral targets. More importantly, this advantage for multisensory stimuli speeded responses to multisensory peripheral target persisted even under conditions of multisensory perceptual load. Indeed, we did not find any differential impact of a high perceptual load task that was unisensory versus multisensory. However, our Bayesian

analyses failed to find sensitive support for the null hypotheses of no difference between multisensory versus unisensory perceptual load on peripheral target detection (although the load by peripheral target modality interaction showed a sensitive null result). For this reason, we conducted a second behavioural experiment to replicate our results, using a stopping rule that would afford sensitivity to differentiate the alternative and null hypotheses.

Behavioural Experiment 2

As mentioned above, the results of Experiment 1 did not reveal any differential effects of multisensory versus unisensory perceptual load on peripheral target responses, with the advantage for multisensory peripheral targets persisting in both conditions. This might initially appear to challenge claims that the capacity for multisensory stimuli is limited to one item (Van der Burg et al., 2013a). However, we note that performance of the central RSVP task was unexpectedly poorer, in terms of slower reaction time and higher error rate, for the multisensory versus visual unisensory task. As this is the opposite of what might be expected if multisensory integration facilitated the central RSVP task, this unexpected finding raises the question of whether multisensory integration was achieved in the primary task (as opposed to simultaneous but independent monitoring of two streams in separate modalities).

In the following experiment, we therefore sought to design a central task of equal difficulty between multisensory and visual load that would also maximise the potential for multisensory integration by using multisensory conjunction targets with established crossmodal correspondences of auditory pitch and visual elevation (e.g. Chiou & Rich, 2012; Evans & Treisman, 2010). In this RSVP paradigm (based on Bahrami et al., 2007), participants were presented with a stream of upright and inverted crosses, combined with either colour (visual load), or pitch (multisensory load). Participants were required to search for an infrequent ‘mismatch’ between the two features.

Additionally, given that the null effect of load modality on RT to peripheral targets did not fall below the Bayes Factor threshold of .30, which would have indicated substantial evidence in favour of the null hypothesis, we adopted a Bayesian optional

stopping rule and inference criteria for the present experiment. Unlike NHST, where a non-significant result cannot distinguish between a true no-difference between conditions versus a lack of sensitivity to measure a difference, Bayesian analysis allows us to conclude that the null hypothesis can be accepted (Dienes, 2014).

We also included both unisensory auditory and unisensory visual, as well as multisensory, peripheral targets, in order to be able to test whether the response to a multisensory stimulus was greater than that which would be predicted by the summed probability of the two unisensory stimuli by testing violations of the race model, which would suggest neural integration of the two sensory stimuli (Miller, 1982, 1986).

Methods

Participants

14 participants (11 female) aged between 18 and 23 years ($M = 20.07$, $SD = 1.59$) were recruited at the University of Sussex. Participants either gained course credits, or were paid, to take part. All participants reported normal or corrected-to-normal vision and hearing. The apriori stopping rule for this experiment only was based on Bayes Factors for the main effect of load and peripheral target type, and all planned comparisons, on reaction time analyses of both the central and peripheral task reaching sensitivity (see Rouder, 2014).

Stimuli and Procedure

The experiment was programmed and presented using E-Prime v2.0, on a 16inch screen. A viewing distance of 57cm was maintained using a chin rest. A speaker was positioned on the left and the right side of the screen. Figure 3 presents example trial sequences. Each trial began with a central fixation dot presented for 500ms, followed by

a continuous stream of upright and inverted crosses presented centrally. 240 crosses were presented in each block, each for 250ms, with a blank screen ITI also lasting 250ms. On 10% of the trials, a central target was presented, and on 7.5% of trials a peripheral target was presented. Participants were instructed to press the spacebar when they saw a central target, and to indicate using two keys on which side a peripheral target was presented. One third of peripheral targets (i.e. presented on 2.5% of trials) were visual only, consisting of a white circle of 1.7° diameter presented to either the left or right of the display (100ms). One third of peripheral targets were auditory only, consisting of a ‘beep’ sound from the left or right speaker (100ms, 1100Hz). The final third were multisensory targets, where the circle and ‘beep’ were presented simultaneously from the same side. The central task involved either visual or multisensory perceptual load. The unisensory visual high load central task involved looking for an incongruency in a regular display of stimuli. 90% of the central crosses followed the rule that yellow crosses are upright, and purple crosses are inverted. The 10% of crosses that mismatched this rule were the targets (i.e. yellow inverted and purple upright crosses, 1.30 and 2.66 relative luminance respectively, based on the CIELAB colour space, measured using the SHINE_colour toolbox for MATLAB; Dal Ben, 2019). In the multisensory central load, crosses were always blue (1.79 relative luminance), and were presented with a high or low pitch ‘ahh’ sound (high pitch was a vocal note E₄, low pitch was vocal note C₄, 58dB). 90% of the central crosses followed the rule that upright crosses play with a high pitch sound, and inverted crosses play with a low pitch sound. Again, the 10% of crosses that mismatched this rule were the targets (i.e. upright crosses with a low pitch sound, inverted crosses with a high pitch sound).

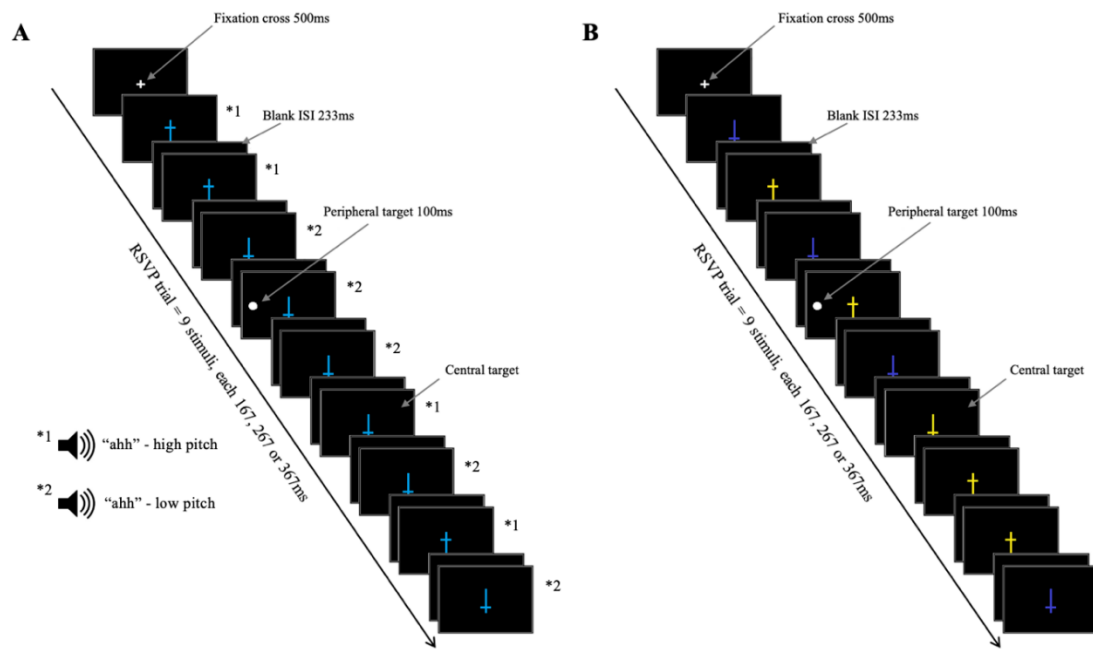


Figure 3. Example trial presentations. A. multisensory central load with unisensory visual peripheral target, B. unisensory visual central load with unisensory visual peripheral target

Participants first completed practice blocks for each of the central load tasks; first responding to the central targets only, then peripheral targets only, before combining both tasks in a single block. They then completed 8 blocks in total, 4 of each different central load task, in the order ABBAABBA or BAABBAAB, which was counterbalanced between participants.

Bayesian Analysis

Bayes Factors were calculated to assess the strength of evidence of for H1 relative to H0 (Wagenmakers et al., 2017). The strength of the evidence for either hypothesis is interpreted according to the classification scheme in Jeffreys (1961). For all t-tests, Bayes Factors were calculated using a half-normal distribution for directional

predictions, here referred to as $B_{H(0, x)}$, and a normal distribution for non-directional predictions, here referred to as $B_{(0, x)}$ where x is the SD of the distribution. SDs were calculated based on Experiment 1 of this paper for central and peripheral task reaction times and error rates, and from Lunn et al., (2019) Experiment 2 for the race model analysis. Given the Bayesian stopping rule, all inferences are drawn from Bayes Factors only. Bayesian ANOVAs were run using the open source statistical software JASP (JASP Team, 2020). For all tests, inferences are drawn only from Bayes Factors given the Bayesian stopping rule, but due to greater familiarity with NHST these results were also reported. For similar practice, see Skelton et al. (2017).

Results

Central Task

Reaction Time. RTs to the central task (correct responses only) were not significantly different under multisensory load ($M = 676$, $SD = 73$) compared to under unisensory load, with the Bayesian analysis confirming substantial evidence for the null hypothesis ($M = 658$, $SD = 66$), $t < 1$, $p = .345$, $B_{(0,93)} = .30$.

Error. Percentage error rates to the central task were similar in multisensory load ($M = 18.36$, $SD = 20.22$) and unisensory load ($M = 20.21$, $SD = 20.29$), $t(13) = 1.50$, $p = .158$, $B_{(0,10)} = .37$, with the difference favouring the null hypothesis but falling just short of a sensitive null.

Peripheral Task

Reaction Time. Correct RTs to the peripheral targets were entered into a 2 x 3 within-subject ANOVA with the factors of central task load (multisensory, high unisensory visual) and peripheral target modality (multisensory, unisensory visual,

unisensory auditory). As seen in Figure 2 and consistent with our first experiment and our prior study (Lunn et al., 2019), there was a decisive evidence for a main effect of peripheral stimulus modality, $F(2,24) = 97.59$, $p < .001$, $\eta^2 = .88$, $B = 2.28 \times 10^{21}$, with multisensory peripheral targets being detected faster than both unisensory visual targets, $t(13) = 4.94$, $p < .001$, $B_{H(0,68)} = 38794.17$, and unisensory auditory targets, $t(13) = 11.20$, $p < .001$, $B_{H(0,68)} = 5.50 \times 10^{25}$, and unisensory visual targets being detected faster than unisensory auditory targets, $t(13) = 9.61$, $p < .001$, $B_{(0,68)} = 4.96 \times 10^{18}$. Critically, there was no main effect of load ($F < 1$, $p = .99$, $\eta^2 = .002$, $B = .21$), nor an interaction between peripheral stimulus modality and central load type ($F < 1$, $p = .63$, $\eta^2 = .04$, $B = .22$), with substantial evidence for the null hypothesis in both cases. RTs to all peripheral target types did not differ depending on whether the central task was unisensory visual or multisensory load ($t < 1$, $p = .718$, $B_{(0,93)} = .13$ for MS targets; $t < 1$, $p = .590$, $B_{(0,93)} = .14$ for USV targets, and $t < 1$, $p = .511$, $B_{(0,93)} = .21$ for USA targets), with Bayesian analysis confirming sensitivity of the null results in both cases. Additionally, there was decisive evidence for the multisensory advantage in both load conditions, with RTs to multisensory peripheral targets being faster than to both unisensory visual and unisensory auditory targets, under conditions of multisensory load ($t(13) = 8.11$, $p < .001$, $B_{H(0,68)} = 7.78 \times 10^{12}$ and $t(13) = 4.42$, $p < .001$, $B_{H(0,68)} = 3782.01$ for multisensory compared with unisensory auditory and visual, respectively) and unisensory visual load ($t(13) = 10.07$, $p < .001$, $B_{H(0,68)} = 4.57 \times 10^{20}$ and $t(13) = 3.54$, $p = .002$, $B_{H(0,68)} = 146.48$ for multisensory compared with unisensory auditory and visual, respectively).

In order to further determine whether multisensory stimuli provide a detection speed advantage consistent with integration, we used the race model (Miller's inequality, Miller, 1982). The race model allows us to investigate whether the RTs in

the multisensory condition exceed the statistical facilitation predicted by probability summation based on two independent unisensory signals. In this model, a theoretical cumulative density function (CDF) is calculated based on the reaction time CDFs of each of the two unimodal stimulus types – F_x and F_y – and the redundant-stimulus, or multisensory, condition, F_z . The race model inequality

$$F_z(t) \leq F_x(t) + F_y(t), t > 0,$$

is examined for every value of t . Where the empirical CDF towards multisensory stimuli is greater than the theoretical CDFs based on the two unisensory components (tested using Bayes Factors), the reaction time advantage can be assumed to be caused by integrative effects. Analyses were carried out using the RMITest software, which applies the algorithm in Ulrich et al. (2007), and then Bayes Factors were calculated for each percentile of the RT distribution.

The results show substantial evidence that RT to multisensory stimuli is faster than the race model bound for three of the fastest percentiles under multisensory load ($B_{H(0,26)}S > 3.18$), and five under unisensory visual load ($B_{H(0,26)}S > 20.43$) (Figure 4). This is consistent with the assumption that the detection time advantage seen in both load conditions might result from multisensory integration.

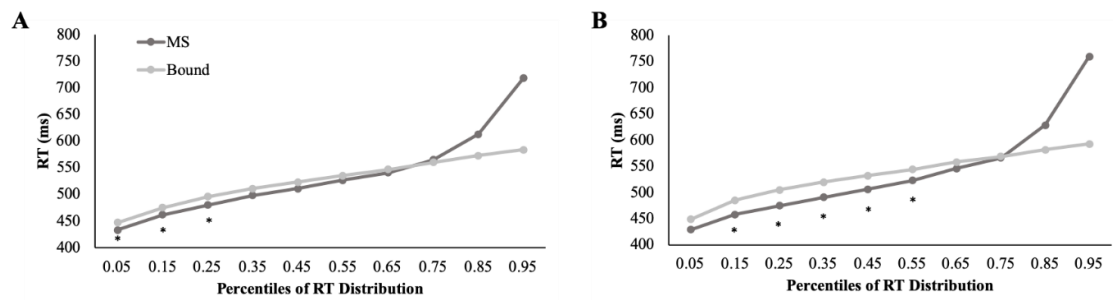


Figure 4. Cumulative probability distribution of reaction times for detection of multisensory (MS) peripheral stimuli, with race model bound for the two unisensory peripheral stimuli predicted by RMITest for A. multisensory central load, B. unisensory visual central load. Asterisks (*) refer to where sensitive evidence was obtained for violation of the race model inequality, based on Ulrich et al., (2007) algorithm

Error. A 2 x 3 within-subject ANOVA with the factors of load (multisensory, high unisensory visual) and peripheral target modality (multisensory, unisensory visual, unisensory auditory) on percentage error rates, revealed anecdotal evidence for a null effect of load ($p = .253$, $\eta^2 = .10$, $B = .51$), with error rates towards peripheral targets being similar regardless of central task type (Table 1). Decisive evidence for a main effect of peripheral target modality on error was found, $F(2, 26) = 7.28$, $p = .003$, $\eta^2 = .36$, $B = 245.92$, with less errors made in detecting multisensory peripheral targets than unisensory auditory peripheral targets $t(13) = 4.64$, $p < .001$, $B_{H(0.5)} = 59.15$, but no sensitive evidence for a difference in errors made detecting multisensory and unisensory visual peripheral targets ($p = .871$, $B_{H(0.5)} = 1.20$), or detecting unisensory visual and unisensory auditory peripheral targets ($p = .145$, $B_{(0.5)} = 2.89$). There was strong evidence in support of no interaction between load and distractor type ($F < 1$, $p = .597$, $\eta^2 = .04$, $B = .21$).

Discussion

In the present experiment, there were no differences in reaction time or error rates for detection of central targets, indicating that the multisensory and unisensory visual tasks were of comparable difficulty. Despite this, as in Experiment 1, the multisensory central task again neither positively or negatively affected reaction time and error rates for detection of peripheral stimuli, with sensitive evidence for the null hypothesis of no difference between load conditions. Furthermore, multisensory peripheral targets once again elicited enhanced attentional capture over unisensory targets consistently across both multisensory and unisensory load tasks, with a race model analysis supporting integration of multisensory peripheral targets.

Experiment 3

In Experiment 2, we again found no difference in either reaction time or error rates in detection of peripheral stimuli during either a multisensory central load task or a unisensory visual central load task, despite the two tasks being of equal difficulty (reflected in the equal RTs to respond and error rates). Hence, the results of these experiments appear to suggest that multisensory presentation of a primary task does not in any way impact the mechanism of perceptual load effects on attention. Furthermore, the persistence of facilitatory multisensory attentional capture in both load conditions appears inconsistent with prior claims that the capacity of audiovisual integration is limited to a single item (Van der Burg et al., 2013a). Our race model analysis further supports the notion that our peripheral multisensory targets were integrated. The present experiment employed ERP methods to provide a further confirmation of multisensory integration of the peripheral task, as well as confirming multisensory integration in the primary task. To this end, we added a third load condition of unisensory auditory load and employed the methods established by Giard & Peronnet (1999) to test for nonlinear effects attributed to multisensory integration, by comparing the sum of the two unisensory responses to responses to the multisensory stimuli ($AV-(A+V)$).

Methods

Participants

18 participants (14 female) aged between 18 and 24 years ($M = 20.88$, $SD = 1.82$) were recruited at the University of Sussex to participate in this experiment. 2 were excluded from analysis for failing to detect over 50% of the auditory targets. Therefore, data from 16 participants (12 female) were included. Participants either gained course

credits, or were paid, to take part. The study was approved by the Sciences and Technology Research Ethics Committee (C-REC) at the University of Sussex. All participants reported normal or corrected-to-normal vision and hearing, and no known skin, neck or head problems. Sample size calculations were conducted prior to data collection, using G*Power software (Faul et al., 2009), revealing that to detect an effect size of $\eta^2p = .23$ ($\alpha = .05$, $1-\beta = .80$), a sample size of 15 was required to detect the frontal multisensory integration effect. The effect size was taken from the integration effect in Talsma et al. (2007).

Stimuli and Procedure

The experiment was programmed and presented using E-Prime v2.0. Participants were sat with a viewing distance of approximately 57cm from a 19 inch CRT monitor, with a resolution of 1600 x 1200 and a refresh rate of 85Hz. Auditory stimuli were delivered through speakers positioned on the left and the right side of the screen. Participants completed 12 blocks of 70 trials in total, 4 blocks of 3 different central load types. The visual and multisensory load blocks were the same as described in Experiment 2. In the auditory load blocks, 90% of the central stimuli followed the rule that the verbal sound ‘ee’ is in a high pitch tone (250ms, note F₄), and the verbal sound ‘ooh’ is in a low pitch tone (250ms, note C₄). The 10% of stimuli that mismatched this rule were the targets (i.e. a low pitch ‘ee’ and a high pitch ‘ooh’). Auditory stimuli for the multisensory and auditory load blocks were presented at 56dB. Participants first completed practice blocks for each of the central load tasks; first responding to the central targets only, then peripheral targets only, before combining both tasks in a single block. The experimental block order was fully counterbalanced between participants.

Bayesian Analysis

Bayes factors were calculated using a half-normal distribution, with SDs based on Experiment 2 of this paper.

ERP Recording and Analysis

Electrical brain activity was continuously digitized using a 64 channel ANT Neuro amplifier and a 1000 Hz sampling rate. Horizontal and vertical eye movements were also recording using channels placed at the outer canthi and inferior orbits of the eye. Data processing was conducted using ERPlab (Lopez-Calderon & Luck, 2014). Following referencing to the average of the left and right mastoids, and using a 0.1Hz (12 db/oct; zero phase) high-pass and 30Hz (24 db/oct; zero phase) low-pass filter, and an 85Hz notch filter (to remove line noise caused by the speakers, visual inspection of the data confirmed that the notch filter did not create or remove components) epochs were baseline correct according to a 100ms pre-stimulus presentation window and neural activity was examined for 500ms post-stimulus presentation. Automatic offline artefact rejection was performed by removing epochs contaminated by eye blinks ($\pm 75\mu\text{V}$), eye movements ($\pm 30\mu\text{V}$), drifts, blocked electrodes or muscle-related potentials ($\pm 200\mu\text{V}$) - on average 22.38% of trials.

In contrasts of multisensory stimuli versus unisensory sum stimuli (AV-(A+V)), any overlapping ERP components would be subtracted twice (see Teder-Sälejärvi et al., 2002). Thus, having baseline corrected the ERPs to -100 to 0ms before stimulus onset, we tested whether the AV and the A+V waveforms differed significantly from each other just before stimulus presentation. Whilst we note that this cannot conclusive rule out potential double subtraction, a time window of -20 to 0ms was chosen for this analysis (as in Talsma et al., 2005), where no evoked activity could possibly be from the

current stimulus, but could contain overlapping ERP from the preceding stimulus in the RSVP stream. For the central task, an ANOVA was conducted on the average baseline activity in each of the three clusters relevant to our analyses (outlined below), with the factors of integration (multisensory, unisensory sum), stimulus type (target, non-target), and electrode cluster (frontal, fronto-central, parietal-occipital). For the secondary task, ERP responses to central non-targets were subtracted from each ERP response to a correctly identified peripheral target and non-target presented concurrently, to give the ERP elicited only by the peripheral target, therefore also removing any overlapping activity given the random and less frequent presentation of these targets. An ANOVA was conducted on mean baseline activity with the factors of integration (multisensory, unisensory sum), load (multisensory, unisensory visual, unisensory auditory), and electrode cluster (frontal, fronto-central, parietal-occipital).

Central Task. ERPs to unisensory visual and unisensory auditory central targets and central non-targets were summed, in order to compare with multisensory central targets/non-targets (correct responses only), to determine whether these elicited a non-linear ERP response (Giard & Peronnet, 1999). Evidence for an early frontal integration effect was analysed using an ANOVA on six frontal electrodes (Fp1, Fp2, AF3, AF4, F3, F4), chosen based on those used to measure this effect by Talsma and Woldorff (2005), using mean amplitude from 80-120ms.

Additionally, as an exploratory analysis to detect reliable differences between ERPs to multisensory (AV) and unisensory sum (A+V) at parietal and occipital scalp sites implicated in the sensory responses, the ERPs from these conditions were submitted to a repeated measures, two-tailed permutation test based on the tmax statistic (Blair & Karniski, 1993) using a family wise alpha level .05. All time points between

100 and 300ms, at 12 parietal-occipital (P7, P5, PO7, PO5, PO3, O1, P8, P6, PO8, PO6, PO4, O2) and 9 fronto-central electrodes (F1, Fz, F2, FC1, FCz, FC2, C1, Cz, C2) were included in the test (i.e. 4221 comparisons for each analysis), chosen a priori on the basis of where we would expect to see visual and auditory evoked potentials. Repeated measures t-tests were performed for each comparison, using the original data and 2500 random within-participant permutations of the data used to estimate the distribution of the null hypothesis. This permutation analysis provides a better spatial and temporal resolution than conventional ANOVAs, whilst correcting for the large number of comparisons. The tmax statistic was chosen because it has been shown to have relatively good power for data where the dimensions are highly correlated, as is the case for ERPs (Hemmelmann et al., 2004). 2500 permutations were used to estimate the distribution of the null hypothesis as it is over twice the number recommended by Manly (1997). All analyses were carried out using the Mass Univariate Analysis ERP Toolbox (Groppe et al., 2011). Based on the estimate of the null hypothesis, critical t-scores of ± 4.48 ($df = 15$) were derived for the analysis of central target data. In other words, any differences in the original data that exceeded a t-score of ± 4.48 were deemed reliable.

In addition to the early frontal integration effect, Talsma and Woldorff (2005) also found evidence of non-linearity at a cluster of centro-medial electrodes. The location of electrodes in the cluster used for this analysis, and those used for the mass univariate analysis to examine the auditory evoked potentials, has significant overlap, and the same pattern of results was identified in each. Given this overlap, we report only our more comprehensive mass univariate analysis in the main body of this paper, but for the purpose of facilitating comparison to prior work the additional analysis following Talsma and Woldorff's method is provided in the Supplementary Materials.

Peripheral Task. As with the central targets, ERPs to unisensory visual and unisensory auditory peripheral targets were then summed for each central load type, in order to compare with corresponding multisensory targets, in order to determine whether these elicited a non-linear ERP response (correct responses only).

Analyses were the same as for the central targets, examining evidence for an early frontal integration effect for each central load condition. Mass univariate analysis was used using the same criteria as previously outlined, to compare the AV and A+V ERPs for each load type and presentation side.

Results

Behavioural Results

Central Task

Reaction Time. RTs to the central task (correct responses only) were entered into a within-subject ANOVA with the factor of load modality (multisensory, unisensory visual, unisensory auditory), which revealed a significant main effect, $F(2, 30) = 6.92, p = .003, \eta^2 = .32, B = 11.25$. There were no significant differences in RT to multisensory ($M = 651, SD = 72.53$) and unisensory auditory ($M = 640, SD = 64.55$) central targets ($t < 1, B_{(0.93)} = .18$) but, in contrast to Experiment 2, responses to the unisensory visual load task ($M = 614, SD = 78.33$) were significantly faster than to either multisensory ($t(15) = 3.51, p = .003, B_{(0.93)} = 48.99$) or unisensory auditory load ($t(15) = 3.02, p = .009, B_{(0.93)} = 8.45$).

Error. Percentage error rates were entered into a within-subject ANOVA with the factor of load (multisensory, unisensory visual, unisensory auditory), which revealed

no significant main effect, $F(2, 30) = 2.30$, $p = .118$, $\eta^2 = .13$, $B = .74$ of the central task load, consistent with the results of Experiment 2.

Peripheral Task

Reaction Time. Correct RTs to the peripheral targets (Figure 2) were entered into a 3 x 3 within-subject ANOVA with the factors of central task load (multisensory, unisensory visual, unisensory auditory) and peripheral target modality (multisensory, unisensory visual, unisensory auditory). Consistent with Experiments 1 and 2, there was a main effect of peripheral stimulus modality, $F(2,30) = 161.50$, $p < .001$, $B = 3.62 \times 10^{41}$, $\eta^2 = .92$, but no main effect of load, $F(2, 30) = 2.67$, $p = .086$, $\eta^2 = .15$, $B = .83$, nor an interaction between load and peripheral target modality, $F < 1$, $\eta^2 = .06$, $B = .12$. Rather, the detection speed advantage for multisensory peripheral stimuli persisted during all three central load conditions, with RTs to multisensory targets being significantly faster than to both unisensory visual targets ($t(15) = 6.22$, $p < .001$, $B_{H(0,38)} = 4.84 \times 10^7$ under multisensory load, $t(15) = 6.56$, $p < .001$, $B_{H(0,40)} = 3.92 \times 10^8$ under unisensory visual load, and $t(15) = 2.31$, $p = .018$, $B_{H(0,40)} = 5.18$ under unisensory auditory load) and unisensory auditory targets ($t(15) = 9.77$, $p < .001$, $B_{H(0,163)} = 6.52 \times 10^{19}$ under multisensory load, $t(15) = 10.13$, $p < .001$, $B_{H(0,149)} = 2.25 \times 10^{21}$ under unisensory visual load and $t(15) = 16.38$, $p < .001$, $B_{H(0,149)} = 1.20 \times 10^{57}$ under unisensory auditory load), for every load type. Unisensory visual targets were also detected significantly faster than unisensory auditory targets, under central task conditions of multisensory load, $t(15) = 8.25$, $p < .001$, $B_{H(0,125)} = 8.19 \times 10^{13}$, unisensory visual load, $t(15) = 7.56$, $p < .001$, $B_{H(0,125)} = 4.06 \times 10^{11}$, and unisensory auditory load, $t(15) = 11.68$, $p < .001$, $B_{H(0,125)} = 4.45 \times 10^{28}$.

We again used the race model (Miller's inequality, Miller, 1982) to further determine whether multisensory stimuli provide a detection speed advantage consistent with integration. As in Behavioural Experiment 2, reaction time to multisensory peripheral targets was significantly faster than the race model bound for the fastest percentiles under multisensory and unisensory visual load (Figure 5), with Bayes Factors confirming the sensitivity of these results ($B_{H(0,30)S} > 20.52$ and $B_{H(0,15)S} > 41.91$, under conditions of multisensory and unisensory visual high load, respectively). However, this was not the case under unisensory auditory load ($p_s > .05$, $B_{H(0,30)S} < .33$ at all but three percentiles, where $B_{H(0,30)S} = .34-.62$).

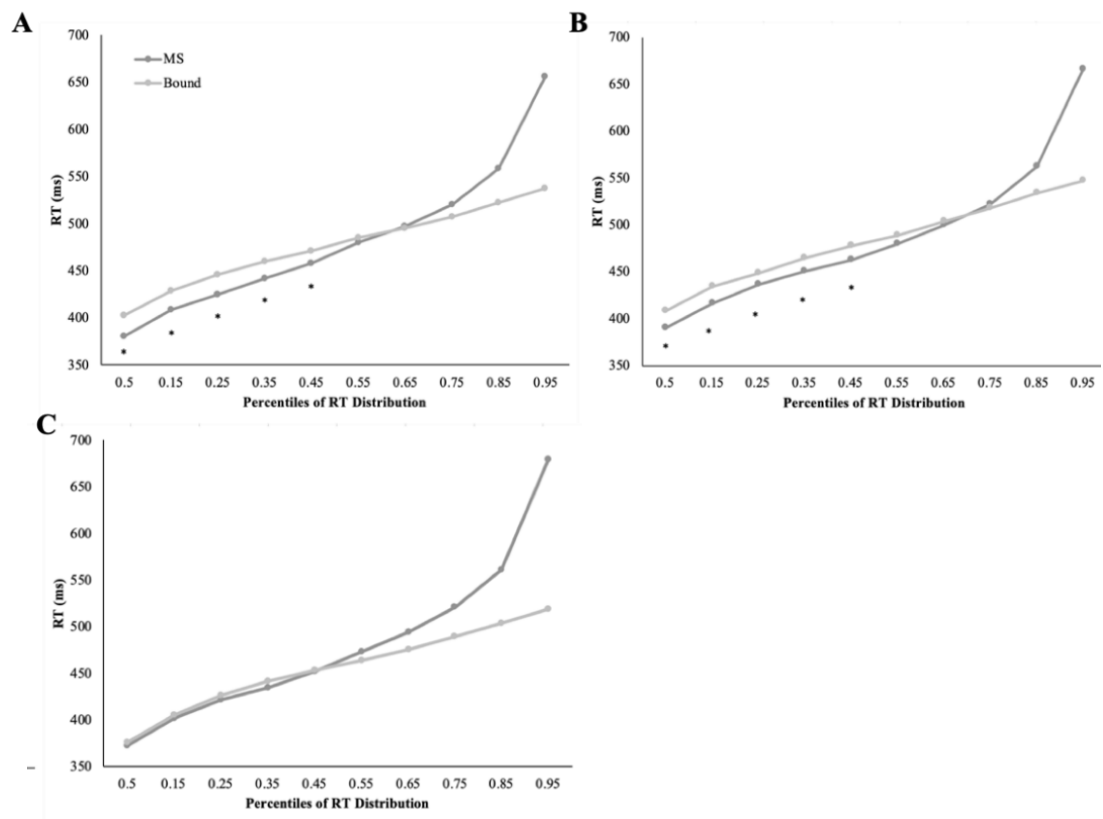


Figure 5. Cumulative probability distribution of reaction times for detection of multisensory peripheral stimuli, with race model bound for the two unisensory peripheral stimuli predicted by RMITest for A. multisensory central load, B. unisensory visual central load, and C. unisensory auditory central load. Asterisks (*) refer to where race model inequality was significantly violated, based on Ulrich et al., (2007) algorithm

Error. A 3 x 3 within-subject ANOVA with the factors of load (multisensory, unisensory visual, unisensory auditory) and peripheral target modality (multisensory, unisensory visual, unisensory auditory) on percentage error rates, revealed a main effect of peripheral target modality $F(2, 30) = 29.58, p < .001, \eta^2 = .66, B = 1.32 \times 10^{11}$ on error (Table 1). Significantly more errors were made in detecting unisensory auditory

peripheral targets than either multisensory peripheral targets $t(15) = 8.29, p < .001$, $B_{H(0,10)} = 9.61 \times 10^{13}$, or unisensory visual peripheral targets, $t(15) = 7.76, p < .001$, $B_{H(0,10)} = 1.62 \times 10^{12}$, but there was no difference in errors made detecting multisensory and unisensory visual peripheral targets ($p = .406$, $B_{H(0,10)} = .55$). As in Experiment 2, there was no main effect of load ($p = .126$, $\eta^2 = .13$, $B = 1.83$, nor a significant interaction between load and distractor type ($F < 1$, $\eta^2 = .03$, $B = .07$).

Overall the behavioural results are consistent with the findings of our previous Experiments in finding no difference between the effects of a multisensory versus unisensory primary task with high perceptual load on secondary task processing. Additionally, multisensory peripheral targets again elicited enhanced attentional capture over unisensory targets, consistently across all three central load tasks, and exceeding that which would be predicted by probability summation based on two independent unisensory signals under conditions of multisensory and visual load.

ERP Results

Central Task. Our initial analysis of baseline activity from -20 to 0ms confirmed that our ability to test for additive ERP responses to the central perceptual load task stimuli was not compromised by overlapping activity in the waveforms at any of the clusters used in our analyses, $F < 1$.

Frontal Integration Effect. The early frontal integration effect was analysed using an ANOVA on the frontal electrode cluster, using mean amplitude from 80-120ms, with the factors of integration (AV, A+V), and central stimulus type (target, non-target). Critically, as can be seen in Figure 6, a significant effect of integration was found, with AV stimuli showing greater frontal positivity than A+V stimuli, $F(1, 15) =$

6.55, $p = .022$, $\eta_p^2 = .30$. No significant difference was found between targets and non-targets ($F < 1$, $\eta_p^2 = .01$), nor was there any interaction between this and the factor integration ($p = .27$, $\eta_p^2 = .08$).

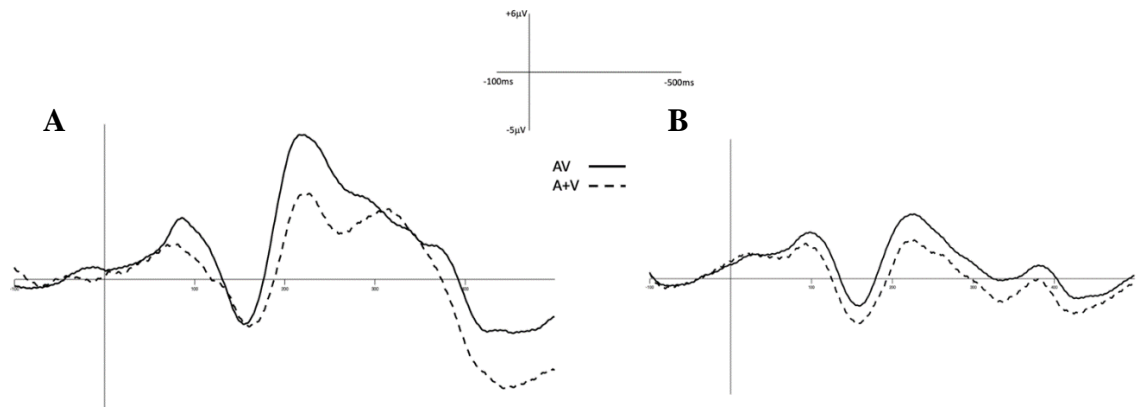


Figure 6. Averaged ERP at the frontal electrode cluster, showing greater early frontal positivity for A. AV central targets compared to A+V central targets, and B. AV central non-targets, compared to A+V central non-targets

Mass Univariate Analysis. Significantly greater parietal-occipital positivity between 199 and 250ms, and significantly greater frontal-central positivity between 208 and 253ms, was found for the AV ERP response, compared with the A+V ERP response ($ps < .049$), for central targets. As can be seen in Figure 7, this reflects a superadditive P2 component, with both the visual and auditory P2 to a multisensory central target being larger than the sum of the unisensory counterparts. For the central non-targets, significantly greater parietal-occipital positivity between 153 and 300ms was found in the AV ERP response, compared with the A+V ERP response (critical t-score ± 4.29 , $df = 15$, $ps < .050$). Figure 6d shows that this parietal positivity is more

sustained across the waveform, rather than corresponding specifically to the P2 component. Indeed, an interaction between integration and target/non-target was also found, with a significant difference in parietal-occipital positivity found for central targets compared with central non-targets, between 212 and 224ms (critical t-score ± 4.67 , $df = 15$, $ps < .049$), suggesting that the superadditive P2 component is enhanced when the central stimulus is a target.

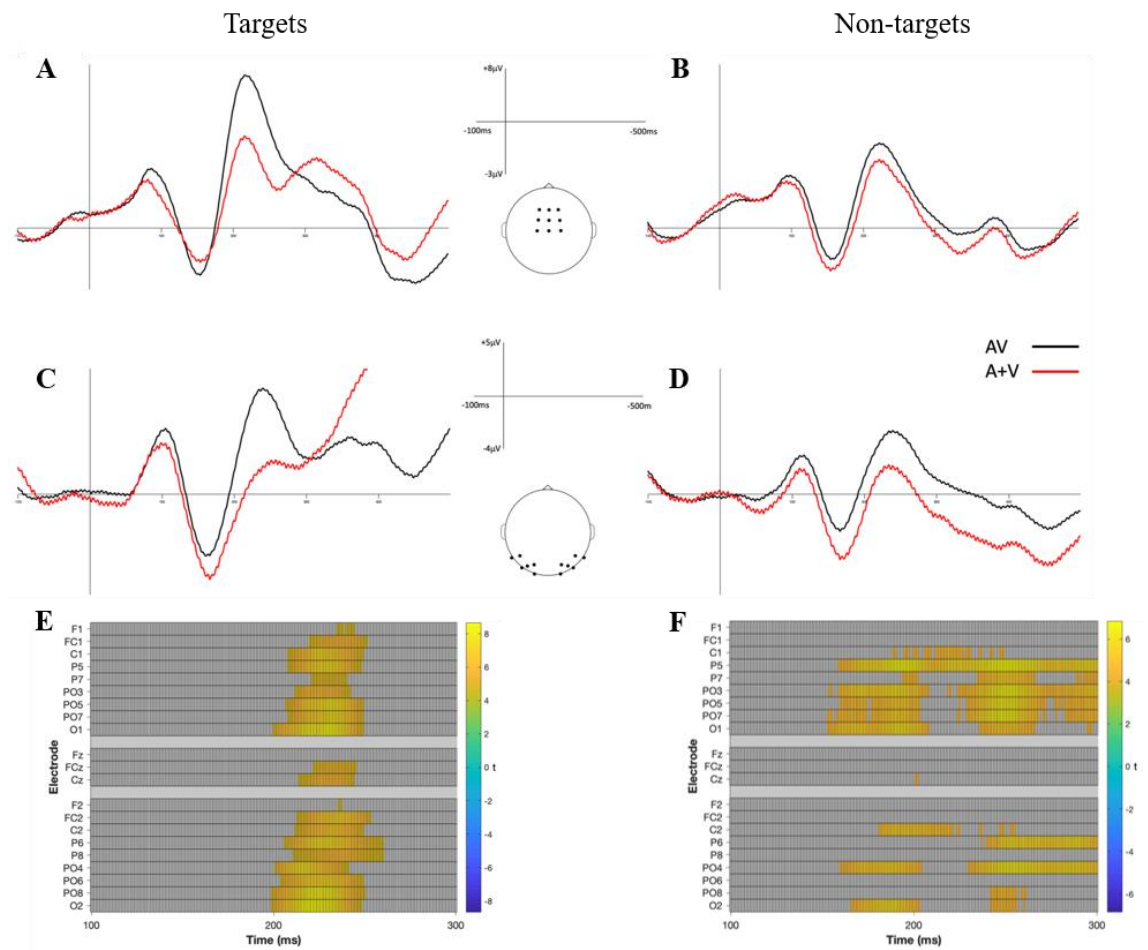


Figure 7. Effect of multisensory integration on central task. Averaged ERP response at the frontal-central electrode cluster for A. targets and B. non-target; averaged ERP response at parietal-occipital electrode cluster for C. targets and D. non-targets. Raster plots to show significant differences between AV and A+V for targets (E) and non-targets (F)

Peripheral Task. The ANOVA on baseline activity from -20 to 0ms confirmed that our ability to test for additive ERP responses to the peripheral task stimuli was not compromised by overlapping activity in the waveforms at any of the clusters used in our analyses, $F < 1$.

Frontal Integration Effect. The early frontal integration effect was analysed using an ANOVA on the frontal electrode cluster, using mean amplitude from 80-120ms, with the factors of integration (AV, A+V), and central load type (multisensory, visual, auditory). No main effects of either integration ($F < 1$, $\eta^2_p = .02$) or load ($p = .376$, $\eta^2_p = .06$) were found, nor an interaction between the two factors ($F < 1$, $\eta^2_p = .030$) (Figure 9).

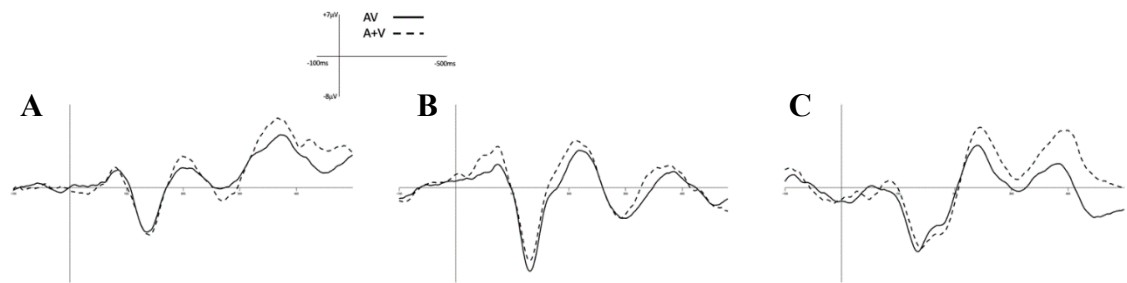


Figure 9. Averaged ERPs at the frontal electrode cluster for AV and A+V peripheral targets, under conditions of A. multisensory, B. visual, and C. auditory load

Mass Univariate Analysis. Under conditions of both multisensory and visual central load, there were no significant differences between the AV and A+V ERPs to peripheral targets presented on the left side of the screen ($ps > .419$ for multisensory load; $ps > .08$ for visual load). Under conditions of auditory central load, significant differences were found, with increased negativity in the AV ERP between 235 and 293ms for parietal-occipital electrodes, and between 215 and 264ms in the frontal cluster (critical t-score ± 4.22 , $df = 15$, $ps < .049$). Figure 10 shows that this corresponds to a sub-additive visual P2, with the response to the multisensory peripheral

stimulus being less than that to the unisensory visual and auditory peripheral stimuli combined.

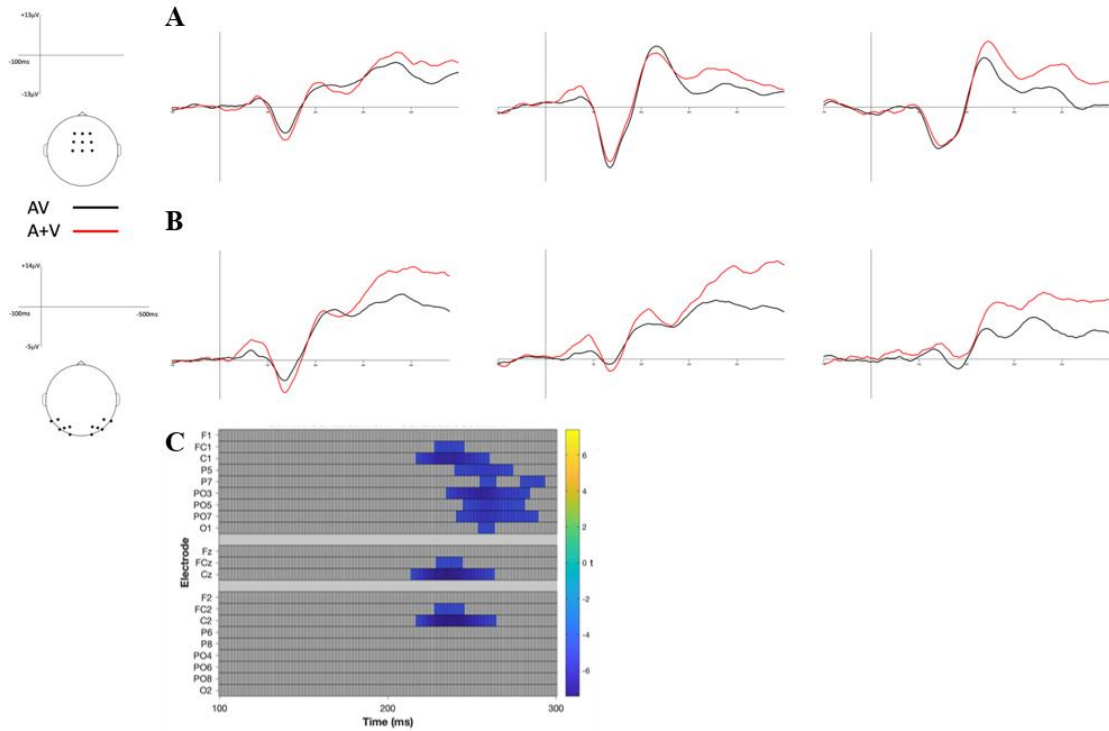


Figure 10. Comparison of multisensory and summed unisensory (auditory + visual) peripheral targets, by central load type. Averaged ERP response at the A. frontal-central electrode cluster and B. parietal-occipital electrode cluster for multisensory, visual and auditory load. C. Raster plot to show significant differences between AV and A+V for auditory central load. No significant differences were found for multisensory or visual central load.

Discussion

Experiment 3 replicated the two key behavioural findings of Experiments 1 and 2, in terms of revealing a similar pattern of performance on the peripheral target

detection task regardless of whether the perceptually loading central task was multisensory or unisensory, and enhanced attentional capture for multisensory versus unisensory peripheral targets. Critically, these behavioural results were demonstrated in the context of compelling ERP evidence that the central perceptual load task involved multisensory integration: We demonstrate superadditivity ($AV > A+V$) in the ERP response to central multisensory stimuli in multiple spatio-temporal clusters; an earlier frontal integration effect, as well as later parieto-occipital and fronto-central clusters.

An unexpected finding was that, in contrast to the super-additive ERP response to central task stimuli, no evidence for super-additivity, in terms of either an early frontal integration effect or the later visual and auditory evoked potentials, was observed in relation to the multisensory peripheral targets. This lack of ERP evidence for neuronal integration of the secondary task comes in contrast to the behavioural results: As mentioned above, response times to multisensory versus unisensory peripheral targets were consistently facilitated across both multisensory and unisensory load tasks, with a race model analysis supporting integration of multisensory peripheral targets under multisensory and unisensory visual load. Indeed, the only case in which any non-linear ERP response to the peripheral task was found was in the auditory load condition, which for this Experiment was the one condition in which the race model assumption was not violated: Here a sub-additive response, rather than super-additive, response was observed at parietal-occipital scalp sites. Although it has been argued that multisensory integration can manifest as sub-additive rather than super-additive ERP responses (Cappe et al., 2010), the lack of an early frontal integration effect combined with the lack of behavioural evidence for integration makes it appear unlikely that this is the case here.

General Discussion

The present study sought to test for the first time how multisensory presentation would impact upon the well-established effects of perceptual load on attention, as well as considering the possibility of a specific impact of multisensory load on attention to other multisensory stimuli. Across all three experiments we used a modified version of a well-established RSVP perceptual load manipulation (e.g. Bahrami et al., 2007; Lunn et al., 2019) to compare conditions of multisensory load to high unisensory visual (Experiments 1-3), as well as high unisensory auditory (Experiment 3), load. Our first key finding was that, in all three experiments, we demonstrate that multisensory versus unisensory presentation of a high perceptual load primary task did not alter secondary task processing, with evidence favouring the null hypothesis in all three experiments, reaching sensitivity in Experiment 2 where a Bayesian stopping rule was applied. Indeed, comparison of our multisensory and unisensory high load conditions to the unisensory low load condition from a closely matched previous study revealed similar modulation of peripheral target responses by both multisensory and unisensory perceptual load (see Figure 2).

Our findings extend the Load Theory literature by clarifying that the proposed ‘special’ properties of multisensory stimuli (e.g. Santangelo & Spence, 2007) do not appear to alter their ability to fill perceptual capacity, thus neither helping nor hindering the speed or accuracy of responses to peripheral targets. In Experiment 3 we demonstrated compelling ERP evidence to support the assumption that the multisensory perceptual load task involved integration, in terms of non-linear ERP responses to multisensory stimuli, with clear periods of superadditivity. Despite this, the multisensory task was still no more effective than a unisensory one in filling perceptual capacity, but also did not impose on capacity any less than high unisensory load, as may

have been expected due to evidence suggesting they are processed more quickly (Colonius & Diederich, 2004; Hughes et al., 1994; Laurienti et al., 2004; Molholm et al., 2002; Murray et al., 2005; Senkowski et al., 2005; Talsma et al., 2007).

Theoretically, this implies that the wealth of data we have on unisensory load will also apply to multisensory load. From an applied perspective, these findings are encouraging in suggesting that the unisensory Load Theory literature remains informative for the large proportion of real-world tasks that involve some multisensory element (e.g. conversation, crossing a road, driving).

The second key finding to emerge from the present study is that multisensory load does not disrupt the facilitatory behavioural advantage afforded by multisensory stimuli in a secondary task. This advantage has previously been found in a series of experiments using two established perceptual load manipulations (Lunn et al., 2019) in terms of faster detection of multisensory versus unisensory stimuli, and in the present study we show across all three experiments that this advantage persists even under conditions of high perceptual load that involve multisensory integration. In Experiments 2 and 3, we show that the responses to multisensory stimuli are not only faster than the unisensory counterparts, but also faster than that which would be predicted by probability summation, even under multisensory load. This has important practical implications for real-world scenarios, as multisensory stimuli have been suggested as good candidates for alerts and warning signals in, for example, driving and aviation. Since these are both multisensory tasks, it is vital that an attentional capture advantage persists under multisensory conditions, and the present research provides evidence that this is indeed the case.

It has previously been suggested that audiovisual integration can occur for only one item at a time (Van der Burg et al., 2013a). Whilst this view has been challenged by research suggesting that capacity may be greater than one based on factors such as perceptual grouping, speed of presentation, and visual field size (Wilbiks et al., 2020; Wilbiks & Dyson, 2016, 2018), even in the most optimum conditions for audiovisual binding, capacity limits did not reach two items on average. Our behavioural evidence of multisensory enhancement of attentional capture by peripheral targets, at the same time as superadditivity in the ERP responses to a central task initially appears to challenge the capacity limit of 1 item. However, this past research employs a paradigm where only one auditory stimulus is presented for integration with multiple visual stimuli, whereas in the present study we presented two pairs of visual and auditory stimuli in the multisensory load condition, with each pair being at a different spatial location to the other. Our research therefore highlights that different paradigms may show simultaneous effects, and thus the number of ‘multisensory items’ that may be integrated at one time could differ from the number of unisensory stimuli that can be integrated with a single stimulus presented in a different modality.

Whilst the behavioural advantage afforded by multisensory stimuli was consistently found across our three experiments, an unexpected caveat is that this appears to come with no superadditive neuronal response, as the aforementioned early frontal integration ERP effect demonstrated in previous literature (e.g. Talsma & Woldorff, 2005) was not found for the peripheral targets. On one hand, this might initially appear to suggest that our high perceptual load primary tasks left insufficient spare attentional capacity to allow for integration (see Alsius et al., 2005, 2014; Talsma et al., 2007). On the other hand, as our findings clearly demonstrate that the behavioural advantage is not always accompanied by ERP superadditivity, it might well be the case

that superadditivity would not be observed even during conditions of low perceptual load (indeed, our prior behavioural study did not find any impact of perceptual load on the multisensory advantage, and found race model violations only under high perceptual load). Future ERP studies comparing the ability of multisensory secondary task stimuli to elicit superadditive responses during high and low load unisensory perceptual load primary tasks are needed to directly test these possibilities. For now, we note that the present work highlights that behavioural evidence for multisensory processing advantages should not be assumed to reflect superadditivity. Indeed, authors of some of the original single-neuron studies highlight that multisensory behavioural advantages are not only due to superadditive responses, and place importance on considering neuronal enhancements due to multisensory integration that are not superadditive, since even linear summation of the two unisensory stimuli would represent an increase in superior colliculus activity compared to a unisensory stimulus alone, and therefore could be expected to enhance behavioural responding (Stanford & Stein, 2007).

In contrast to the secondary peripheral task, our multisensory primary task did elicit superadditive neuronal responses, but no behavioural benefit occurred. This double dissociation is intriguing as both the primary and secondary tasks required attention to both modalities - in the case of the primary task, due to the need to detect a multisensory conjunction, and in the case of the secondary task, because target modality was unpredictable - which has been highlighted as a critical condition for the physiological effects of integration (Talsma et al., 2007). However, it might be argued that the primary task provided a stronger attentional setting for multisensory integration due to the target definition being itself based on a multisensory conjunction (whereas the secondary targets were defined on the basis of either unisensory feature). On the other hand, the behavioural effect in the secondary, peripheral task but not the central

primary task is consistent with past research demonstrating a behavioural benefit where the task requires spatial reorientation to a target (e.g. Molholm et al., 2002; Senkowski et al., 2005; Talsma & Woldorff, 2005), but not where the task is presented such that targets appear at fixation (e.g. Talsma et al., 2007), suggesting that multisensory integration facilitates this reorientation of attention, rather than the selection of a target from a central stream. Nevertheless, this double dissociation between non-linearity of ERPs and a behavioural benefit, highlights the importance of exercising caution when inferring this link in future research.

In conclusion, our results show clear evidence that a task high in multisensory perceptual load – despite being shown to elicit a non-linear neuronal response - does not appear to differ in its ability to fill perceptual capacity from a unisensory load. Secondly, our behavioural results in all three experiments consistently show that multisensory stimuli are detected faster in a secondary task than unisensory stimuli, irrespective of bimodality of the primary task. There was no evidence of non-linear ERP responses to the multisensory secondary task stimuli, nevertheless the clear behavioural advantage they afford in terms of faster detection provides further support for their use in applied, multisensory contexts. Finally, our results highlight that behavioural advantages for multisensory stimuli do not always go hand in hand with non-linear ERP responses and as such should not be assumed to reflect neuronal super-additivity.

Supplementary Materials

Centro-Medial Integration Effect

Central Task

Centro-medial effects of integration were analysed at seven electrodes (FC1, FC2, FCz, C1, C2, Cz, CPz) using ANOVAs on mean voltage at successive 20ms intervals from 0-500ms, with the factors of integration (AV, A+V), central stimulus type (target, non-target), and electrode site (seven levels, corresponding to each channel). As in Talsma & Woldorff (2005), to correct for the increased possibility of Type 1 errors, only results that were statistically significant in two or more consecutive time windows will be reported.

At these electrodes, a significant main effect of the factor integration was found in the latency range 180-280ms (Table S1). As in Talsma and Woldorff (2005), this centro-medially distributed effect of integration corresponded to an enhanced positivity elicited by multisensory stimuli. A significant main effect was also found in between 400-500ms, here corresponding to less positivity elicited by the multisensory stimulus compared with the summed unisensory response. There was also a significant interaction between integration and central stimulus type at this latency range , and follow up t-tests confirm that from 380ms, there were no significant differences between the AV and A+V waveforms for non-targets (ps .070-.949), but for the central targets the AV waveform was less positive than the A+V response (ps <.001-.013). Therefore, this reduced positivity in the AV waveform as compared to the A+V is likely due to the A+V waveform reflecting a summed P300 to both unisensory visual and auditory targets. There was also a significant interaction between the two factors in the

latency ranges 200-240ms, and 280-500ms. Follow up t-tests show that between 200-240ms the AV waveform was significantly more positive than the A+V waveform for both targets and non-targets ($p < .001$ for targets, $p < .001-.002$ for non-targets), but then between 280-360ms this was only the case for the non-targets ($p .072-.431$ for targets, $p .003-.037$ for non-targets). As can be seen in Figure 7, the AV non-target was generally more positive than the A+V non-target across the entire waveform, whereas for the central targets this multisensory enhancement occurred at specific time points. The factor of central stimulus type was significant between 180-380ms, with central targets eliciting more positive waveforms than central non-targets.

Time (ms)	Integration		Top-down set		Int*Set	
	<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>
0-20	< 1	.539	2.44	.139	< 1	.497
20-40	< 1	.798	1.99	.179	1.05	.397
40-60	< 1	.785	< 1	.995	< 1	.662
60-80	< 1	.996	3.45	.083	< 1	.669
80-100	1.77	.204	< 1	.602	< 1	.525
100-120	3.19	.094	1.30	.272	< 1	.712
120-140	1.27	.278	1.24	.283	2.46	.138
140-160	< 1	.953	< 1	.680	4.42	.053
160-180	1.00	.334	1.38	.258	1.34	.265
180-200	8.56	.010*	8.81	.010*	< 1	.600
200-220	24.15	< .001*	10.19	.006*	8.18	.012*
220-240	51.04	< .001*	6.03	.027*	14.64	.002*
240-260	39.48	< .001*	4.36	.054	4.56	.050
260-280	6.66	.021*	6.31	.024*	< 1	.857
280-300	< 1	.602	10.60	.005*	4.59	.049*
300-320	< 1	.729	12.26	.003*	10.91	.005*
320-340	< 1	.769	13.25	.002*	14.07	.002*
340-360	< 1	.632	10.27	.006*	11.94	.004*
360-380	< 1	.386	5.46	.034*	9.50	.008*
380-400	2.10	.168	1.51	.238	8.95	.009*
400-420	5.42	.034*	< 1	.496	11.53	.004*
420-440	9.96	.007*	< 1	.525	13.23	.002*
440-460	13.95	.002*	< 1	.628	16.12	.001*
460-480	29.34	< .001*	< 1	.546	24.21	< .001*
480-500	40.18	< .001*	< 1	.492	9.17	.008*

Table S1 Overview of ANOVA results over centro-medial scalp sites for the effects of

multisensory integration, and central stimulus type, from 0-500ms. All tests were

conducted with (1, 15) degrees of freedom, * indicates significance at $p < .05$

Peripheral Task

Effects of integration and top-down attentional setting over centro-medial scalp sites were analysed using ANOVAs on mean voltage at successive 20ms intervals from 0-500ms, with the factors of integration (AV, A+V), load (multisensory, unisensory visual, unisensory auditory), and electrode site (6 levels, corresponding to each channel).

At these electrodes, a significant main effect of the factor integration was found only in the latency range 300-500ms, with the response to A+V targets being more positive than to AV targets (Table S2). Similarly to the central targets, this is due to the A+V waveform summing together two P300 components. There were no other centro-medially distributed effects of integration. A main effect of load was demonstrated at 100-280ms and at 420-500ms, with a significant interaction between integration and load at 240-280ms. Follow up t-tests show that there were no significant differences between AV and A+V targets under conditions of either multisensory or visual load between 240-280ms ($ps > .391$), but under conditions of auditory load, the response to a multisensory target was significantly less positive than to the unisensory sum.

Time (ms)	Integration		Load		Int*Load	
	<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>
0-20	< 1	.963	6.82	.004*	< 1	.524
20-40	< 1	.814	2.99	.066	< 1	.494
40-60	1.40	.491	2.96	.067	1.84	.177
60-80	< 1	.256	3.90	.031*	1.03	.369
80-100	< 1	.609	1.15	.330	< 1	.908
100-120	< 1	.994	3.95	.030*	< 1	.433
120-140	< 1	.658	14.87	< .001*	1.52	.235
140-160	< 1	.903	9.63	< .001*	< 1	.381
160-180	< 1	.911	17.08	< .001*	< 1	.695
180-200	< 1	.758	11.97	< .001*	< 1	.632
200-220	< 1	.992	10.36	< .001*	< 1	.810
220-240	< 1	.936	21.67	< .001*	1.99	.155
240-260	< 1	.450	14.25	< .001*	5.54	.009*
260-280	2.49	.135	3.74	.035*	6.54	.004*
280-300	4.01	.064	< 1	.893	2.91	.070
300-320	6.83	.020*	1.11	.343	1.49	.243
320-340	8.13	.012*	2.61	.090	1.60	.218
340-360	7.24	.017*	3.31	.050	1.02	.373
360-380	5.91	.028*	4.36	.022*	< 1	.547
380-400	6.87	.019*	3.04	.063	< 1	.485
400-420	8.34	.011*	2.23	.125	1.75	.191
420-440	8.72	.010*	3.95	.030*	1.80	.183
440-460	8.03	.013*	4.54	.019*	< 1	.433
460-480	6.99	.018*	6.90	.003*	< 1	.518
480-500	5.10	.039*	8.17	.001*	1.58	.223

Table S2. Overview of ANOVA results over centro-medial scalp sites for the effects of multisensory integration and central load condition, from 0-500ms. The main effect of

integration was conducted with (1, 15) degrees of freedom, and the effect of load and the interaction with (2, 60) degrees of freedom. * indicates significance at $p < .05$

Chapter 6: Blinded or deafened by our thoughts? Testing for dissociable effects of auditory and visual mental imagery on early perceptual processing

Chapter Summary

Throughout this thesis, I have thus far considered the effects of perceptual capacity and the allocation of attentional resources to sensory stimuli presented externally. Given that previous research has provided evidence for an attentional capacity limit that is shared crossmodally, as well as between representations of externally presented and internally generated stimuli, this chapter examines the attenuation of visual and auditory perceptual processing of external stimuli, by visual and auditory mental imagery. It will seek to explore whether such attenuation can occur crossmodally, as well as if there are modality specific effects, and how early in the processing of external stimuli it may occur. Where additional analyses will be included in the article submission as supplementary materials, these have been provided at the end of the chapter.

Abstract

In the light of increasing evidence of shared neural mechanisms underlying both external perception and mental imagery, an important question is whether engagement in mental imagery (e.g. during mind wandering) undermines the brain's ability to construct a representation of our external environment. Here we test for both amodal and modality-specific attenuation of visual and auditory perceptual processing by naturalistic visual and auditory mental imagery. Engagement in mental imagery was associated with a reduced response in the sensory cortices to both visual and auditory external events. Attenuation of visual sensory processing by both auditory and visual mental imagery was found from primary visual cortex onwards, with the degree of attenuation increasing in successive stages of processing and being consistently stronger for visual versus auditory imagery. In contrast, attenuation of the auditory cortical response was confined to A3 and A4, and occurred irrespective of the modality of the imagery. Our findings suggest that the effects of mental imagery on both visual and auditory processing may involve shared attentional resources, but modality-specific competition for neural resources occurs only for stimuli occurring in the visual domain.

Introduction

Neuroscientific advances increasingly highlight that the distinction between imagination and real-world perceptual experience is less clear-cut than might be assumed. There appears substantial overlap in the neural mechanisms involved in creating a conscious percept of an external sensory stimulus versus generating mental imagery: Visual mental imagery is well established to activate the visual cortex, even as early as VI (Chen et al., 1998; Ganis et al., 2004; Ishai et al., 2002; Klein et al., 2000; Lambert et al., 2002; Le Bihan et al., 1993; Naselaris et al., 2015), and recent multivariate analysis reveals shared representations of visual imagery and visual perception originating from parietal-occipital regions (Xie et al., 2020). These effects are strongest when the imagery itself is more vivid (Cui et al., 2007), or where a high-resolution detail is inspected (Kosslyn & Thompson, 2003). Neural responses in the auditory cortex have also been demonstrated during auditory mental imagery (e.g. Halpern et al., 2004), though the extent to which the primary auditory cortex is recruited remains unknown (see Zatorre & Halpern, 2005). Given evidence that we spend approximately half of our waking lives thinking about matters unrelated to our immediate external environment (Killingsworth & Gilbert, 2010), an interesting question is how mental imagery relating to task-unrelated mind wandering might impact our ability to perceptually process the world around us.

The possibility that mental imagery might impede external perception is supported by several lines of evidence. First, within the mind wandering literature it is fairly well documented that participants show reduced sensory and cognitive processing of external events during periods in which they report off-task mind wandering, often described as ‘decoupling’ of inner and outer experience (see Kam & Handy, 2013, for a review). Reduced event related potential (ERP) responses to external events during

mind wandering have been found during paradigms presented in the visual (e.g. Barron et al., 2011; Kam et al., 2013; Smallwood et al., 2008), and auditory domains (e.g. Braboszcz & Delorme, 2011; though see Kam et al., 2013). Interestingly, other evidence reveals the opposite pattern in terms of suppression of mind wandering by increased external perceptual processing, consistent with two-way competition between internal and external representations. Forster and Lavie (2009) demonstrated that increasing the visual perceptual load of an external task modulates the occurrence of internally generated, task-unrelated thoughts, and this finding has since been replicated in an applied context, with a decreased occurrence of mind wandering whilst driving in a more perceptually demanding scenario (Geden et al., 2018). Hence, although the research reviewed above does not directly address mental imagery (as this was not typically measured), it is consistent with the notion of competition for representation of information that is internally generated versus derived from external sources.

In considering potential mechanisms for mental imagery effects on external perception, a key question is whether or not these are modality-specific – for example, is visual perception more affected by imagery generated in the ‘mind’s eye’ versus the ‘mind’s ear’. On one hand, prominent theoretical models of attention might lead us to expect amodal effects, occurring irrespective of modality. Work in support of the Load Theory of attention implies that information processing is limited by the availability of a ‘perceptual capacity’, recently proposed to reflect limited cellular metabolic energy across the brain (Bruckmaier et al., 2020; Lavie, 2010). This limited capacity resource appears to be shared across the senses, insofar as filling capacity with visual information appears to similarly impact awareness of visual, auditory, tactile and even olfactory stimuli (Forster & Spence, 2018; Macdonald & Lavie, 2011; S. Murphy & Dalton, 2016; Raveh & Lavie, 2015). Critically, although this theory has been primarily

tested in relation to external stimuli, as mentioned above more recent initial work has implied that this central capacity may also be shared between internal and external representations (Forster & Lavie, 2009; Konstantinou et al., 2012; Morris et al., 2020). Hence, if mental imagery effects on external perceptual processing are primarily driven by this central attentional resource, we may anticipate a general imagery induced attenuation of external perceptual processing, with the direction of attention to internal sources disrupting processing of external stimuli regardless of sensory modality.

On the other hand, given the reliance on shared resources within the sensory cortices, we might expect that this competition for resources would lead to modality-specific patterns of disruption. Initial evidence in support of modality-specific effects comes from an ERP study demonstrating that whilst both freely generated visual imagery and inner speech reduce the P1 component generated by an external visual stimulus, this reduction was greater for the visual imagery condition (Villena-González et al., 2016). Conversely, a subsequent study found no effect of either of the internal thought conditions on the ERP response to an external auditory stimulus (Villena-González et al., 2018). The recent studies of Villena-González and colleagues might initially appear to imply that the effects of mental imagery are not only modality-specific, but also limited to some sensory domains (i.e. visual). However, as these studies focused on the comparison of freeform visual imagery and inner speech, it is unclear how well the two conditions were matched in terms of the amount of and clarity of sensory imagery, and the difficulty in producing these types of thought. Such variation could impact both the perceptual (or visual short-term memory) and cognitive load associated with the imagery, factors that are well established to powerfully modulate the allocation of attentional resources (Lavie et al., 2004). As such, the results of these prior studies could potentially reflect various differing qualities of the imagery

produced by participants in response to the instructions to generate visual imagery versus inner speech, that go beyond sensory modality.

The present study was designed to establish the ability of visual versus auditory mental imagery to attenuate external visual and auditory perception, as well as whether this may extend even to the earliest stages of perceptual processing. In order to better control the mental imagery generated by participants, we developed a set of naturalistic scenarios, piloted in a large sample of participants to confirm that these were not only successful in eliciting larger amounts of the intended imagery type, but were also matched for the difficulty in producing this imagery. Whilst imagining these scenarios, participants were presented with stimuli that are established to elicit robust neuronal responses across the visual and auditory cortices. In this manner, the present study goes beyond previous ERP research to map out the impact of imagery across the visual and auditory cortices, using the superior spatial resolution afforded by fMRI. Evidence suggests that visual short-term memory load modulates irrelevant external visual processing as early as V1, with the effects increasing across successive visual areas (Konstantinou et al., 2012). Given the proposed overlap between visual working memory and visual mental imagery (Keogh & Pearson, 2011; Pearson et al., 2015), this raises the possibility that even naturalistic imagery, of the type that could occur during mind wandering, may induce attenuation of visual external stimuli even as early as V1. Using a similar approach will also afford us increased sensitivity to look for parallel effects in successive regions of the auditory cortex.

Methods

Participants

19 healthy, right-handed participants (13 female) aged between 20 and 30 years ($M = 25.42$, $SD = 3.66$) were recruited to participate in this experiment and were reimbursed for their time. One participant was excluded from analysis due to excessive movement throughout all functional runs. The study was approved by the Research Governance Ethics Committee (RGEC) at Brighton and Sussex Medical School. All participants reported normal or corrected-to-normal (with contact lenses) vision and hearing, fluency in English, capable of producing both visual and auditory mental imagery, and did not have any MRI contraindications.

Stimuli and Procedure

The experiment was programmed and presented using MATLAB R2018a (Mathworks, Inc.) with Psychtoolbox (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997), backprojected onto a screen and viewed through an angled mirror on the head coil. Auditory stimuli were delivered over MRI safe in-ear headphones, and a LiveTrack AV Eye Tracker was manually adjusted for optimum viewing for each participant. Each run began with a screen which read 'Waiting for scanner' and a central countdown from 5. At the start of every trial, a central fixation cross was presented alone on screen for 500ms, and an imagery instruction was then presented centrally for 5000ms. Scenario presentation time was determined based off an online pilot study ($N = 38$, see Supplementary Materials), such that this would be an adequate time to read the majority of scenarios whilst maintaining a maximal number of trials. These instructions either asked participants to imagine specific scenarios designed to induce either visual or

auditory mental imagery (e.g. “imagine your journey from home to work/university or “picture which of your friends’ voices you like the most and the least”), or to direct attention to the upcoming stimuli (“please pay attention to the visual and auditory stimuli”). The 48 visual and auditory scenarios were selected from a larger set on the basis of an online pilot ratings study ($N = 62$, see Supplementary materials), so that they were matched, on average, for level of difficulty in creating the mental imagery, and such that visual scenarios induced more visual imagery than auditory, and vice versa for the auditory scenario.

A second fixation cross was then presented for 500ms to draw the participants eyes back to the centre of the screen. Participants had a total of 11 seconds to imagine the scenario (participants had previously been instructed to picture it as vividly and in as much detail as possible). During this, visual and auditory stimuli were presented. In each 11 second period, 5 stimuli were presented for 1000ms each with a 1000ms interstimulus interval (ITI). There were 4 possible stimuli (2 visual and 2 auditory) which were presented in a pseudorandomised order, so that they each appeared 45 times for each mental imagery condition during the overall study. The two visual stimuli were Gabor patches, consisting of a 0.05 cycles per pixel sinusoidal grating modulated by a Gaussian envelope ($SD = 44$), one oriented at 45° and the other at 135° . Auditory stimuli were based off the ‘Morse code’ like pattern of sounds used by (Thomas et al., 2015) for tonotopic mapping of the auditory cortex, designed to increase the perceptual salience of auditory stimuli over the background scanner noise. Each auditory stimulus consisted of a burst of 8 pure tones of either 50ms or 200ms in length, alternated in a pseudo-randomised order with an ITI of 50ms to maintain consistent overall length. The exact pitch of each pure tone was also randomly jittered in proportion to a base frequency, with steps of 1% away from the base frequency up to $\pm 5\%$. The low pitch

auditory stimulus therefore consisted of a burst of tones jittered around 141Hz (133.95-148.05Hz) and the high pitch auditory stimulus jittered around 3200Hz (3040-3360Hz). The sound intensity was adjusted according to a standard equal-loudness curve to approximate equal perceived loudness across the two frequencies. Volume was individually adjusted to be clearly and comfortably audible for each subject.

Participants completed five functional runs with the scanner; one run consisting of a recap of the instructions followed by 20 scenarios, and four runs of 22 scenarios. 30 scenarios were presented across the experiment for each of the mental imagery conditions, and 48 for the direct attention condition.

Outside the scanner, following completion of the mental imagery task, participants rated every scenario on a 5-point Likert Scale for difficulty (where 1 = extremely easy, 5 = extremely difficult), amount of visual and auditory mental imagery experienced (where 1 = no visual/auditory mental imagery, 5 = visual/auditory mental imagery all the time), and clarity of visual and auditory mental imagery (where 1 = extremely unclear – would not be able to tell what the visual images/sounds were, 5 = extremely clear – as if it were real life). For the purpose of confirming the ability to generate some mental imagery (i.e. the absence of aphantasia), participants also completed the Vividness of Visual Imagery Questionnaire (VVIQ; Marks, 1973), and the Plymouth Sensory Imagery Questionnaire (Psi-Q; Andrade et al., 2014). Note that we used a version of the VVIQ where higher scores are indicative of more mental imagery, which is more intuitive, particularly when completed in combination with the Psi-Q and all other post-scan scales where 5 indicated higher ratings of the construct being measured (for similar practice, see Keogh & Pearson, 2011; Zeman et al., 2015).

fMRI Acquisition and Analysis

A 3T Siemens Prisma scanner with a 64-channel head-coil was used to acquire all images. Over-ear protectors were placed over the in-ear headphones to block out scanner noise. Where this was not possible due to head size, soft cushions were inserted into the head coil, which additionally provided comfort and reduced head movement. Functional images were acquired using the Human Connectome Project (HCP) gradient-echo EPI sequence, with a multiband acceleration factor of 8; TR = 0.8s; TE = 33.1ms; 52 degree flip angle; FOV = 208 x 180mm; 72 slices with slice thickness of 2mm and isotropic 2mm voxels. Two SpinEcho Field maps with reversed phase-encode blips in both Anterior to Posterior and Posterior to Anterior were acquired with the same parameters as the functional images. A high-resolution structural T1-weighted image was acquired with 3D MPRAGE sequence (TR = 2.4s; TE = 2.14s; 8 degree flip angle; FOV = 224 x 224mm and 0.8mm isotropic voxels).

Images were pre-processed using a combination of SPM12 (Wellcome Department of Imaging Neuroscience, London, UK), FSL (Smith et al., 2004), and FreeSurfer (Fischl, 2012). Images from all runs were spatially realigned to the mean image. Field maps were estimated and applied using FSL, to correct for image distortions (Andersson et al., 2001, 2003). Cortical reconstruction and volumetric segmentation of images was performed using FreeSurfer (freely available for download online at <http://surfer.nmr.mgh.harvard.edu>). Procedures are described in detail in previous publications (Dale et al., 1999; Dale & Sereno, 1993; Fischl et al., 2004; Fischl & Dale, 2000; Ségonne et al., 2004). These include motion correction, removal of non-brain tissue, segmentation of the subcortical white matter and deep grey matter volumetric structures, surface deformation, surface inflation, and registration to a

surface-based atlas, in this case the Human Connectome Project (HCP) atlas. Functional images were aligned to the anatomical image, and a 4mm FWHM smoothing kernel was applied to the functional images for the GLM analyses. Each stimulus type (visual 45°, visual 135°, auditory high tones, auditory low tones) was modelled in a separate one second regressor for each imagery condition (visual, auditory, no imagery), giving 12 regressors of interest. GLM-denoise was applied (available MATLAB code at <http://kendrickkay.net/GLMdenoise/>) which derives the optimal number of principle components to use as noise regressors (Kay et al., 2013).

Data analysis was conducted using custom scripts in MATLAB (Version 2019b; The MathWorks, Inc., MA, USA). Beta-images were created for each of the three imagery conditions (visual, auditory, no imagery) and the two external stimuli modalities (visual, auditory). Whole-brain contrast images were evaluated with one sample t-tests, thresholded at $p < .01$. Regions of interest (ROI; Figure 1) were chosen on the basis of examining the effect of mental imagery on the processing of visual and auditory stimuli, decided a priori and defined using the HCP atlas (Glasser et al., 2016). In the visual cortex, these regions were V1, V2, V3 and V4. In the auditory cortex, they were A1 (the auditory core, located on medial and posterior Heschl's gyrus, Glasser et al., 2016), A2 (comprised of Lateral Belt and Medial Belt, henceforth L-Belt and M-Belt, on the superior surface of the superior temporal gyrus, located laterally and antero-medially to A1 respectively, Beauchamp et al., 2019; Glasser et al., 2016), A3 (Para Belt, henceforth P-Belt, corresponding with von Economo & Koskinas (1925)'s atlas area TA1, Glasser et al., 2016; Triarhou, 2007) and A4 (likely overlaps with Te3 in the lateral superior temporal gyrus, Glasser et al., 2016; Morosan et al., 2005) in the auditory cortex.

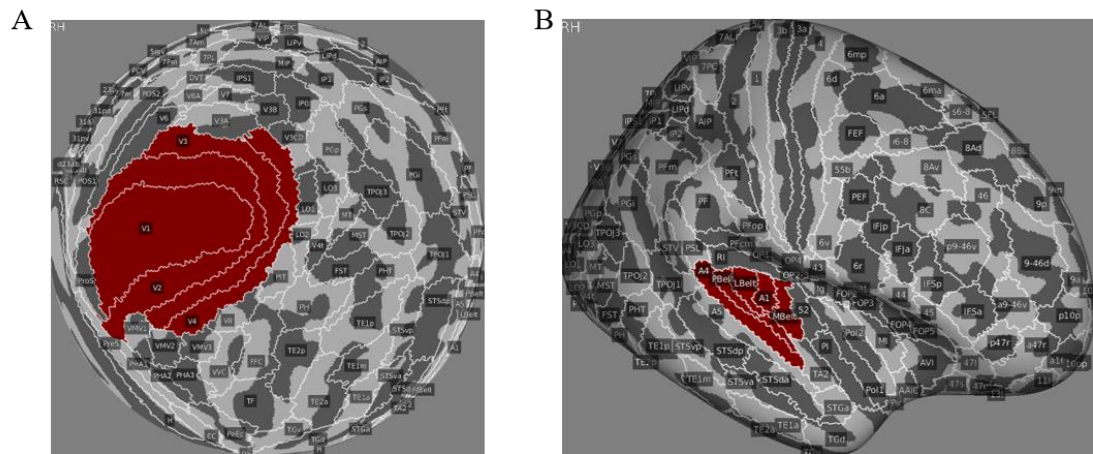


Figure 1. HCP atlas (Glasser et al., 2016) visualised on the inflated cortical surface of the right hemisphere, in FreeSurfer's fsaverage space. A. occipital view, B., lateral view, with ROIs shown filled in red. Scripts for visualisation available at <http://github.com/kendrickkay/>

For each separate ROI, the mean signal over all voxels in both hemispheres was extracted for each condition. Using repeated measures ANOVAs with the factors of imagery condition (visual, auditory, direct attention) and cortical region (V1-V4 or A1-A4 for visual and auditory stimuli, respectively) and follow up paired sample t-tests, we examined whether there was a significant reduction in activity in each ROI in the presence of visual or auditory mental imagery compared to direct attention to the visual and auditory external stimuli. The reduction (or, cost) in signal due to both imagery modalities was calculated at each ROI by subtracting the signal during visual or auditory mental imagery from the direct attention condition, and then the cost in activity elicited by a visual external stimulus during mental imagery was entered into a 2 x 4 repeated measures ANOVA, with the factors of imagery modality (visual, auditory) and visual cortex region (V1, V2, V3, V4). The cost in activity elicited by an

auditory external stimulus during mental imagery was entered into a similar ANOVA, but with the second factor of auditory cortex region (A1, A2, A3, A4).

Results

Behavioural Results

VVIQ scores ($M = 57.44$, $SD = 8.43$) for our participants were similar to those previously found for non-aphantasic controls ($M = 57.92$; Zeman et al., 2015) and in a meta-analysis of 1869 participants ($M = 59.20$; McKelvie, 1995). They were also substantially greater than scores for aphantasics defined as having ‘no imagery’ (obtaining the minimum score of 16), or ‘minimal imagery’ (scores of 17-30; Zeman et al., 2015). Scores on the Psi-Q ($M = 7.28$, $SD = 1.38$), were similar to that found in a sample of 404 participants in the original study establishing this questionnaire ($M = 7.05$; Andrade et al., 2014). These results confirm that all participants were able to produce sensory mental imagery.

A paired samples t-test revealed that, consistent with the online pilot study, participants in the main study rated no significant difference between the average difficulty in imagining a visual scenario ($M = 2.41$, $SD = .43$) and the difficulty imagining an auditory scenario ($M = 2.45$, $SD = .51$), $t < 1$ (Table 1).

Analysis of imagery ratings also confirmed the success of the visual versus auditory scenarios in elicited the desired modality of imagery: The mean amount of imagery experienced by participants was entered into a 2 x 2 repeated measures ANOVA, with the factors of scenario type (visual, auditory) and imagery experienced (visual, auditory). A main effect of scenario type was found, $F(1, 17) = 22.28$, $p < .001$, $\eta^2_p = .57$, with overall more imagery experienced for the auditory scenarios. There was also a significant main effect of imagery experienced, $F(1, 17) = 17.96$, $p < .001$, $\eta^2_p = .51$, with overall more visual imagery experienced than auditory. Both main effects are due to auditory scenarios eliciting more visual imagery than visual scenarios did

auditory imagery ($t(17) = 4.77, p < .001$). Critically, there was a significant interaction between scenario type and imagery experienced, $F(1, 17) = 72.35, p < .001, \eta^2_p = .81$. Follow up t-tests confirmed that when the scenario was visual, participants experienced significantly more visual imagery than auditory imagery $t(17) = 9.51, p < .001$, and when the scenario was auditory, participants experienced significantly more auditory imagery than visual, $t(17) = 5.09, p < .001$. Additionally, participants experienced significantly more visual imagery when the scenario was visual compared to when it was auditory, $t(17) = 5.93, p < .001$, and significantly more auditory imagery when the scenario was auditory compared to when it was visual, $t(17) = 9.17, p < .001$.

From the 15 participants who experienced, on average, at least some of both imagery modalities in the two different scenario types, the clarity of this imagery was entered into a 2 x 2 repeated measures ANOVA, with the factors of scenario type (visual, auditory) and imagery experienced (visual, auditory). No main effect of scenario type was found, $F(1, 14) = 1.63, p = .223, \eta^2_p = .10$, with equal clarity of the overall imagery experienced between the two scenarios. There was also no significant main effect of clarity of the modality of the imagery, $F < 1, \eta^2_p = .02$, with overall equal clarity of visual and auditory imagery across the experiment. Critically, there was again a significant interaction between scenario type and imagery experienced, $F(1, 14) = 14.45, p = .002, \eta^2_p = .51$. Follow up t-tests confirmed that when the scenario was visual, participants experienced significantly clearer visual imagery than auditory imagery $t(14) = 2.87, p = .006$, and when the scenario was auditory, participants experienced significantly clearer auditory imagery than visual, $t(16) = 4.77, p < .001$. Additionally, participants experienced significantly clearer visual imagery when the scenario was visual compared to when it was auditory, $t(16) = 4.46, p < .001$, and

significantly more auditory imagery when the scenario was auditory compared to when it was visual, $t(14) = 3.31, p = .003$.

Scenario	Difficulty	Imagery Amount		Imagery Clarity	
		Visual	Auditory	Visual	Auditory
Visual	2.41	3.75	1.64	3.61	2.99
	(.43)	(.53)	(.66)	(.54)	(.81)
Auditory	2.45	2.57	3.71	3.04	3.64
	(.51)	(.86)	(.56)	(.65)	(.52)

Table 1. Summary of mean ratings of visual and auditory scenarios on the difficulty to produce mental imagery of the scenario provided, amount of imagery produced in each modality, and the clarity of imagery produced in each modality (SDs in parentheses). All ratings made on a Likert scale of 1-5, where 1 indicates none of the construct and 5 indicates maximum levels of the construct.

To summarise, the pattern of responses shows that our visual and auditory scenarios were well matched for difficulty, and successfully resulted in the primary imagery experienced being that which the scenario was designed to induce. In addition, the imagery experienced was clearer when it was in the modality that the scenario was designed for.

Imaging Results

Mental imagery effects on visual cortex response to external stimuli

The external visual stimuli elicited activation, as expected, in each of the visual cortex ROIs. As can be seen in Figure 2, this was significantly reduced under conditions of mental imagery, with both imagery conditions significantly impacting the visual processing from the earliest stage in V1, through to V4 (p s < .001 - .023). However, the impact of mental imagery on activity in the visual cortex varied between regions, $F(2.08, 35.41) = 25.56, p < .001, \eta^2_p = .60$, reflecting an increase in cost from V2 to V3, and V3 to V4. Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(2) = .28, p = .001$), therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\epsilon = .69$). Additionally, there was overall a greater cost to visual cortex activation as a result of engaging in visual mental imagery, compared with auditory mental imagery, $F(1, 17) = 14.87, p < .001, \eta^2_p = .47$. There was no significant interaction between visual region and imagery modality, $F(1.61, 27.35) = 1.61, p = .072, \eta^2_p = .15$. Again, Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(2) = .16, p < .001$), therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\epsilon = .54$).

Follow up tests revealed that the magnitude of the cost for each imagery modality was equivalent in V1 and V2 (t s < 1), but successively increased from V2 to V3 ($t(17) = 4.47, p < .001$ and $t(17) = 2.88, p = .005$ for visual and auditory imagery, respectively) and from V3 to V4 ($t(17) = 5.86, p < .001$ and $t(17) = 4.21, p < .001$ for visual and auditory imagery, respectively). Additionally, there was significantly greater cost to the signal in response to external visual stimuli when participants were engaged

in visual mental imagery as compared with auditory mental imagery, at every visual ROI examined ($ps < .001 - .002$).

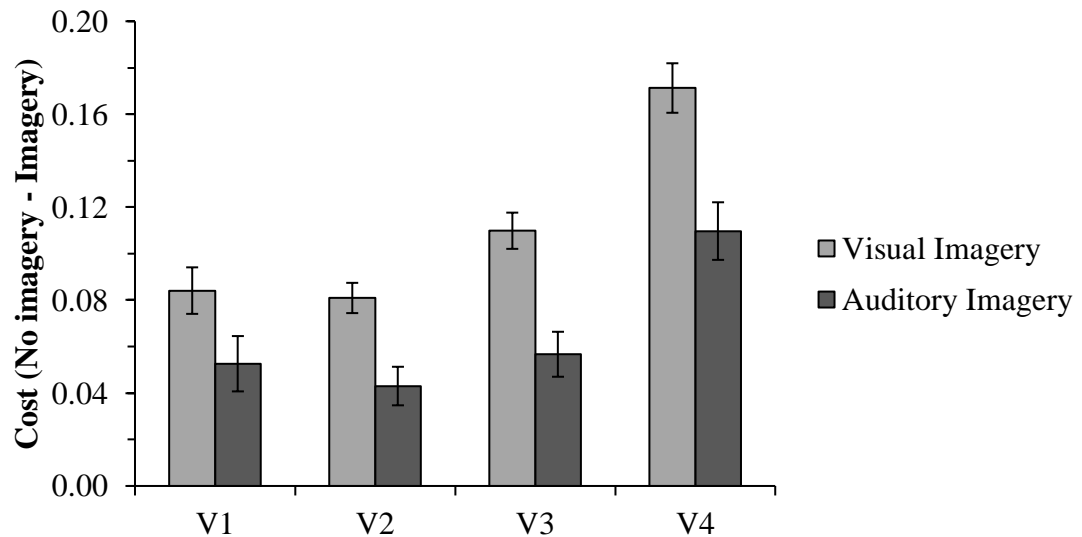


Figure 2. Mean cost to activity (no imagery minus imagery), in response to external visual stimuli, for both visual and auditory mental imagery, for each visual ROI. Error bars show ± 1 SEM with Cousineau-Morey correction (Cousineau, 2005; Morey, 2008)

Mental imagery effects on auditory cortex response to external stimuli.

In contrast to the pattern seen in visual cortex, mental imagery suppression of response to external auditory stimuli was confined to later regions of auditory cortex (see Figure 3). Across the two earlier auditory cortex ROIs (A1 and A2), the average signal to auditory stimuli was not significantly reduced when participants were engaged in either visual or auditory imagery, as compared with the no imagery condition ($ps = .091 - .541$), however there was a significant reduction in both A3 and A4, when

engaged in imagery in either modality ($ps < .001 - .002$). Overall, the impact of mental imagery did not differ dependent on the modality of the imagery, $F < 1$, $\eta^2_p = .01$, but it did differ by auditory cortex region, $F(1.94, 32.93) = 14.81$, $p < .001$, $\eta^2_p = .46$, reflecting an increase in cost from A2 to A3. Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(2) = .35$, $p = .006$), therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\epsilon = .65$). There was no significant interaction between imagery modality and visual region, $F(3, 51) = 1.46$, $p = .237$, $\eta^2_p = .08$.

Follow up tests revealed that the magnitude of the cost due to auditory imagery did not increase between A3 and A4 ($t < 1$), but did increase for visual imagery ($t(17) = 2.52$, $p = .011$). The cost to the signal in response to external auditory stimuli when participants were engaged in auditory mental imagery was no greater than the cost due to engaging in visual mental imagery at either V3 or V4 ($ts < 1$).

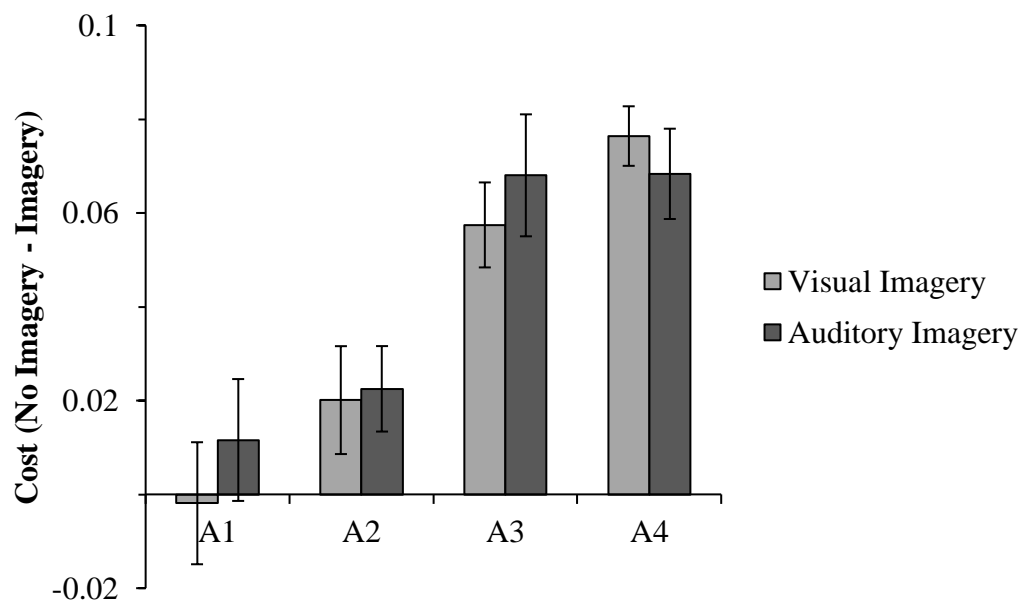


Figure 3. Mean cost to activity (no imagery minus imagery), in response to external auditory stimuli, for both visual and auditory mental imagery, for each auditory ROI. Error bars show ± 1 SEM with Cousineau-Morey correction (Cousineau, 2005; Morey, 2008)

Discussion

The present research demonstrates several key findings. First, engaging in either visual or auditory naturalistic mental imagery induced attenuation of visual perceptual processing across the visual cortex, even as early as V1. In parallel to prior findings using task manipulations of visual short-term memory load or external perceptual load, we demonstrate a successive increase in disruption due to imagery occurring in both modalities across the cortical areas up to V4 (Konstantinou et al., 2012; Schwartz et al., 2005). As such our findings confirm that a powerful and early suppression of the visual cortical response can arise in response to naturalistic mental imagery of the type that might occur during daily life mind wandering.

Our second key finding is that we demonstrate for the first time that mental imagery can also disrupt the auditory cortex response to external auditory stimuli. This finding comes in contrast to the ERP research conducted by Villena-González et al., (2016, 2018), who found while the P1 response to external visual stimuli was reduced when participants were instructed to engage in both freeform visual imagery or inner speech, neither visual imagery nor inner speech impacted ERP measures of auditory sensory processing. The component examined by the authors, the auditory N1, is thought to be comprised of activity from multiple neural generators in the temporal lobe (Näätänen & Picton, 1987) and thus our finding of a cost to auditory processing during internal imagery conditions may have been revealed due to the superior spatial resolution afforded by fMRI to specifically examine individual areas of the auditory cortex.

Nevertheless, our findings are to some extent consistent with the broader picture identified by these authors that the effect of imagery is different for auditory versus

visual perceptual processing in two key respects. First, we demonstrate that, while visual cortical suppression was seen from V1 onwards, the attenuation of auditory cortex only occurs in later cortical regions, with no reduction in activity in the primary auditory cortex. Second, our findings suggest that visual perceptual processing, but not auditory processing, may be subject to an additional modality-specific form of suppression above and beyond the general amodal effects from both forms of imagery: While the auditory cortex response was similarly attenuated by both visual and auditory imagery, visual cortex suppression was consistently greater, across all four ROIs, during visual versus auditory imagery.

Given that the form of imagery induced attenuation common to both visual and auditory perception appeared to occur amodally (i.e. during both auditory and visual imagery), we propose that such effects of imagery are driven by central attentional resources shared between internal and external processing. As noted above, work on the Load Theory has highlighted a ‘perceptual capacity’ limit that appears to be shared crossmodally (Forster & Spence, 2018; Macdonald & Lavie, 2011; S. Murphy & Dalton, 2016; Raveh & Lavie, 2015) , and there is some initial evidence to suggest this capacity limit might also be shared by not only representations of external perceptual stimuli but also internally generated stimuli such as intrusive thoughts or stimuli held in visual short term memory (Forster & Lavie, 2009; Konstantinou et al., 2012; Morris et al., 2020). We note that effects of visual perceptual load on visual processing have been demonstrated to occur as early as V1 (e.g. Schwartz et al., 2005) whereas, to date, attentional effects on auditory processing have only been found relatively later (Molloy et al., 2015). In this respect, the established effects of loading perceptual capacity through external stimuli are consistent with the pattern seen in the current data, where

the reduction in neuronal activity began in the earliest stages of visual processing, but later in the auditory cortex.

In the visual cortex only, an additional modality-specific competition was observed above and beyond attentional effects. This effect could potentially be attributed to competition arising from the overlap in neural substrates involved in generating visual mental imagery and conscious perception (Chen et al., 1998; Ganis et al., 2004; Halpern et al., 2004; Ishai et al., 2002; Klein et al., 2000; Lambert et al., 2002; Le Bihan et al., 1993; Naselaris et al., 2015). However, it is an intriguing question why similar competition induced suppression did not appear to occur in auditory cortex in response to auditory imagery. Our participants' ratings of their experienced imagery suggests that our scenarios were effective in eliciting very similar levels of imagery in the intended modality ($M = 3.75$ and $M = 3.71$ for amount of visual imagery in a visual scenario and auditory imagery in an auditory scenario, respectively), as well as being well matched for difficulty. In addition, we note that both our behavioural and online pilot ratings show a cleaner manipulation of auditory imagery compared with visual: While some (albeit lower) levels of visual imagery were reported even in response to the scenarios designed to elicit auditory imagery, almost no auditory imagery was elicited by the visual scenarios ($M = 1.64$ on a scale where 1 represents no imagery of that modality). It is therefore all the more striking that the relatively smaller difference in the level of visual imagery was sufficient to elicit robust modality-specific effects across the visual cortex. In contrast, despite almost no auditory imagery being reported in the visual imagery condition, this condition was as effective as the auditory imagery condition in attenuating the auditory cortex response. As noted in the introduction, the neural correlates of auditory mental imagery are less well established than their visual counterparts and there are differing opinions as to whether imagery related activation

extends to A1 (Zatorre & Halpern, 2005). As such, one possibility is that the neural mechanisms underlying auditory mental imagery show less overlap with mechanisms of external auditory perception, compared to the visual domain.

In the present research, no behavioural measure of perceptual processing was obtained, and thus an interesting direction for future research is to determine whether mental imagery can cause inattentional blindness or deafness (e.g. Macdonald & Lavie, 2008; Raveh & Lavie, 2015). Engaging in mind wandering, which may contain mental imagery, is a common occurrence during daily activities such as driving, particularly when traveling along a familiar route (Burdett et al., 2016). Our findings highlight that mental imagery can powerfully suppress external perceptual processing, with the potential implication that, when attention is directed towards our thoughts, the extent to which we see and hear external events (e.g. a red light or a siren) might be impeded (see Briggs et al., 2016, for a behavioural demonstration of such effects in the visual modality). Future research could examine whether any resulting behavioural effects of this reduction in processing are modality-specific, for example in term of awareness, or encoding an event into memory.

A second interesting question for future research is how these findings may be modulated by individual differences in the ability to produce mental imagery. There has been increasing recent research interest in the condition of *aphantasia*, which involves an inability to generate visual imagery, as well as reduced imagery in other sensory modalities including auditory (Dawes et al., 2020). One intriguing possibility is that this condition may confer an advantage, in terms of no reduction in sensory processing due to internal thoughts. Conversely, given that mental imagery has been demonstrated to be particularly likely to elicit responses in V1 when it is more vivid (Cui et al., 2007), we

might expect that exaggerated modality-specific competition would be experienced by individuals who experience vivid mental imagery, such as those with hyperphantasia.

In conclusion, we extend previous research through demonstrations of an imagery-induced attenuation of visual external perception as well as, for the first time, demonstrations of this effect on auditory perception. For visual sensory processing, this attenuation extends to the earliest stages of perceptual processing, and involves both amodal and modality-specific effects. On the other hand, attenuation of auditory external stimuli occurs only at later stages of perceptual processing, and does not appear to depend on the modality of the imagery. Thus, the present study implies two mechanism underlying the effects of mental imagery on external processing, in terms of competition for shared attentional resources and modality-specific resources, which have profound yet distinct effects on visual versus auditory perception.

Supplementary Materials

Online Scenario Timings

Participants

45 participants (43 female) aged between 18 and 27 years ($M = 19$, $SD = .82$) were recruited to participate in this experiment. All participants were undergraduate students at the University of Sussex, participating for course credits. 7 participants were excluded for not complying with the study instructions. The study was approved by the Science and Technology Cross-Schools Research Ethics Committee at the University of Sussex. All participants reported normal or corrected-to-normal, vision and hearing, fluency in English, and capability in producing both visual and auditory mental imagery.

Stimuli and Procedure

After providing demographic information and confirming that they met the inclusion criteria for the study, participants were instructed to read each scenario that would be presented to them throughout the experiment and press the space bar as soon as they had finished reading it, which would clear the screen to blank for 10 seconds whilst they imagined it for 10 seconds. If participants did not press the space bar, the questionnaire would move on after 10 seconds of it being presented. They were then asked to briefly describe what they had imagined. In total, 72 scenarios were presented – 36 designed to elicit visual imagery, and 36 designed to elicit auditory imagery. All descriptions were read by the experimenter and it was determined whether or not they had imagined something in line with the presented scenario, showing that they had read and understood it.

Results

Trials where participants did not press the space bar indicating they had finished reading it, were excluded. Mean time to read a scenario was 3.87 seconds ($SD = 2.08$), and 75% of all scenarios across all participants were read in under 5 seconds. Table S1 presents the cumulative frequency distribution for the time to read all scenarios, for all included participants.

Time (s)	Frequency
0-0.999	1
1-1.999	355
2-2.999	497
3-3.999	387
4-4.999	288
5-5.999	185
6-6.999	125
7-7.999	85
8-8.999	76
9-9.999	49

Table S1. Frequency table presenting the time in seconds to read all scenarios in the online experiment, for all participants

Online Scenario Ratings

Participants

69 participants (56 female) aged between 18 and 34 years ($M = 20.57$, $SD = 3.54$) were recruited to participate in this experiment. 9 were recruited from Prolific Academic (www.prolific.co) and were reimbursed for their time, and the remaining 60 were undergraduate students at the University of Sussex, participating for course

credits. 7 participants were excluded for reporting that they spent on average less than 3 seconds engaging in mental imagery per scenario. The study was approved by the Science and Technology Cross-Schools Research Ethics Committee at the University of Sussex. All participants reported normal or corrected-to-normal, vision and hearing, fluency in English, and capability in producing both visual and auditory mental imagery.

Stimuli and Procedure

After providing demographic information and confirming that they met the inclusion criteria for the study, participants were instructed to read each scenario that would be presented to them throughout the experiment, and then to imagine it as vividly as possible. Each scenario was presented in the centre of the screen for five seconds, and then ten seconds were allocated for imagining it whilst the screen was blank. After each scenario, participants provided ratings on a 5-point Likert Scale for difficulty in imagining it, amount of visual and auditory mental imagery experienced, and clarity of visual and auditory mental imagery. They were also asked about the contents of this imagery, whether or not they had any task unrelated thoughts throughout the time imagining the scenario, what percentage of the 10 seconds they spent engaging in mental imagery, and whether it was too much or too little time to think about it. In total, 72 scenarios were presented – 36 designed to elicit visual imagery, and 36 designed to elicit auditory imagery.

Results

A paired samples t-test revealed that there was no significant difference between the average difficulty in imagining a visual scenario ($M = 2.65$, $SD = .54$) and the difficulty imagining an auditory scenario ($M = 2.59$, $SD = .59$), $t(61) = 1.48$, $p = .143$.

The mean amount of visual imagery experienced by participants was entered into a 2 x 2 repeated measures ANOVA, with the factors of scenario type (visual, auditory) and imagery experienced (visual, auditory). The same pattern of results was observed as in the main study, with a significant main effect of scenario type, $F(1, 61) = 67.67$, $p < .001$, $\eta^2_p = .53$, a significant main effect of imagery experienced, $F(1, 61) = 193.29$, $p < .001$, $\eta^2_p = .76$, and a significant interaction between scenario type and imagery experienced, $F(1, 61) = 329.68$, $p < .001$, $\eta^2_p = .84$. Follow up t-tests confirmed that when the scenario was visual, participants experienced significantly more visual imagery than auditory imagery $t(61) = 19.54$, $p < .001$, and when the scenario was auditory, participants experienced significantly more auditory imagery than visual, $t(61) = 17.43$, $p < .001$. Additionally, participants experienced significantly more visual imagery when the scenario was visual compared to when it was auditory, $t(61) = 12.44$, $p < .001$, and significantly more auditory imagery when the scenario was auditory compared to when it was visual, $t(61) = 17.43$, $p = .002$.

From the 60 participants who experienced, on average, at least some of both imagery modalities in the two different scenario types, the clarity of this imagery was entered into a 2 x 2 repeated measures ANOVA, with the factors of scenario type (visual, auditory) and imagery experienced (visual, auditory). In this experiment, a main effect of scenario type was found, $F(1, 59) = 5.44$, $p = .023$, $\eta^2_p = .08$, with increased

overall clarity of the mental imagery for auditory scenarios than visual. There was also a significant main effect of clarity of the modality of the imagery, $F(1, 59) = 10.35, p = .002, \eta^2_p = .15$, with overall equal clarity of visual and auditory imagery across the experiment. Critically, as in the main study, there was a significant interaction between scenario type and imagery experienced, $F(1, 59) = 44.77, p < .001, \eta^2_p = .43$. Follow up t-tests confirmed that when the scenario was visual, participants experienced significantly clearer visual imagery than auditory imagery $t(59) = 5.74, p < .001$, and when the scenario was auditory, participants experienced significantly clearer auditory imagery than visual, $t(61) = 3.71, p < .001$. Additionally, participants experienced significantly clearer visual imagery when the scenario was visual compared to when it was auditory, $t(59) = 5.78, p < .001$, and significantly more auditory imagery when the scenario was auditory compared to when it was visual, $t(61) = 5.30, p < .001$.

Chapter 7: General Discussion

Throughout this thesis, I broadly aimed to investigate the extent to which multisensory interactions and integration impact upon attention, awareness, and perceptual processing. More specifically, throughout Chapters 2-5 I have investigated the proposed ‘special’ attentional status of multisensory stimuli, and the extent to which different levels of attention may influence their integration. Then in Chapter 6, I considered how the generation of sensory mental imagery may also affect processing of external stimuli. In this final chapter, I will summarise the key findings from the empirical chapters, relating them to the aims of the research set out in Chapter 1. The implications of this research, both theoretical and in real-world terms, will be discussed, and suggestions for future research will be made.

Are multisensory stimuli immune to the effects of perceptual load?

Given that multisensory stimuli have been proposed to have a ‘special’ attentional status, a key aim was to investigate whether or not attentional capture by these stimuli is dependent on the availability of perceptual capacity, as is the case for unisensory visual stimuli. In Chapter 2 we implemented, for the first time, established and controlled manipulations of perceptual load comparable to the ones traditionally used in unisensory perceptual load studies (e.g. Bahrami et al., 2007; Cunningham & Egeth, 2018; Dalton & Lavie, 2004; Forster et al., 2014; Forster & Lavie, 2008, 2009, 2011, 2014; He & Chen, 2010; Morris et al., 2020; Rees et al., 1997; Schwartz et al., 2005). In three of the experiments in this chapter, multisensory stimuli were presented peripherally as targets in a secondary task, along with unisensory visual and/or unisensory auditory stimuli. We found that, similarly to unisensory peripheral events,

reaction time to multisensory peripheral targets was slowed down when the central task was high in perceptual load, compared to low perceptual load. This modulation was found regardless of whether the peripheral targets were lower salience ‘flash and beep’ type events, or larger, colourful, and meaningful animal images with their corresponding sound. In the final experiment of this paper, the animal images were presented as irrelevant distractors, with instructions to ignore them. Again, just like distractor costs associated with unisensory visual distractors, costs due to multisensory distractors were reduced under high perceptual load. Hence, across both manipulations, and regardless of their presentation as secondary targets or distractors, the processing of multisensory stimuli was not immune to the effect of perceptual load.

Chapter 4 extends the research into the immunity of multisensory stimuli to these effects, in terms of awareness. Combining previously applied inattentional blindness and deafness paradigms to measure detection sensitivity (Macdonald & Lavie, 2008; Raveh & Lavie, 2015), this chapter examined whether a multisensory stimulus may be special in its ability to reach conscious awareness under perceptually demanding conditions. Following the principle of inverse effectiveness (e.g. Senkowski et al., 2011; Stevenson et al., 2012; Stevenson & James, 2009) we may have expected any special status of multisensory stimuli to be demonstrated here, where the unisensory constituent parts of the stimulus were weak and difficult to detect. However, the detection sensitivity of a multisensory critical stimulus did decrease under conditions of high perceptual load. In summary, it appears that both attentional capture by, and awareness of, multisensory stimuli is subject to perceptual capacity limits, to a similar extent to unisensory stimuli.

Does multisensory integration itself depend on attention?

As considered in Chapter 1, attention may be allocated to a stimulus for a number of different reasons, for example because we have remaining perceptual capacity when a task is less demanding (see Lavie, 1995), due to the stimulus itself being a target or matching the attentional set of a search task in some way (see Bacon & Egeth, 1994; Folk & Remington, 1998), or due to it being presented at a task-relevant location (i.e. inside our ‘attentional window’, Belopsky & Theeuwes, 2010). This thesis therefore aimed to explore what degree of attention may be sufficient for multisensory integration to occur.

In Chapter 2, whilst reaction time to detect a multisensory stimulus was slowed under conditions of high perceptual load, clear evidence of facilitatory attentional capture was observed even in this condition, in terms of both faster and more accurate detection compared to unisensory stimuli. Thus, despite more attentional resources being allocated to the primary task, multisensory stimuli still appeared able to integrate and elicit this behavioural advantage. In fact, significant race model violations were found only under high perceptual load, indicating that this advantage may be particularly pronounced in perceptually demanding situations.

In the same Chapter, we also manipulated the allocation of top-down attention towards multisensory stimuli. In Experiments 1-3, peripheral stimuli were presented as secondary search targets, and facilitatory attentional capture was afforded by the multisensory stimuli. On the other hand, in Experiment 4 participants were instructed to ignore these stimuli, and their peripheral location of presentation was entirely irrelevant. Here, we did not find any evidence of greater distractor interference by a multisensory stimulus over a unisensory visual one.

In Chapter 3, we furthered the research conducted on distraction by multisensory stimuli, using ERP measures. Despite finding no increased distractor interference due to multisensory stimuli in Chapter 2, it remained possible that our behavioural measure was not sensitive enough to detect faster attentional capture by multisensory stimuli. If multisensory stimuli do capture attention faster than unisensory stimuli, but also captivate this attention for longer, then this could result in a comparable distraction effect measured through reaction times. Therefore, we adapted the same distractor paradigm as used in Chapter 2 to be appropriate for measuring the N2pc and P_D components – respectively indexing a shift in spatial attention to, and spatial suppression of, a target (e.g. Hickey et al., 2008; Kiss et al., 2008). The results showed that, whilst an entirely irrelevant multisensory distractor elicits both a P_D and its auditory equivalent, the P_{AD}, indicative of active spatial suppression, these components do not occur at either a different latency or at an increased amplitude to those elicited by the unisensory constituents separately. This suggests that the two sensory modalities are suppressed independently, with no multisensory integration occurring which would likely require an enhanced, or even non-linear, suppression.

Taken together, these Chapters suggest a distinction between top-down attentional settings and perceptual load in multisensory integration. Whilst some degree of endogenous attention is required before integration (and hence any resulting attentional enhancement) may occur, the attentional resources available to be allocated does not, in contrast, appear to have an impact. Whilst reaction times to, and conscious awareness of, multisensory stimuli are modulated by perceptual load in a similar fashion to unisensory stimuli, a filled perceptual capacity does not appear to disrupt integration. In other words, whilst the effects of multisensory integration cannot prevent the effects of perceptual load, the effects of perceptual load also do not appear to prevent

multisensory integration – at least as measured by the race model. This conclusion was questioned to some extent by the ERP evidence in Chapter 5: Here, similar multisensory targets did not elicit superadditivity during a high load task. Thus, whilst we cannot conclude definitively that integration occurred using this more conservative criterion, they nevertheless can still show ‘special’ behavioural effects under these more demanding conditions.

Do multisensory stimuli differ in their ability to fill perceptual capacity?

Given that many tasks encountered in our daily lives occur in more than one modality at a time, Chapter 5 explored the effects of multisensory perceptual load. Here, we felt that two opposite outcomes were possible. Firstly, if multisensory stimuli elicit a superadditive response which is able to boost the biasing of top down attention towards the primary load task, this could make a particularly engaging form of perceptual load, resulting in an even more effective reduction of the processing of stimuli in a secondary task compared to unisensory perceptual load. On the other hand, given evidence that multisensory stimuli are processed more quickly (Colonius & Diederich, 2004; Hughes et al., 1994; Laurienti et al., 2004; Molholm et al., 2002; Murray et al., 2005; Senkowski et al., 2005) this could lower the perceptual load of the primary task, which would result in an increased availability of capacity to process the stimuli presented as peripheral targets. In fact, neither of these options appear to be the case. Multisensory stimuli seemingly do not differ in their ability to fill perceptual capacity compared with unisensory stimuli. In Experiment 3 of this chapter, we demonstrated non-linear, superadditive, responses to the multisensory load stimuli, and yet across three separate experiments we found no differences in the speed or accuracy of responses to peripheral targets during multisensory load as compared to unisensory load. Therefore, despite

forming the primary load task, and certainly being allocated top down attention, we have found multisensory stimuli to be no more effective than unisensory load in filling perceptual capacity, but neither do they impose on this capacity any less.

Additionally, in this chapter we sought to investigate whether the effects of multisensory load would selectively affect the response to multisensory stimuli in a secondary task, which may be expected if only one multisensory stimulus can be attentionally ‘boosted’ at a time (see Van der Burg et al., 2013). Across all three experiments we found that the multisensory behavioural advantage persists even under conditions of high multisensory perceptual load. Critically, in Experiments 2 and 3 we demonstrated that even under multisensory load reaction times to multisensory stimuli were even faster than that which would be predicted by the race model, though we note that this might not reflect integration, at least in terms of ERP superadditivity.

Can sensory mental imagery impose perceptual load?

Having explored the effects of perceptual capacity and the allocation of attentional resources to sensory stimuli presented externally, we then explored how this capacity, which appears to be both amodal and shared across the senses, may be filled by internally generated stimuli. Here we tested for two potential effects, an amodal effect driven by central perceptual load limits, and a modality-specific effect, whereby visual imagery would affect visual perceptual processing more than auditory imagery and vice versa. An account based on theoretical models of attention may predict that, with a limited perceptual capacity (Lavie, 2010) shared between internal and external sources (Forster & Lavie, 2009; Konstantinou et al., 2012; Morris et al., 2020), processing of external stimuli would be generally reduced when resources are diverted to internal sources, irrespective of modality. On top of this, a modality-specific

competition was proposed due to shared resources in the visual and auditory cortices between mental imagery and conscious perception (e.g. Chen et al., 1998; Ganis et al., 2004; Halpern et al., 2004; Ishai et al., 2002; Klein et al., 2000; Lambert et al., 2002; Le Bihan et al., 1993; Naselaris et al., 2015). Evidence in support of this comes from an ERP study demonstrating that a reduction in the P1 component during inner thought conditions compared with external attention to a visual stimulus is greater when the inner thoughts contain visual imagery as opposed to inner speech (Villena-González et al., 2016). However, a subsequent study found no effect of either of the internal thought conditions on the ERP response to an external auditory stimulus (Villena-González et al., 2018). On the other hand,

In Chapter 6, we found that neuronal responses to both visual and auditory external stimuli were reduced by both visual and auditory mental imagery, providing evidence for the amodal effect driven by perceptual capacity limits. The reduction in visual processing furthers the aforementioned ERP research (Villena-González et al., 2016), and the demonstration of an effect on auditory processing is, to the best of our knowledge, the first time this has been established. For external stimuli presented in the visual domain only, we additionally found evidence of a modality-specific effect, with neuronal responses to the visual stimuli being reduced more when imagery is also visual, compared to when it is auditory. Thus, the effect of imagery appears to be different for auditory versus visual perceptual processing. Whilst the effects of mental imagery on external processing of both visual and auditory stimuli may involve shared attentional capacity, it is only the effects on visual processing which additionally reflect competition in modality specific resources.

Implications

Theoretical implications for the attention literature

As outlined above, the modulation of multisensory stimuli by perceptual load, in terms of both reaction time (Chapter 2) and detection sensitivity (Chapter 4), extends the perceptual load literature by demonstrating that its effects extend to stimuli occurring in more than one modality at the same time. They do not appear to belong to the ‘special’ class of stimuli which are fully immune to the effects of perceptual load (e.g. human faces; Lavie et al., 2003), contradicting suggestions of this as a possibility (Santangelo & Spence, 2007). In addition to providing support for the load modulation of multisensory stimuli in terms of conscious awareness, in Chapter 4 we established the phenomenon of inattentional deafblindness. Furthering previous demonstrations of reduced detection sensitivity of both visual and auditory stimuli presented in the context of a secondary task, whilst engaged in a more demanding primary visual search task (Macdonald & Lavie, 2008; Raveh & Lavie, 2015), we show that both can occur simultaneously.

In Chapter 5, we also extended the Load Theory literature in terms of the effect of a multisensory load. As discussed above, multisensory stimuli did not appear to differ in their ability to fill perceptual capacity from unisensory stimuli, with responses to a secondary detection task being equal regardless of bimodality of the primary task. This is of great theoretical interest as it implies that the wealth of data already obtained on unisensory perceptual load would also apply to a multisensory, audiovisual load. Furthermore, the facilitatory effect of a multisensory stimulus presented as a secondary target also remains under these load conditions.

Further implications of the research presented in this thesis come from the results of Chapter 3. It has long been debated whether task-relevance is required for attentional capture to occur (e.g. see Folk et al., 1994; Theeuwes, 1992, 2010), and in this chapter we provided, to the best of our knowledge, the first demonstration that an entirely irrelevant visual distractor, both in terms of feature settings and location, elicited the P_D component – reflecting suppression following a spatial ‘attend-to-me’ signal (Hickey et al., 2008; Sawaki & Luck, 2010). Secondly, we established an auditory analogue of this component – the P_{AD} – which we propose reflects the same spatial suppression response to an entirely irrelevant auditory distractor. This finding supports previous demonstrations of distraction by entirely irrelevant auditory stimuli (Dalton & Lavie, 2004).

Theoretical implications for the multisensory integration literature

Our demonstrations of a behavioural advantage elicited by multisensory stimuli when they are allocated top-down attention due to being part of a secondary task, but no increased distractor interference when they were to be ignored (Chapter 2), provides further support for the suggestion that multisensory integration may be compromised when the stimuli are not attended (e.g. Alsius et al., 2005; Talsma et al., 2007). Chapter 3 was also consistent with this, demonstrating a lack of multisensory enhancement when a stimulus was presented as an irrelevant distractor, using ERP indices of active spatial suppression. Neither constituent part of a multisensory, audiovisual, stimulus elicited a greater (or earlier) ERP component indicative of this suppression, thus suggesting that they are processed independently, and neither sensory stimulus ‘boosts’ the other to require greater suppression to avoid distraction.

As noted above, in Chapter 5 we demonstrated that multisensory load does not disrupt the facilitatory behavioural advantage afforded by multisensory stimuli in a secondary task. The theoretical implications of this finding for the multisensory integration literature are twofold. Firstly, multisensory stimuli presented in the context of a primary task were found to elicit a superadditive neuronal response, but despite this, were not identified faster than targets in the unisensory perceptual load conditions. In contrast, no evidence of neuronal superadditivity was found for the peripheral targets, and yet reaction times towards these were faster than would be predicted by the race model (Raab, 1962). Whilst superadditivity is a commonly used metric to look for evidence of neural integration using ERPs, it has been noted that enhanced behavioural responding may be expected to occur even in the absence of a non-linear response (Stanford & Stein, 2007). The authors note that even linear summation of the two unisensory counterparts in the processing of a multisensory stimulus would represent an increase in activity in the superior colliculus compared to a unisensory stimulus alone. Our results are supportive of this and highlight that behavioural evidence of a multisensory advantage should not be assumed to reflect neuronal superadditivity, even in cases of race model violation. The second implication of this research speaks to the question of how many multisensory stimuli may be integrated at one time. Van der Burg et al. (2013b) proposed that integration can only occur for one item at a time, using a modified version of the ‘pip and pop’ paradigm, whereby a spatially uninformative auditory signal facilitates detection of a visual event. To test capacity limits, the number of visual events that were synchronised with the auditory signal was manipulated, and the authors found that only one instance of multisensory integration could occur at any one time. In Chapter 5, our experimental paradigm differed in that, rather than one auditory signal being presented for integration with multiple visual events, in the

multisensory load condition two pairs of visual and auditory events could be presented. Thus, our finding that a multisensory peripheral stimulus may be attentionally ‘boosted’ at the same time that a multisensory central stimulus demonstrates superadditivity may call into question the capacity limit of one item for integration at a time. However, given the differences in paradigms, we note that ‘capacity’ may differ depending on the stimuli presented, with the number of entirely independent multisensory items that can be integrated potentially differing from the number of unisensory stimuli that can be integrated with a single other stimulus presented in a different modality.

Real-world implications

In terms of real word significance, there are clear practical implications of the research conducted. Given that we rarely encounter purely unisensory tasks or purely unisensory events, it is critical to understand how multisensory stimuli capture our attention, as well as how competition from internal sensory stimuli may disrupt our perception. One consistent example mentioned throughout the literature is the proposal that multisensory stimuli may be particularly strong candidates for alerts and warning signals while driving (e.g. Ho et al., 2007; see Spence & Soto-Faraco, 2020 for a review). Our findings are very compatible with this proposal. Firstly, Chapter 2 suggests that if an alert comes from a location the driver is already monitoring, for example a place on the dashboard, then this being a multisensory signal could result in a faster reaction time to detect it, compared to a unisensory signal of equivalent strength. Secondly, Chapters 2 and 3 suggest that when driving a car an irrelevant multisensory event may be no more distracting than an already distracting unisensory one. Additionally, according to Chapter 2, whilst detection may still be slower when driving through a busy town (high perceptual load) than down an empty lane (low perceptual load) for either unisensory or multisensory warning signals, the multisensory advantage

would still be present in both circumstances. Chapter 3 suggests that this advantage may also present in terms of conscious awareness of the signal, though we note that very low salience items such as those employed in the inattention blindness and deafness literature would not be used as alerts. In Chapter 5 evidence is also provided to show that multisensory facilitation of attentional capture persists even under conditions of multisensory load, which is vital given that driving is in itself a multisensory task. Finally, given the increased prevalence of mind wandering whilst driving down familiar routes (Burdett et al., 2016), an act which has been correlated with increased likelihood of being involved in an accident (Qu et al., 2015), it is important to ascertain how sensory mental imagery which may be produced during these periods could disrupt our perception of the road ahead. The results presented in Chapter 6 suggest that sensory mental imagery may hinder the extent to which we see or hear critical events occurring on the road, particularly for events in the visual domain when the mental imagery is also visual, due to additional modality specific competition.

Future Directions

What level of task relevance is required for multisensory integration?

Whilst multisensory stimuli may not be quite as attentionally ‘special’ as has been proposed (Santangelo & Spence, 2007), there are further behavioural and neuronal facilitatory effects which should be explored. Firstly, given that it appears it is top-down attentional setting and not resources available due to perceptual demand that determine multisensory integration, it would be of interest to identify the degree of task relevance that is sufficient for this to occur and hence enhancement of attentional capture to be demonstrated. In Chapter 2, distractors were entirely irrelevant, sharing no features with

the targets and being presented in entirely irrelevant locations. If distractors were presented as non-targets within a search array, and thus appearing inside the attentional window (Theeuwes, 2010), this would extend the findings to determine whether there is a requirement for a stimulus to be part of the top-down attentional setting, or whether a more general direction of attention to the location of a multisensory event is enough for integration to occur. Similarly, presenting multisensory distractors that share a feature with the targets would establish whether matching the top-down attentional setting in some way would also be sufficient (i.e. ‘contingent capture’, see Folk et al., 1994).

If integration capacity is limited, what factors determine which stimuli will be combined?

Earlier in this chapter, I outlined the difference between the paradigm implemented in Chapter 5, and that employed in previous research to examine multisensory integration capacity. In the paradigm developed by Van der Burg et al. (2013b), one auditory event was presented for potential integration to occur with multiple visual events, and future research could examine the factors involved in determining which of the visual events the auditory cue is most likely to integrate with. This could be examined with particular reference to the principles of multisensory integration – inverse effectiveness, spatial and temporal matching, and semantic congruence. In the real world, it seems most logical that the latter three factors would influence which stimulus we would select for integration with another. For example, if we heard the sound of a dog at the same time as seeing both a dog and a cat, we would automatically assume the sound would be integrated with the sight of the dog. If presented with the sight of two dogs, we may be more likely to rely on the temporal and spatial coincidence of the sight and sound to determine which dog is barking. However,

there may also be other factors at play, such as crossmodal correspondences between pitch and size (Gallace & Spence, 2006; Spence, 2011). Wilbiks et al. (2020) determined that crossmodal congruency between brightness of a visual stimulus and pitch of an auditory tone effects integration capacity, opening the door for further exploration into the aforementioned factors, particularly in the context of more ecologically valid stimuli, such as the animal stimuli we selected for Chapters 2 and 3, which are more familiar and meaningful than the commonly used ‘flash and beep’ events.

Are there individual differences in the effects demonstrated in this thesis?

Individual differences have been examined within the context of unisensory perceptual load and attention. For example, individuals who experience greater distraction in their daily lives also demonstrate increased distractibility in a lab-based visual search task with irrelevant distractors (Forster & Lavie, 2007), as do those who self-report an increased susceptibility to engage in mind wandering (Forster & Lavie, 2014). Additionally, whilst high perceptual load reduced distraction to a similar extent, adults with ADHD demonstrate greater distraction in a visual search task compared with controls (Forster et al., 2014). On the other hand, for individuals with autism spectrum disorder (ASD), a higher level of perceptual load is required to reduce distraction than for controls (Remington et al., 2009), indicative of an increased perceptual capacity for these individuals, which also results in enhanced performance under conditions of high load in an awareness paradigm (Remington et al., 2012). Throughout this thesis, we demonstrate both that multisensory stimuli are modulated by perceptual load, in terms of both attentional capture and awareness, to a similar extent to unisensory stimuli. Additionally, they do not appear to fill capacity any differently to

unisensory stimuli when presented in a loading task. For this reason, we would likely predict that the findings of individual differences studies conducted on unisensory stimuli would also apply to those when stimuli are presented in two different modalities at the same time.

In Chapter 6, we consider that individual differences in the ability to produce mental imagery may affect the level of reduction in response to external stimuli observed due to this imagery. We may expect that all individuals capable of producing mental imagery would demonstrate some level of attenuation, due to the competition for attentional resources induced by attending to internal thought. However, the additional modality specific competition resulting in an increased reduction of visual processing when the imagery is also visual, attributed to shared sensory resources in the visual cortex (e.g. Chen et al., 1998; Ganis et al., 2004; Halpern et al., 2004; Ishai et al., 2002; Klein et al., 2000; Lambert et al., 2002; Le Bihan et al., 1993; Naselaris et al., 2015) may differ depending on the clarity of the imagery that an individual experiences, given that the likelihood of imagery eliciting responses in the primary visual cortex is greater when the imagery is more vivid.

Do the findings of this thesis apply to multisensory stimuli that are not audiovisual?

Finally, throughout this thesis I have examined multisensory integration and interactions, where the multisensory stimulus was always audiovisual, and primary tasks were also always presented in either one, or both, of these modalities. Given that the effects of multisensory integration on orienting responses seem to be similar for different combinations of modalities (see Spence, 2001, for a review), we may expect that the findings of this thesis would also extend to other types of bimodal stimuli. Additionally, research has demonstrated that trimodal (visual, auditory, haptic) stimuli

show an advantage over bimodal stimuli in terms of reaction time (Diederich & Colonius, 2004; Hecht et al., 2008). However, given the evidence presented in this thesis regarding a load modulation of multisensory stimuli even in the absence of a disruption to behavioural enhancement resulting from integration, gives us no reason to assume that integrating more stimuli at a time would confer an immunity to perceptual load. Nevertheless, further research would be required to confirm this.

Conclusions

In summary, this thesis contributes to research examining the attentional status of multisensory stimuli, within the Perceptual Load Theory framework. Speed of detection, distractor interference, and conscious awareness of multisensory stimuli were all modulated by perceptual load, as is the case for unisensory stimuli. Additionally, multisensory stimuli do not appear to impose on perceptual capacity any differently than unisensory stimuli, when they are themselves presented in a primary task, despite eliciting a superadditive neuronal response. We support suggestions of an interplay between multisensory integration and attention through demonstrations of a bidirectional relationship between these two processes, whereby multisensory integration may enhance spatial attentional capture, provided we are already exerting some degree of endogenous attention towards them. Finally, we reveal a general effect of perceptual capacity limits on the attenuation of external sensory stimuli due to engagement in mental imagery, as well as a modality specific attenuation occurring only for visual external stimuli, and differences in how early in the processing stream such effects occur.

References

- Alsius, A., Möttönen, R., Sams, M. E., Soto-Faraco, S., & Tiippana, K. (2014). Effect of attentional load on audiovisual speech perception: Evidence from ERPs. *Frontiers in Psychology*, 5. <https://doi.org/10.3389/fpsyg.2014.00727>
- Alsius, A., Navarra, J., Campbell, R., & Soto-Faraco, S. (2005). Audiovisual Integration of Speech Alters under High Attention Demands. *Current Biology*, 15(9), 839–843. <https://doi.org/10.1016/j.cub.2005.03.046>
- Amlôt, R., Walker, R., Driver, J., & Spence, C. (2003). Multimodal visual–somatosensory integration in saccade generation. *Neuropsychologia*, 41(1), 1–15. [https://doi.org/10.1016/S0028-3932\(02\)00139-2](https://doi.org/10.1016/S0028-3932(02)00139-2)
- Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences*, 108(25), 10367–10371. <https://doi.org/10.1073/pnas.1104047108>
- Andersson, J. L. R., Hutton, C., Ashburner, J., Turner, R., & Friston, K. (2001). Modeling Geometric Deformations in EPI Time Series. *NeuroImage*, 13(5), 903–919. <https://doi.org/10.1006/nimg.2001.0746>
- Andersson, J. L. R., Skare, S., & Ashburner, J. (2003). How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging. *NeuroImage*, 20(2), 870–888. [https://doi.org/10.1016/S1053-8119\(03\)00336-7](https://doi.org/10.1016/S1053-8119(03)00336-7)
- Andrade, J., May, J., Deeprase, C., Baugh, S.-J., & Ganis, G. (2014). Assessing vividness of mental imagery: The Plymouth Sensory Imagery Questionnaire. *British Journal of Psychology*, 105(4), 547–563. <https://doi.org/10.1111/bjop.12050>

- Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics*, 55(5), 485–496.
<https://doi.org/10.3758/BF03205306>
- Bahrami, B., Carmel, D., Walsh, V., Rees, G., & Lavie, N. (2007). Attentional load modulates subconscious orientation processing. *Journal of Vision*, 7(9), 788–788. <https://doi.org/10.1167/7.9.788>
- Barron, E., Riby, L. M., Greer, J., & Smallwood, J. (2011). Absorbed in Thought: The Effect of Mind Wandering on the Processing of Relevant and Irrelevant Events. *Psychological Science*, 22(5), 596–601.
<https://doi.org/10.1177/0956797611404083>
- Beauchamp, M. S. (2019). Using Multisensory Integration to Understand the Human Auditory Cortex. In A. K. C. Lee, M. T. Wallace, A. B. Coffin, A. N. Popper, & R. R. Fay (Eds.), *Multisensory Processes: The Auditory Perspective* (pp. 161–176). Springer International Publishing. https://doi.org/10.1007/978-3-030-10461-0_8
- Belopolsky, A. V., & Theeuwes, J. (2010). No capture outside the attentional window. *Vision Research*, 50(23), 2543–2550.
<https://doi.org/10.1016/j.visres.2010.08.023>
- Berggren, N., & Eimer, M. (2018). Electrophysiological correlates of active suppression and attentional selection in preview visual search. *Neuropsychologia*, 120, 75–85. <https://doi.org/10.1016/j.neuropsychologia.2018.10.016>
- Bertelson, P., Pavani, F., Ladavas, E., Vroomen, J., & de Gelder, B. (2000). Ventriloquism in patients with unilateral visual neglect. *Neuropsychologia*, 38(12), 1634–1642. [https://doi.org/10.1016/S0028-3932\(00\)00067-1](https://doi.org/10.1016/S0028-3932(00)00067-1)

- Bertelson, P., & Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Perception & Psychophysics*, 29(6), 578–584. <https://doi.org/10.3758/BF03207374>
- Bertelson, P., Vroomen, J., De Gelder, B., & Driver, J. (2000). The ventriloquist effect does not depend on the direction of deliberate visual attention. *Perception & Psychophysics*, 62(2), 321–332. <https://doi.org/10.3758/BF03205552>
- Besle, J., Fort, A., & Giard, M.-H. (2004). Interest and validity of the additive model in electrophysiological studies of multisensory interactions. *Cognitive Processing*, 5(3). <https://doi.org/10.1007/s10339-004-0026-y>
- Besle, J., Bertrand, O., & Giard, M.-H. (2009). Electrophysiological (EEG, sEEG, MEG) evidence for multiple audiovisual interactions in the human auditory cortex. *Hearing Research*, 258(1), 143–151. <https://doi.org/10.1016/j.heares.2009.06.016>
- Blair, R. C., & Karniski, W. (1993). An alternative method for significance testing of waveform difference potentials. *Psychophysiology*, 30(5), 518–524. <https://doi.org/10.1111/j.1469-8986.1993.tb02075.x>
- Braboszcz, C., & Delorme, A. (2011). Lost in thoughts: Neural markers of low alertness during mind wandering. *NeuroImage*, 54(4), 3040–3047. <https://doi.org/10.1016/j.neuroimage.2010.10.008>
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433–436.
- Briggs, G. F., Hole, G. J., & Land, M. F. (2016). Imagery-inducing distraction leads to cognitive tunnelling and deteriorated driving performance. *Transportation Research Part F: Traffic Psychology and Behaviour*, 38, 106–117. <https://doi.org/10.1016/j.trf.2016.01.007>

- Bruckmaier, M., Tachtsidis, I., Phan, P., & Lavie, N. (2020). Attention and capacity limits in perception: A cellular metabolism account. *Journal of Neuroscience*. <https://doi.org/10.1523/JNEUROSCI.2368-19.2020>
- Burdett, B. R. D., Charlton, S. G., & Starkey, N. J. (2016). Not all minds wander equally: The influence of traits, states and road environment factors on self-reported mind wandering during everyday driving. *Accident Analysis & Prevention*, 95, 1–7. <https://doi.org/10.1016/j.aap.2016.06.012>
- Burra, N., & Kerzel, D. (2014). The distractor positivity (Pd) signals lowering of attentional priority: Evidence from event-related potentials and individual differences. *Psychophysiology*, 51(7), 685–696. <https://doi.org/10.1111/psyp.12215>
- Calvert, G. A. (2001). Crossmodal Processing in the Human Brain: Insights from Functional Neuroimaging Studies. *Cerebral Cortex*, 11(12), 1110–1123. <https://doi.org/10.1093/cercor/11.12.1110>
- Calvert, G. A., Campbell, R., & Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current Biology*, 10(11), 649–657. [https://doi.org/10.1016/S0960-9822\(00\)00513-3](https://doi.org/10.1016/S0960-9822(00)00513-3)
- Cappe, C., Thut, G., Romei, V., & Murray, M. M. (2010). Auditory-Visual Multisensory Interactions in Humans: Timing, Topography, Directionality, and Sources. *Journal of Neuroscience*, 30(38), 12572–12580. <https://doi.org/10.1523/JNEUROSCI.1099-10.2010>
- Cartwright-Finch, U., & Lavie, N. (2007). The role of perceptual load in inattention blindness. *Cognition*, 102(3), 321–340. <https://doi.org/10.1016/j.cognition.2006.01.002>

- Chen, W., Kato, T., Zhu, X.-H., Ogawa, S., Tank, D. W., & Ugurbil, K. (1998). Human primary visual cortex and lateral geniculate nucleus activation during visual imagery. *NeuroReport*, 9(16), 3669–3674.
- Cherry, E. C. (1953). Some Experiments on the Recognition of Speech, with One and with Two Ears. *The Journal of the Acoustical Society of America*, 25(5), 975–979. <https://doi.org/10.1121/1.1907229>
- Chiou, R., & Rich, A. N. (2012). Cross-Modality Correspondence between Pitch and Spatial Location Modulates Attentional Orienting. *Perception*, 41(3), 339–353. <https://doi.org/10.1068/p7161>
- Colonus, H., & Arndt, P. (2001). A two-stage model for visual-auditory interaction in saccadic latencies. *Perception & Psychophysics*, 63(1), 126–147. <https://doi.org/10.3758/BF03200508>
- Colonus, H., & Diederich, A. (2004). Multisensory Interaction in Saccadic Reaction Time: A Time-Window-of-Integration Model. *Journal of Cognitive Neuroscience*, 16(6), 1000–1009. <https://doi.org/10.1162/0898929041502733>
- Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, 1(1), 42–45. <https://doi.org/10.20982/tqmp.01.1.p042>
- Crosse, M. J., Foxe, J. J., & Molholm, S. (2019). Developmental Recovery of Impaired Multisensory Processing in Autism and the Cost of Switching Sensory Modality. *BioRxiv*, 565333. <https://doi.org/10.1101/565333>
- Cui, X., Jeter, C. B., Yang, D., Montague, P. R., & Eagleman, D. M. (2007). Vividness of mental imagery: Individual variability can be measured objectively. *Vision Research*, 47(4), 474–478. <https://doi.org/10.1016/j.visres.2006.11.013>

- Cunningham, C. A., & Egeth, H. E. (2018). The capture of attention by entirely irrelevant pictures of calorie-dense foods. *Psychonomic Bulletin & Review*, 25(2), 586–595. <https://doi.org/10.3758/s13423-017-1375-8>
- Dal Ben, R. (2019). *SHINE_color and Lum_fun: A set of tools to control luminance of colored images* (Version 0.2) [Computer program]. <https://doi.org/10.17605/OSF.IO/AUZJY>. Retrieved from <https://osf.io/auzjy/>
- Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical Surface-Based Analysis: I. Segmentation and Surface Reconstruction. *NeuroImage*, 9(2), 179–194. <https://doi.org/10.1006/nimg.1998.0395>
- Dale, A. M., & Sereno, M. I. (1993). Improved Localization of Cortical Activity by Combining EEG and MEG with MRI Cortical Surface Reconstruction: A Linear Approach. *Journal of Cognitive Neuroscience*, 5(2), 162–176. <https://doi.org/10.1162/jocn.1993.5.2.162>
- Dalton, P., & Fraenkel, N. (2012). Gorillas we have missed: Sustained inattentional deafness for dynamic events. *Cognition*, 124(3), 367–372. <https://doi.org/10.1016/j.cognition.2012.05.012>
- Dalton, P., & Lavie, N. (2004). Auditory Attentional Capture: Effects of Singleton Distractor Sounds. *Journal of Experimental Psychology: Human Perception and Performance*, 30(1), 180–193. <https://doi.org/10.1037/0096-1523.30.1.180>
- Dawes, A. J., Keogh, R., Andriillon, T., & Pearson, J. (2020). A cognitive profile of multi-sensory imagery, memory and dreaming in aphantasia. *Scientific Reports*, 10(1), 10022. <https://doi.org/10.1038/s41598-020-65705-7>
- Debener, S., Herrmann, C. S., Kranczioch, C., Gembris, D., & Engel, A. K. (2003). Top-down attentional processing enhances auditory evoked gamma band activity. *NeuroReport*, 14(5), 683–686.

- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9–21.
<https://doi.org/10.1016/j.jneumeth.2003.10.009>
- Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review*, 70(1), 80–90. <https://doi.org/10.1037/h0039515>
- Diederich, A., & Colonius, H. (2004). Bimodal and trimodal multisensory enhancement: Effects of stimulus onset and intensity on reaction time. *Perception & Psychophysics*, 66(8), 1388–1404.
<https://doi.org/10.3758/BF03195006>
- Diederich, A., Colonius, H., Bockhorst, D., & Tabeling, S. (2003). Visual-tactile spatial interaction in saccade generation. *Experimental Brain Research*, 148(3), 328–337. <https://doi.org/10.1007/s00221-002-1302-7>
- Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, 5. <https://doi.org/10.3389/fpsyg.2014.00781>
- Doehrmann, O., & Naumer, M. J. (2008). Semantics and the multisensory brain: How meaning modulates processes of audio-visual integration. *Brain Research*, 1242, 136–150. <https://doi.org/10.1016/j.brainres.2008.03.071>
- Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review*, 87(3), 272–300. <https://doi.org/10.1037/0033-295X.87.3.272>
- Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalography and Clinical Neurophysiology*, 99(3), 225–234.
[https://doi.org/10.1016/0013-4694\(96\)95711-9](https://doi.org/10.1016/0013-4694(96)95711-9)

- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 16(1), 143–149. <https://doi.org/10.3758/BF03203267>
- Evans, K. K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *Journal of Vision*, 10(1), 6–6. <https://doi.org/10.1167/10.1.6>
- Fairnie, J., Moore, B. C. J., & Remington, A. (2016). Missing a trick: Auditory load modulates conscious awareness in audition. *Journal of Experimental Psychology: Human Perception and Performance*, 42(7), 930–938. <https://doi.org/10.1037/xhp0000204>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Fischl, B. (2012). FreeSurfer. *NeuroImage*, 62(2), 774–781. <https://doi.org/10.1016/j.neuroimage.2012.01.021>
- Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences*, 97(20), 11050–11055. <https://doi.org/10.1073/pnas.200033797>
- Fischl, B., Salat, D. H., van der Kouwe, A. J. W., Makris, N., Ségonne, F., Quinn, B. T., & Dale, A. M. (2004). Sequence-independent segmentation of magnetic resonance images. *NeuroImage*, 23, S69–S84. <https://doi.org/10.1016/j.neuroimage.2004.07.016>
- Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, 18(4), 1030–1044.

- Folk, Charles L., & Remington, R. (1998). Selectivity in distraction by irrelevant featural singletons: Evidence for two forms of attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*; Washington, 24(3), 847–858. <http://dx.doi.org/10.1037/0096-1523.24.3.847>
- Folk, Charles L., Remington, R. W., & Wright, J. H. (1994). The structure of attentional control: Contingent attentional capture by apparent motion, abrupt onset, and color. *Journal of Experimental Psychology: Human Perception and Performance*, 20(2), 317–329. <http://dx.doi.org/10.1037/0096-1523.20.2.317>
- Forster, S. (2013). Distraction and Mind-Wandering Under Load. *Frontiers in Psychology*, 4. <https://doi.org/10.3389/fpsyg.2013.00283>
- Forster, S., & Lavie, N. (2007). High Perceptual Load Makes Everybody Equal. *Psychological Science*, 18(5), 377–381. <https://doi.org/10.1111/j.1467-9280.2007.01908.x>
- Forster, S., & Lavie, N. (2008a). Failures to ignore entirely irrelevant distractors: The role of load. *Journal of Experimental Psychology: Applied*, 14(1), 73–83. <https://doi.org/10.1037/1076-898X.14.1.73>
- Forster, S., & Lavie, N. (2008b). Attentional capture by entirely irrelevant distractors. *Visual Cognition*, 16(2–3), 200–214. <https://doi.org/10.1080/13506280701465049>
- Forster, S., & Lavie, N. (2009). Harnessing the wandering mind: The role of perceptual load. *Cognition*, 111(3), 345–355. <https://doi.org/10.1016/j.cognition.2009.02.006>
- Forster, S., & Lavie, N. (2011). Entirely irrelevant distractors can capture and captivate attention. *Psychonomic Bulletin & Review*, 18(6), 1064–1070. <https://doi.org/10.3758/s13423-011-0172-z>

- Forster, S., & Lavie, N. (2014). Distracted by your mind? Individual differences in distractibility predict mind wandering. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(1), 251–260.
<https://doi.org/10.1037/a0034108>
- Forster, S., & Lavie, N. (2016). Establishing the Attention-Distractibility Trait. *Psychological Science*, 27(2), 203–212.
<https://doi.org/10.1177/0956797615617761>
- Forster, S., Robertson, D. J., Jennings, A., Asherson, P., & Lavie, N. (2014). Plugging the attention deficit: Perceptual load counters increased distraction in ADHD. *Neuropsychology*, 28(1), 91–97. <https://doi.org/10.1037/neu0000020>
- Forster, S., & Spence, C. (2018). “What Smell?” Temporarily Loading Visual Attention Induces a Prolonged Loss of Olfactory Awareness. *Psychological Science*, 29(10), 1642–1652. <https://doi.org/10.1177/0956797618781325>
- Fort, A., Delpuech, C., Pernier, J., & Giard, M.-H. (2002). Dynamics of Cortico-subcortical Cross-modal Operations Involved in Audio-visual Object Detection in Humans. *Cerebral Cortex*, 12(10), 1031–1039.
<https://doi.org/10.1093/cercor/12.10.1031>
- Fortier-Gauthier, U., Moffat, N., Dell’Acqua, R., McDonald, J. J., & Jolicœur, P. (2012). Contralateral cortical organisation of information in visual short-term memory: Evidence from lateralized brain activity during retrieval. *Neuropsychologia*, 50(8), 1748–1758.
<https://doi.org/10.1016/j.neuropsychologia.2012.03.032>
- Foxe, J. J., Morocz, I. A., Murray, M. M., Higgins, B. A., Javitt, D. C., & Schroeder, C. E. (2000). Multisensory auditory–somatosensory interactions in early cortical

- processing revealed by high-density electrical mapping. *Cognitive Brain Research*, 10(1), 77–83. [https://doi.org/10.1016/S0926-6410\(00\)00024-0](https://doi.org/10.1016/S0926-6410(00)00024-0)
- Foxe, J. J., & Snyder, A. C. (2011). The Role of Alpha-Band Brain Oscillations as a Sensory Suppression Mechanism during Selective Attention. *Frontiers in Psychology*, 2. <https://doi.org/10.3389/fpsyg.2011.00154>
- Franconeri, S. L., & Simons, D. J. (2003). Moving and looming stimuli capture attention. *Perception & Psychophysics*, 65(7), 999–1010. <https://doi.org/10.3758/BF03194829>
- Frassinetti, F., Bolognini, N., & Làdavas, E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Experimental Brain Research*, 147(3), 332–343. <https://doi.org/10.1007/s00221-002-1262-y>
- Freeman, L. C. A., Wood, K. C., & Bizley, J. K. (2018). Multisensory stimuli improve relative localisation judgments compared to unisensory auditory or visual stimuli. *The Journal of the Acoustical Society of America*, 143(6), EL516–EL522. <https://doi.org/10.1121/1.5042759>
- Frens, M. A., Van Opstal, A. J., & Van Der Willigen, R. F. (1995). Spatial and temporal factors determine auditory-visual interactions in human saccadic eye movements. *Perception & Psychophysics*, 57(6), 802–816. <https://doi.org/10.3758/BF03206796>
- Gallace, A., & Spence, C. (2006). Multisensory synesthetic interactions in the speeded classification of visual size. *Perception & Psychophysics*, 68(7), 1191–1203. <https://doi.org/10.3758/BF03193720>
- Gamble, M. L., & Luck, S. J. (2011). N2ac: An ERP component associated with the focusing of attention within an auditory scene. *Psychophysiology*, 48(8), 1057–1068. <https://doi.org/10.1111/j.1469-8986.2010.01172.x>

- Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2004). Brain areas underlying visual mental imagery and visual perception: An fMRI study. *Cognitive Brain Research*, 20(2), 226–241. <https://doi.org/10.1016/j.cogbrainres.2004.02.012>
- Gaspar, J. M., Christie, G. J., Prime, D. J., Jolicœur, P., & McDonald, J. J. (2016). Inability to suppress salient distractors predicts low visual working memory capacity. *Proceedings of the National Academy of Sciences*, 113(13), 3693–3698. <https://doi.org/10.1073/pnas.1523471113>
- Gaspar, J. M., & McDonald, J. J. (2014). Suppression of Salient Objects Prevents Distraction in Visual Search. *Journal of Neuroscience*, 34(16), 5658–5666. <https://doi.org/10.1523/JNEUROSCI.4161-13.2014>
- Gaspelin, N., Leonard, C. J., & Luck, S. J. (2015). Direct Evidence for Active Suppression of Salient-but-Irrelevant Sensory Inputs. *Psychological Science*, 26(11), 1740–1750. <https://doi.org/10.1177/0956797615597913>
- Gaspelin, N., & Luck, S. J. (2018). Combined Electrophysiological and Behavioral Evidence for the Suppression of Salient Distractors. *Journal of Cognitive Neuroscience*, 30(9), 1265–1280. https://doi.org/10.1162/jocn_a_01279
- Geden, M., Staicu, A.-M., & Feng, J. (2018). The impacts of perceptual load and driving duration on mind wandering in driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 57, 75–83. <https://doi.org/10.1016/j.trf.2017.07.004>
- Giard, M. H., & Peronnet, F. (1999). Auditory-Visual Integration during Multimodal Object Recognition in Humans: A Behavioral and Electrophysiological Study. *Journal of Cognitive Neuroscience*, 11(5), 473–490. <https://doi.org/10.1162/089892999563544>

- Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C. F., Jenkinson, M., Smith, S. M., & Van Essen, D. C. (2016). A multi-modal parcellation of human cerebral cortex. *Nature*, 536(7615), 171–178. <https://doi.org/10.1038/nature18933>
- Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology*, 48(12), 1711–1725. <https://doi.org/10.1111/j.1469-8986.2011.01273.x>
- Halpern, A. R., Zatorre, R. J., Bouffard, M., & Johnson, J. A. (2004). Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia*, 42(9), 1281–1292. <https://doi.org/10.1016/j.neuropsychologia.2003.12.017>
- Hartcher-O'Brien, J., Soto-Faraco, S., & Adam, R. (2017). Editorial: A Matter of Bottom-Up or Top-Down Processes: The Role of Attention in Multisensory Integration. *Frontiers in Integrative Neuroscience*, 11. <https://doi.org/10.3389/fnint.2017.00005>
- He, C., & Chen, A. (2010). Interference from familiar natural distractors is not eliminated by high perceptual load. *Psychological Research PRPF*, 74(3), 268–276. <https://doi.org/10.1007/s00426-009-0252-0>
- Hecht, D., Reiner, M., & Karni, A. (2008). Enhancement of response times to bi- and tri-modal sensory stimuli during active movements. *Experimental Brain Research*, 185(4), 655–665. <https://doi.org/10.1007/s00221-007-1191-x>
- Hein, G., Doehrmann, O., Muller, N. G., Kaiser, J., Muckli, L., & Naumer, M. J. (2007). Object Familiarity and Semantic Congruency Modulate Responses in Cortical Audiovisual Integration Areas. *Journal of Neuroscience*, 27(30), 7881–7887. <https://doi.org/10.1523/JNEUROSCI.1740-07.2007>

- Hemmelmann, C., Horn, M., Reiterer, S., Schack, B., Süsse, T., & Weiss, S. (2004). Multivariate tests for the evaluation of high-dimensional EEG data. *Journal of Neuroscience Methods*, 139(1), 111–120.
<https://doi.org/10.1016/j.jneumeth.2004.04.013>
- Hershenson, M. (1962). Reaction time as a measure of intersensory facilitation. *Journal of Experimental Psychology*, 63(3), 289–293. <https://doi.org/10.1037/h0039516>
- Hickey, C., Di Lollo, V., & McDonald, J. J. (2008). Electrophysiological Indices of Target and Distractor Processing in Visual Search. *Journal of Cognitive Neuroscience*, 21(4), 760–775. <https://doi.org/10.1162/jocn.2009.21039>
- Hickey, C., McDonald, J. J., & Theeuwes, J. (2006). Electrophysiological Evidence of the Capture of Visual Attention. *Journal of Cognitive Neuroscience*, 18(4), 604–613. <https://doi.org/10.1162/jocn.2006.18.4.604>
- Hillyard, S. A., Teder-Sälejärvi, W. A., & Münte, T. F. (1998). Temporal dynamics of early perceptual processing. *Current Opinion in Neurobiology*, 8(2), 202–210.
[https://doi.org/10.1016/S0959-4388\(98\)80141-4](https://doi.org/10.1016/S0959-4388(98)80141-4)
- Hipp, J. F., Engel, A. K., & Siegel, M. (2011). Oscillatory Synchronization in Large-Scale Cortical Networks Predicts Perception. *Neuron*, 69(2), 387–396.
<https://doi.org/10.1016/j.neuron.2010.12.027>
- Ho, C., Reed, N., & Spence, C. (2007). Multisensory In-Car Warning Signals for Collision Avoidance. *Human Factors*, 49(6), 1107–1114.
<https://doi.org/10.1518/001872007X249965>
- Holmes, N. P., & Spence, C. (2005). Multisensory Integration: Space, Time and Superadditivity. *Current Biology*, 15(18), R762–R764.
<https://doi.org/10.1016/j.cub.2005.08.058>

- Hughes, H. C., Reuter-Lorenz, P. A., Nozawa, G., & Fendrich, R. (1994). Visual-auditory interactions in sensorimotor processing: Saccades versus manual responses. *Journal of Experimental Psychology. Human Perception and Performance*, 20(1), 131–153.
- Hughes, Howard C, Nelson, M. D., & Aronchick, D. M. (1998). Spatial characteristics of visual-auditory summation in human saccades. *Vision Research*, 38(24), 3955–3963. [https://doi.org/10.1016/S0042-6989\(98\)00036-4](https://doi.org/10.1016/S0042-6989(98)00036-4)
- Ishai, A., Haxby, J. V., & Ungerleider, L. G. (2002). Visual Imagery of Famous Faces: Effects of Memory and Attention Revealed by fMRI. *NeuroImage*, 17(4), 1729–1741. <https://doi.org/10.1006/nimg.2002.1330>
- JASP Team. (2020). *JASP* (0.13.1) [Computer software].
- Jeffreys, H. (1961). *Theory of Probability*.
- Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics*, 43(4), 346–354. <https://doi.org/10.3758/BF03208805>
- Kadunce, D., Vaughan, W., Wallace, M., & Stein, B. (2001). The influence of visual and auditory receptive field organization on multisensory integration in the superior colliculus. *Experimental Brain Research*, 139(3), 303–310. <https://doi.org/10.1007/s002210100772>
- Kahneman, D. (1973). *Attention and effort*. Prentice-Hall.
- Kam, J. W. Y., Dao, E., Stanciulescu, M., Tildesley, H., & Handy, T. C. (2013). Mind Wandering and the Adaptive Control of Attentional Resources. *Journal of Cognitive Neuroscience*, 25(6), 952–960. https://doi.org/10.1162/jocn_a_00375

- Kam, J. W. Y., & Handy, T. C. (2013). The neurocognitive consequences of the wandering mind: A mechanistic account of sensory-motor decoupling. *Frontiers in Psychology*, 4. <https://doi.org/10.3389/fpsyg.2013.00725>
- Kay, K., Rokem, A., Winawer, J., Dougherty, R., & Wandell, B. (2013). GLMdenoise: A fast, automated technique for denoising task-based fMRI data. *Frontiers in Neuroscience*, 7. <https://doi.org/10.3389/fnins.2013.00247>
- Keogh, R., & Pearson, J. (2011). Mental Imagery and Visual Working Memory. *PLOS ONE*, 6(12), e29221. <https://doi.org/10.1371/journal.pone.0029221>
- Killingsworth, M. A., & Gilbert, D. T. (2010). A Wandering Mind Is an Unhappy Mind. *Science*, 330(6006), 932–932. <https://doi.org/10.1126/science.1192439>
- Kinchla, R. A. (1974). Detecting target elements in multielement arrays: A confusability model. *Perception & Psychophysics*, 15(1), 149–158. <https://doi.org/10.3758/BF03205843>
- Kiss, M., Velzen, J. V., & Eimer, M. (2008). The N2pc component and its links to attention shifts and spatially selective visual processing. *Psychophysiology*, 45(2), 240–249. <https://doi.org/10.1111/j.1469-8986.2007.00611.x>
- Klein, I., Paradis, A.-L., Poline, J.-B., Kosslyn, S. M., & Le Bihan, D. (2000). Transient Activity in the Human Calcarine Cortex During Visual-Mental Imagery: An Event-Related fMRI Study. *Journal of Cognitive Neuroscience*, 12(supplement 2), 15–23. <https://doi.org/10.1162/089892900564037>
- Kleiner, M., Brainard, D. H., & Pelli, D. G. (2007). “What’s new in Psychtoolbox-3?” *Perception 36 ECVF Abstract Supplement*.
- Konstantinou, N., Bahrami, B., Rees, G., & Lavie, N. (2012). Visual Short-term Memory Load Reduces Retinotopic Cortex Response to Contrast. *Journal of*

Cognitive Neuroscience, 24(11), 2199–2210.

https://doi.org/10.1162/jocn_a_00279

Kosslyn, S. M., & Thompson, W. L. (2003). When is early visual cortex activated during visual mental imagery? *Psychological Bulletin*, 129(5), 723–746.

<https://doi.org/10.1037/0033-2909.129.5.723>

Kvasova, D., & Soto-Faraco, S. (2019). *Not so automatic: Task relevance and perceptual load modulate cross-modal semantic congruence effects on spatial orienting* [Preprint]. Neuroscience. <https://doi.org/10.1101/830679>

Lambert, S., Sampaio, E., Scheiber, C., & Mauss, Y. (2002). Neural substrates of animal mental imagery: Calcarine sulcus and dorsal pathway involvement — an fMRI study. *Brain Research*, 924(2), 176–183. [https://doi.org/10.1016/S0006-8993\(01\)03232-2](https://doi.org/10.1016/S0006-8993(01)03232-2)

Larson, G. E., Alderton, D. L., Neideffer, M., & Underhill, E. (1997). Further evidence on dimensionality and correlates of the Cognitive Failures Questionnaire. *British Journal of Psychology*, 88(1), 29–38. <https://doi.org/10.1111/j.2044-8295.1997.tb02618.x>

Laurienti, P. J., Burdette, J. H., Wallace, M. T., Yen, Y.-F., Field, A. S., & Stein, B. E. (2002). Deactivation of Sensory-Specific Cortex by Cross-Modal Stimuli. *Journal of Cognitive Neuroscience*, 14(3), 420–429.

<https://doi.org/10.1162/089892902317361930>

Laurienti, P. J., Kraft, R. A., Maldjian, J. A., Burdette, J. H., & Wallace, M. T. (2004). Semantic congruence is a critical factor in multisensory behavioral performance. *Experimental Brain Research*, 158(4), 405–414. <https://doi.org/10.1007/s00221-004-1913-2>

Laurienti, P. J., Perrault, T. J., Stanford, T. R., Wallace, M. T., & Stein, B. E. (2005).

On the use of superadditivity as a metric for characterizing multisensory integration in functional neuroimaging studies. *Experimental Brain Research*, 166(3–4), 289–297. <https://doi.org/10.1007/s00221-005-2370-2>

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention.

Journal of Experimental Psychology: Human Perception and Performance, 21(3), 451–468. <https://doi.org/10.1037/0096-1523.21.3.451>

Lavie, N. (2010). Attention, Distraction, and Cognitive Control Under Load. *Current*

Directions in Psychological Science, 19(3), 143–148.

<https://doi.org/10.1177/0963721410370295>

Lavie, N., & Cox, S. (1997). On the Efficiency of Visual Selective Attention: Efficient

Visual Search Leads to Inefficient Distractor Rejection. *Psychological Science*, 8(5), 395–396. <https://doi.org/10.1111/j.1467-9280.1997.tb00432.x>

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load Theory of Selective

Attention and Cognitive Control. *Journal of Experimental Psychology: General*, 133(3), 339–354. <https://doi.org/10.1037/0096-3445.133.3.339>

Lavie, N., Ro, T., & Russell, C. (2003). The Role of Perceptual Load in Processing

Distractor Faces. *Psychological Science*, 14(5), 510–515.

<https://doi.org/10.1111/1467-9280.03453>

Le Bihan, D., Turner, R., Zeffiro, T. A., Cuénod, C. A., Jezzard, P., & Bonnerot, V.

(1993). Activation of human primary visual cortex during visual recall: A magnetic resonance imaging study. *Proceedings of the National Academy of Sciences*, 90(24), 11802–11805. <https://doi.org/10.1073/pnas.90.24.11802>

Leo, F., Bertini, C., di Pellegrino, G., & Làdavas, E. (2008). Multisensory integration

for orienting responses in humans requires the activation of the superior

colliculus. *Experimental Brain Research*, 186(1), 67–77.

<https://doi.org/10.1007/s00221-007-1204-9>

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, 8.

<https://doi.org/10.3389/fnhum.2014.00213>

Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique*. MIT Press.

Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, 31(3), 291–308.

<https://doi.org/10.1111/j.1469-8986.1994.tb02218.x>

Lunn, J., Sjoblom, A., Ward, J., Soto-Faraco, S., & Forster, S. (2019). Multisensory enhancement of attention depends on whether you are already paying attention.

Cognition, 187, 38–49. <https://doi.org/10.1016/j.cognition.2019.02.008>

Macdonald, J. S. P., & Lavie, N. (2008). Load induced blindness. *Journal of Experimental Psychology: Human Perception and Performance*, 34(5), 1078–1091. <https://doi.org/10.1037/0096-1523.34.5.1078>

Macdonald, J. S. P., & Lavie, N. (2011). Visual perceptual load induces inattentional deafness. *Attention, Perception, & Psychophysics*, 73(6), 1780–1789.

<https://doi.org/10.3758/s13414-011-0144-4>

Mack, A., & Rock, I. (1998). Inattention blindness: Perception without attention. In *Visual attention* (pp. 55–76). Oxford University Press.

Mack, A., Rock, I., & Research), A. (Professor M., The New School for Social. (1998). *Inattentional Blindness*. MIT Press.

Mackay, D. G. (1973). *Aspects of the Theory of Comprehension, Memory and Attention*.

- Manly, B. F. J. (1997). *Randomization, Bootstrap, and Monte Carlo Methods in Biology* (2nd ed.). Chapman & Hall.
- Marks, D. F. (1973). Visual Imagery Differences in the Recall of Pictures. *British Journal of Psychology*, 64(1), 17–24. <https://doi.org/10.1111/j.2044-8295.1973.tb01322.x>
- MATLAB (Version 2019B). (2019). [Computer software]. The MathWorks Inc.
- Matusz, P. J., Broadbent, H., Ferrari, J., Forrest, B., Merkley, R., & Scerif, G. (2015). Multi-modal distraction: Insights from children's limited attention. *Cognition*, 136, 156–165. <https://doi.org/10.1016/j.cognition.2014.11.031>
- McDonald, J. J., Green, J. J., Jannati, A., & Di Lollo, V. (2013). On the electrophysiological evidence for the capture of visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, 39(3), 849–860. <http://dx.doi.org/10.1037/a0030510>
- McEvoy, S. P., Stevenson, M. R., & Woodward, M. (2007). The prevalence of, and factors associated with, serious crashes involving a distracting activity. *Accident Analysis & Prevention*, 39(3), 475–482. <https://doi.org/10.1016/j.aap.2006.09.005>
- Mcgurk, H., & Macdonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264(5588), 746–748. <https://doi.org/10.1038/264746a0>
- Meredith, M. A., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *Journal of Neuroscience*, 7(10), 3215–3229. <https://doi.org/10.1523/JNEUROSCI.07-10-03215.1987>
- Meredith, M. A., & Stein, B. E. (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration.

Journal of Neurophysiology, 56(3), 640–662.

<https://doi.org/10.1152/jn.1986.56.3.640>

Meredith, M. A., & Stein, B. E. (1996). Spatial determinants of multisensory integration in cat superior colliculus neurons. *Journal of Neurophysiology*, 75(5), 1843–1857. <https://doi.org/10.1152/jn.1996.75.5.1843>

Meredith, M., & Stein, B. (1983). Interactions among converging sensory inputs in the superior colliculus. *Science*, 221(4608), 389–391. <https://doi.org/10.1126/science.6867718>

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

Miller, J. (1986). Timecourse of coactivation in bimodal divided attention. *Perception & Psychophysics*, 40(5), 331–343. <https://doi.org/10.3758/BF03203025>

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). Multisensory auditory–visual interactions during early sensory processing in humans: A high-density electrical mapping study. *Cognitive Brain Research*, 14(1), 115–128. [https://doi.org/10.1016/S0926-6410\(02\)00066-6](https://doi.org/10.1016/S0926-6410(02)00066-6)

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. <https://doi.org/10.1523/JNEUROSCI.2931-15.2015>

Morey, R. D. (2008). Confidence Intervals from Normalized Data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, 4(2), 61–64. <https://doi.org/10.20982/tqmp.04.2.p061>

Morís Fernández, L., Visser, M., Ventura-Campos, N., Ávila, C., & Soto-Faraco, S. (2015). Top-down attention regulates the neural expression of audiovisual integration. *NeuroImage*, *119*, 272–285.

<https://doi.org/10.1016/j.neuroimage.2015.06.052>

Morris, J., Vi, C. T., Obrist, M., Forster, S., & Yeomans, M. R. (2020). Ingested but not perceived: Response to satiety cues disrupted by perceptual load. *Appetite*, *155*, 104813. <https://doi.org/10.1016/j.appet.2020.104813>

Morosan, P., Schleicher, A., Amunts, K., & Zilles, K. (2005). Multimodal architectonic mapping of human superior temporal gyrus. *Anatomy and Embryology*, *210*(5–6), 401–406. <https://doi.org/10.1007/s00429-005-0029-1>

Most, S. B., Simons, D. J., Scholl, B. J., Jimenez, R., Clifford, E., & Chabris, C. F. (2001). How not to be Seen: The Contribution of Similarity and Selective Ignoring to Sustained Inattentional Blindness. *Psychological Science*, *12*(1), 9–17. <https://doi.org/10.1111/1467-9280.00303>

Murphy, G., & Greene, C. M. (2015). High perceptual load causes inattentional blindness and deafness in drivers. *Visual Cognition*, *23*(7), 810–814. <https://doi.org/10.1080/13506285.2015.1093245>

Murphy, G., & Greene, C. M. (2016). Perceptual Load Induces Inattentional Blindness in Drivers. *Applied Cognitive Psychology*, *30*(3), 479–483. <https://doi.org/10.1002/acp.3216>

Murphy, S., & Dalton, P. (2016). Out of touch? Visual load induces inattentional numbness. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(6), 761. <https://doi.org/10.1037/xhp0000218>

Murphy, S., Fraenkel, N., & Dalton, P. (2013). Perceptual load does not modulate auditory distractor processing. *Cognition*, 129(2), 345–355.

<https://doi.org/10.1016/j.cognition.2013.07.014>

Murphy, S., Spence, C., & Dalton, P. (2017). Auditory perceptual load: A review.

Hearing Research, 352, 40–48. <https://doi.org/10.1016/j.heares.2017.02.005>

Murray, M. M., Molholm, S., Michel, C. M., Heslenfeld, D. J., Ritter, W., Javitt, D. C.,

Schroeder, C. E., & Foxe, J. J. (2005). Grabbing Your Ear: Rapid Auditory–

Somatosensory Multisensory Interactions in Low-level Sensory Cortices Are

Not Constrained by Stimulus Alignment. *Cerebral Cortex*, 15(7), 963–974.

<https://doi.org/10.1093/cercor/bhh197>

Näätänen, R., & Picton, T. (1987). The N1 Wave of the Human Electric and Magnetic

Response to Sound: A Review and an Analysis of the Component Structure.

Psychophysiology, 24(4), 375–425. <https://doi.org/10.1111/j.1469->

8986.1987.tb00311.x

Naselaris, T., Olman, C. A., Stansbury, D. E., Ugurbil, K., & Gallant, J. L. (2015). A

voxel-wise encoding model for early visual areas decodes mental images of

remembered scenes. *NeuroImage*, 105, 215–228.

<https://doi.org/10.1016/j.neuroimage.2014.10.018>

Newby, E. A., & Rock, I. (1998). Inattentional Blindness as a Function of Proximity to

the Focus of Attention. *Perception*, 27(9), 1025–1040.

<https://doi.org/10.1068/p271025>

Otto, T. U., & Mamassian, P. (2012). Noise and Correlations in Parallel Perceptual

Decision Making. *Current Biology*, 22(15), 1391–1396.

<https://doi.org/10.1016/j.cub.2012.05.031>

- Pannunzi, M., Pérez-Bellido, A., Pereda-Baños, A., López-Moliner, J., Deco, G., & Soto-Faraco, S. (2014). Deconstructing multisensory enhancement in detection. *Journal of Neurophysiology*, *113*(6), 1800–1818.
<https://doi.org/10.1152/jn.00341.2014>
- Pápai, M. S., & Soto-Faraco, S. (2017). Sounds can boost the awareness of visual events through attention without cross-modal integration. *Scientific Reports*, *7*, 41684.
<https://doi.org/10.1038/srep41684>
- Pearson, J., Naselaris, T., Holmes, E. A., & Kosslyn, S. M. (2015). Mental Imagery: Functional Mechanisms and Clinical Applications. *Trends in Cognitive Sciences*, *19*(10), 590–602. <https://doi.org/10.1016/j.tics.2015.08.003>
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.
- Pérez-Bellido, A., Soto-Faraco, S., & López-Moliner, J. (2012). Sound-driven enhancement of vision: Disentangling detection-level from decision-level contributions. *Journal of Neurophysiology*, *109*(4), 1065–1077.
<https://doi.org/10.1152/jn.00226.2012>
- Posner, M. I. (1980). Orienting of Attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. <https://doi.org/10.1080/00335558008248231>
- Purkis, H. M., Lester, K. J., Link to external site, this link will open in a new window, Field, A. P., & Link to external site, this link will open in a new window. (2011). But what about the Empress of Racnoss? The allocation of attention to spiders and Doctor Who in a visual search task is predicted by fear and expertise. *Emotion*, *11*(6), 1484–1488. <http://dx.doi.org/10.1037/a0024415>
- Qu, W., Ge, Y., Xiong, Y., Carciofo, R., Zhao, W., & Zhang, K. (2015). The relationship between mind wandering and dangerous driving behavior among

Chinese drivers. *Safety Science*, 78, 41–48.

<https://doi.org/10.1016/j.ssci.2015.04.016>

Raab, D. H. (1962). Division of Psychology: Statistical Facilitation of Simple Reaction Times*. *Transactions of the New York Academy of Sciences*, 24(5 Series II), 574–590. <https://doi.org/10.1111/j.2164-0947.1962.tb01433.x>

Raveh, D., & Lavie, N. (2015). Load-induced inattentional deafness. *Attention, Perception, & Psychophysics*, 77(2), 483–492. <https://doi.org/10.3758/s13414-014-0776-2>

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, 18(3), 849–860. <http://dx.doi.org/10.1037/0096-1523.18.3.849>

Rees, G., Frith, C. D., & Lavie, N. (1997). Modulating Irrelevant Motion Perception by Varying Attentional Load in an Unrelated Task. *Science*, 278(5343), 1616–1619. <https://doi.org/10.1126/science.278.5343.1616>

Remington, A. M., Swettenham, J. G., & Lavie, N. (2012). Lightening the load: Perceptual load impairs visual detection in typical adults but not in autism. *Journal of Abnormal Psychology*, 121(2), 544. <https://doi.org/10.1037/a0027670>

Remington, A., Swettenham, J., Campbell, R., & Coleman, M. (2009). Selective Attention and Perceptual Load in Autism Spectrum Disorder. *Psychological Science*, 20(11), 1388–1393. <https://doi.org/10.1111/j.1467-9280.2009.02454.x>

Roa Romero, Y., Senkowski, D., & Keil, J. (2015). Early and late beta-band power reflect audiovisual perception in the McGurk illusion. *Journal of Neurophysiology*, 113(7), 2342–2350. <https://doi.org/10.1152/jn.00783.2014>

- Rouder, J. N. (2014). Optional stopping: No problem for Bayesians. *Psychonomic Bulletin & Review*, 21(2), 301–308. <https://doi.org/10.3758/s13423-014-0595-4>
- Santangelo, V., Botta, F., Lupiáñez, J., & Spence, C. (2011). The time course of attentional capture under dual-task conditions. *Attention, Perception, & Psychophysics*, 73(1), 15–23. <https://doi.org/10.3758/s13414-010-0017-2>
- Santangelo, V., Ho, C., & Spence, C. (2008). Capturing spatial attention with multisensory cues. *Psychonomic Bulletin & Review*, 15(2), 398–403. <https://doi.org/10.3758/PBR.15.2.398>
- Santangelo, V., & Spence, C. (2007). Multisensory cues capture spatial attention regardless of perceptual load. *Journal of Experimental Psychology: Human Perception and Performance*, 33(6), 1311–1321. <https://doi.org/10.1037/0096-1523.33.6.1311>
- Sawaki, R., Geng, J. J., & Luck, S. J. (2012). A Common Neural Mechanism for Preventing and Terminating the Allocation of Attention. *Journal of Neuroscience*, 32(31), 10725–10736. <https://doi.org/10.1523/JNEUROSCI.1864-12.2012>
- Sawaki, R., Kreither, J., Leonard, C. J., Kaiser, S. T., Hahn, B., Gold, J. M., & Luck, S. J. (2017). Hyperfocusing of attention on goal-related information in schizophrenia: Evidence from electrophysiology. *Journal of Abnormal Psychology*, 126(1), 106–116. <http://dx.doi.org/10.1037/abn0000209>
- Sawaki, R., & Luck, S. J. (2010). Capture versus suppression of attention by salient singletons: Electrophysiological evidence for an automatic attend-to-me signal. *Attention, Perception, & Psychophysics*, 72(6), 1455–1470. <https://doi.org/10.3758/APP.72.6.1455>

- Sawaki, R., & Luck, S. J. (2013). Active suppression after involuntary capture of attention. *Psychonomic Bulletin & Review*, 20(2), 296–301.
<https://doi.org/10.3758/s13423-012-0353-4>
- Schwartz, S., Vuilleumier, P., Hutton, C., Maravita, A., Dolan, R. J., & 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(6), 770–786.
<https://doi.org/10.1093/cercor/bhh178>
- Ségonne, F., Dale, A. M., Busa, E., Glessner, M., Salat, D., Hahn, H. K., & Fischl, B. (2004). A hybrid approach to the skull stripping problem in MRI. *NeuroImage*, 22(3), 1060–1075. <https://doi.org/10.1016/j.neuroimage.2004.03.032>
- Senkowski, D., Saint-Amour, D., Höfle, M., & Foxe, J. J. (2011). Multisensory interactions in early evoked brain activity follow the principle of inverse effectiveness. *NeuroImage*, 56(4), 2200–2208.
<https://doi.org/10.1016/j.neuroimage.2011.03.075>
- Senkowski, D., Talsma, D., Herrmann, C. S., & Woldorff, M. G. (2005). Multisensory processing and oscillatory gamma responses: Effects of spatial selective attention. *Experimental Brain Research*, 166(3), 411–426.
<https://doi.org/10.1007/s00221-005-2381-z>
- Simons, D. J. (2000). Attentional capture and inattention blindness. *Trends in Cognitive Sciences*, 4(4), 147–155. [https://doi.org/10.1016/S1364-6613\(00\)01455-8](https://doi.org/10.1016/S1364-6613(00)01455-8)
- Simons, D. J., & Chabris, C. F. (1999). Gorillas in Our Midst: Sustained Inattentional Blindness for Dynamic Events. *Perception*, 28(9), 1059–1074.
<https://doi.org/10.1068/p281059>

- Skelton, A. E., Catchpole, G., Abbott, J. T., Bosten, J. M., & Franklin, A. (2017). Biological origins of color categorization. *Proceedings of the National Academy of Sciences*, 114(21), 5545–5550. <https://doi.org/10.1073/pnas.1612881114>
- Smallwood, J., McSpadden, M., & Schooler, J. W. (2008). When attention matters: The curious incident of the wandering mind. *Memory & Cognition*, 36(6), 1144–1150. <https://doi.org/10.3758/MC.36.6.1144>
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M., & Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, 23, S208–S219. <https://doi.org/10.1016/j.neuroimage.2004.07.051>
- Soto-Faraco, S., Kvasova, D., Biau, E., Ikumi, N., Ruzzoli, M., Morís-Fernández, L., & Torralba, M. (2019, August). *Multisensory Interactions in the Real World*. Elements in Perception; Cambridge University Press. <https://doi.org/10.1017/9781108578738>
- Soto-Faraco, S., Navarra, J., & Alsius, A. (2004). Assessing automaticity in audiovisual speech integration: Evidence from the speeded classification task. *Cognition*, 92(3), B13–B23. <https://doi.org/10.1016/j.cognition.2003.10.005>
- Spence, C. (2001). Crossmodal attentional capture: A controversy resolved? In *Attraction, distraction and action: Multiple perspectives on attentional capture* (pp. 231–262). Elsevier Science. [https://doi.org/10.1016/S0166-4115\(01\)80012-](https://doi.org/10.1016/S0166-4115(01)80012-1)

- Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, 73(4), 971–995. <https://doi.org/10.3758/s13414-010-0073-7>
- Spence, C. (2013). Just how important is spatial coincidence to multisensory integration? Evaluating the spatial rule. *Annals of the New York Academy of Sciences*, 1296(1), 31–49. <https://doi.org/10.1111/nyas.12121>
- Spence, C., & Frings, C. (2019). Multisensory feature integration in (and out) of the focus of spatial attention. *Attention, Perception, & Psychophysics*. <https://doi.org/10.3758/s13414-019-01813-5>
- Spence, C., & Santangelo, V. (2009). Capturing spatial attention with multisensory cues: A review. *Hearing Research*, 258(1), 134–142. <https://doi.org/10.1016/j.heares.2009.04.015>
- Spence, C., & Soto-Faraco, S. (2020). *Crossmodal Attention Applied: Lessons for and from Driving* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/9781108919951>
- Stanford, T. R., Quessy, S., & Stein, B. E. (2005). Evaluating the Operations Underlying Multisensory Integration in the Cat Superior Colliculus. *Journal of Neuroscience*, 25(28), 6499–6508. <https://doi.org/10.1523/JNEUROSCI.5095-04.2005>
- Stanford, Terrence R., & Stein, B. E. (2007). Superadditivity in multisensory integration: Putting the computation in context. *Neuroreport*, 18(8), 787–792. <https://doi.org/10.1097/WNR.0b013e3280c1e315>
- Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses* (pp. xv, 211). The MIT Press.

- Stein, B. E., Meredith, M. A., Huneycutt, W. S., & McDade, L. (1989). Behavioral Indices of Multisensory Integration: Orientation to Visual Cues is Affected by Auditory Stimuli. *Journal of Cognitive Neuroscience*, 1(1), 12–24.
<https://doi.org/10.1162/jocn.1989.1.1.12>
- Stein, B. E., Scott Huneycutt, W., & Alex Meredith, M. (1988). Neurons and behavior: The same rules of multisensory integration apply. *Brain Research*, 448(2), 355–358. [https://doi.org/10.1016/0006-8993\(88\)91276-0](https://doi.org/10.1016/0006-8993(88)91276-0)
- Stein, B. E., Stanford, T. R., Ramachandran, R., Perrault, T. J., & Rowland, B. A. (2009). Challenges in quantifying multisensory integration: Alternative criteria, models, and inverse effectiveness. *Experimental Brain Research*, 198(2–3), 113. <https://doi.org/10.1007/s00221-009-1880-8>
- Stein, B. E., Stanford, T. R., & Rowland, B. A. (2014). Development of multisensory integration from the perspective of the individual neuron. *Nature Reviews. Neuroscience*, 15(8), 520–535.
- Stevenson, R. A., Bushmakina, M., Kim, S., Wallace, M. T., Puce, A., & James, T. W. (2012). Inverse Effectiveness and Multisensory Interactions in Visual Event-Related Potentials with Audiovisual Speech. *Brain Topography*, 25(3), 308–326. <https://doi.org/10.1007/s10548-012-0220-7>
- Stevenson, R. A., & James, T. W. (2009). Audiovisual integration in human superior temporal sulcus: Inverse effectiveness and the neural processing of speech and object recognition. *NeuroImage*, 44(3), 1210–1223. <https://doi.org/10.1016/j.neuroimage.2008.09.034>
- Sundfør, H. B., Sagberg, F., & Høy, A. (2019). Inattention and distraction in fatal road crashes – Results from in-depth crash investigations in Norway. *Accident Analysis & Prevention*, 125, 152–157. <https://doi.org/10.1016/j.aap.2019.02.004>

- Talsma, D., Doty, T. J., & Woldorff, M. G. (2007). Selective Attention and Audiovisual Integration: Is Attending to Both Modalities a Prerequisite for Early Integration? *Cerebral Cortex*, *17*(3), 679–690. <https://doi.org/10.1093/cercor/bhk016>
- Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences*, *14*(9), 400–410. <https://doi.org/10.1016/j.tics.2010.06.008>
- Talsma, D., & Woldorff, M. G. (2005). Selective Attention and Multisensory Integration: Multiple Phases of Effects on the Evoked Brain Activity. *Journal of Cognitive Neuroscience*, *17*(7), 1098–1114. <https://doi.org/10.1162/0898929054475172>
- Teder-Sälejärvi, W. A., McDonald, J. J., Di Russo, F., & Hillyard, S. A. (2002). An analysis of audio-visual crossmodal integration by means of event-related potential (ERP) recordings. *Cognitive Brain Research*, *14*(1), 106–114. [https://doi.org/10.1016/S0926-6410\(02\)00065-4](https://doi.org/10.1016/S0926-6410(02)00065-4)
- Tellinghuisen, D. J., & Nowak, E. J. (2003). The inability to ignore auditory distractors as a function of visual task perceptual load. *Perception & Psychophysics*, *65*(5), 817–828. <https://doi.org/10.3758/BF03194817>
- ten Oever, S., Romei, V., van Atteveldt, N., Soto-Faraco, S., Murray, M. M., & Matusz, P. J. (2016). The COGs (context, object, and goals) in multisensory processing. *Experimental Brain Research*, *234*(5), 1307–1323. <https://doi.org/10.1007/s00221-016-4590-z>
- Theeuwes, J. (1991). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. *Perception & Psychophysics*, *49*(1), 83–90. <https://doi.org/10.3758/BF03211619>

- Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, 51(6), 599–606. <https://doi.org/10.3758/BF03211656>
- Theeuwes, J. (2010). Top–down and bottom–up control of visual selection. *Acta Psychologica*, 135(2), 77–99. <https://doi.org/10.1016/j.actpsy.2010.02.006>
- Theeuwes, J., Kramer, A. F., Hahn, S., & Irwin, D. E. (1998). Our Eyes do Not Always Go Where we Want Them to Go: Capture of the Eyes by New Objects. *Psychological Science*, 9(5), 379–385. <https://doi.org/10.1111/1467-9280.00071>
- Thomas, J. M., Huber, E., Stecker, G. C., Boynton, G. M., Saenz, M., & Fine, I. (2015). Population receptive field estimates of human auditory cortex. *NeuroImage*, 105, 428–439. <https://doi.org/10.1016/j.neuroimage.2014.10.060>
- Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, 12(4), 242–248. <https://doi.org/10.1080/17470216008416732>
- Triarhou, L. C. (2007). A Proposed Number System for the 107 Cortical Areas of Economo and Koskinas, and Brodmann Area Correlations. *Stereotactic and Functional Neurosurgery*, 85(5), 204–215. <https://doi.org/10.1159/000103259>
- Ulrich, R., Miller, J., & Schröter, H. (2007). Testing the race model inequality: An algorithm and computer programs. *Behavior Research Methods*, 39(2), 291–302. <https://doi.org/10.3758/BF03193160>
- Van der Burg, E., Awh, E., & Olivers, C. N. L. (2013a). The Capacity of Audiovisual Integration Is Limited to One Item. *Psychological Science*, 24(3), 345–351. <https://doi.org/10.1177/0956797612452865>
- Van der Burg, E., Awh, E., & Olivers, C. N. L. (2013b). The Capacity of Audiovisual Integration Is Limited to One Item. *Psychological Science*, 24(3), 345–351. <https://doi.org/10.1177/0956797612452865>

- Van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W., & Theeuwes, J. (2008). Pip and pop: Nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 34(5), 1053–1065. <https://doi.org/10.1037/0096-1523.34.5.1053>
- Van der Burg, E., Talsma, D., Olivers, C. N. L., Hickey, C., & Theeuwes, J. (2011). Early multisensory interactions affect the competition among multiple visual objects. *NeuroImage*, 55(3), 1208–1218. <https://doi.org/10.1016/j.neuroimage.2010.12.068>
- Vetter, P., Smith, F. W., & Muckli, L. (2014). Decoding Sound and Imagery Content in Early Visual Cortex. *Current Biology*, 24(11), 1256–1262. <https://doi.org/10.1016/j.cub.2014.04.020>
- Villena-González, M., & Cosmelli, D. (2020). Chapter 5 - Imagination and mind wandering: Two sides of the same coin? A brain dynamics perspective. In D. D. Preiss, D. Cosmelli, & J. C. Kaufman (Eds.), *Creativity and the Wandering Mind* (pp. 93–120). Academic Press. <https://doi.org/10.1016/B978-0-12-816400-6.00005-5>
- Villena-González, M., López, V., & Rodríguez, E. (2016). Orienting attention to visual or verbal/auditory imagery differentially impairs the processing of visual stimuli. *NeuroImage*, 132, 71–78. <https://doi.org/10.1016/j.neuroimage.2016.02.013>
- Villena-González, M., Palacios-García, I., Rodríguez, E., & López, V. (2018). Beta Oscillations Distinguish Between Two Forms of Mental Imagery While Gamma and Theta Activity Reflects Auditory Attention. *Frontiers in Human Neuroscience*, 12. <https://doi.org/10.3389/fnhum.2018.00389>

- von Economo, C. F., & Koskinas, G. N. (1925). *Die cytoarchitektonik der hirnrinde des erwachsenen menschen*. J. Springer.
- Wagenmakers, E.-J., Verhagen, J., Ly, A., Matzke, D., Steingroever, H., Rouder, J. N., & Morey, R. D. (2017). The need for Bayesian hypothesis testing in psychological science. In *Psychological science under scrutiny: Recent challenges and proposed solutions* (pp. 123–138). Wiley-Blackwell.
<https://doi.org/10.1002/9781119095910.ch8>
- Wallace, Mark T., Meredith, M. A., & Stein, Barry E. (1992). Integration of multiple sensory modalities in cat cortex. *Experimental Brain Research*, 91(3).
<https://doi.org/10.1007/BF00227844>
- Weaver, M. D., van Zoest, W., & Hickey, C. (2017). A temporal dependency account of attentional inhibition in oculomotor control. *NeuroImage*, 147, 880–894.
<https://doi.org/10.1016/j.neuroimage.2016.11.004>
- Wilbiks, J. M. P., & Dyson, B. J. (2016). The Dynamics and Neural Correlates of Audio-Visual Integration Capacity as Determined by Temporal Unpredictability, Proactive Interference, and SOA. *PLOS ONE*, 11(12), e0168304.
<https://doi.org/10.1371/journal.pone.0168304>
- Wilbiks, J. M. P., & Dyson, B. J. (2018). The contribution of perceptual factors and training on varying audiovisual integration capacity. *Journal of Experimental Psychology: Human Perception and Performance*, 44(6), 871–884.
<https://doi.org/10.1037/xhp0000503>
- Wilbiks, J. M. P., Pavilanis, A. D. S., & Rioux, D. M. (2020). Audiovisual integration capacity modulates as a function of illusory visual contours, visual display circumference, and sound type. *Attention, Perception, & Psychophysics*, 82(4), 1971–1986. <https://doi.org/10.3758/s13414-019-01882-6>

- Wundersitz, L. (2019). Driver distraction and inattention in fatal and injury crashes: Findings from in-depth road crash data. *Traffic Injury Prevention*, 20(7), 696–701. <https://doi.org/10.1080/15389588.2019.1644627>
- Xie, S., Kaiser, D., & Cichy, R. M. (2020). Visual Imagery and Perception Share Neural Representations in the Alpha Frequency Band. *Current Biology*, 30(13), 2621–2627.e5. <https://doi.org/10.1016/j.cub.2020.04.074>
- Zatorre, R. J., & Halpern, A. R. (2005). Mental Concerts: Musical Imagery and Auditory Cortex. *Neuron*, 47(1), 9–12. <https://doi.org/10.1016/j.neuron.2005.06.013>
- Zeman, A., Dewar, M., & Della Sala, S. (2015). Lives without imagery – Congenital aphantasia. *Cortex*, 73, 378–380. <https://doi.org/10.1016/j.cortex.2015.05.019>