Mechanisms of unimodal and cross-modal interference control

Rebecca J. Hirst, BSc., MSc.

Thesis submitted to the University of Nottingham for the degree of *Doctor of Philosophy*

November 2018

Thesis Abstract

Background: *In everyday life we must frequently ignore distractions arising from multiple senses. However, most of our understanding about this cognitive process (known as interference control) is derived from unimodal paradigms, in which relevant and irrelevant information are presented in the same sense. Thus, it remains unclear whether the mechanisms proposed to underpin unimodal interference control generalise to cross-modal contexts.*

Aims: *The aim of this thesis was to identify whether similar mechanisms underlie unimodal and cross-modal interference control. To answer this question, I compared patterns of unimodal and cross-modal interference in development and ageing, and compared the processing levels at which unimodal and cross-modal interference occurred. Furthermore, I explored whether some senses are given priority over others (i.e. sensory dominance), whether this differs between age groups, and whether this is associated with susceptibility to distraction from different senses.*

Experimental chapters (3 – 9): *In Chapter* ***3****, I introduce a Stroop paradigm designed to separate interference occurring at stimulus-encoding and response-selection. Using this paradigm I compared unimodal and cross-modal Stroop interference in children, young adults and older adults. In Chapter* ***4****, I investigated whether some ageing effects in Stroop performance can be attributed to colour vision loss, and set up the experimental protocol for deriving event-related potentials (ERPs) associated with different stages of processing. In Chapter* ***5****, I used ERPs in a group of young adults to assess whether unimodal and cross-modal interference occurred at different processing stages. In Chapter* ***6****, I validated whether portable electroencephalography (EEG) could be used to assess interference control in the real world.*

Following this, I focus on sensory dominance measures and how they might be related to cross-modal interference. In Chapter ***7****, I perform a meta-analysis of studies using one measure of sensory dominance, the Colavita* *effect, to assess the robustness of this measure and whether it is influenced by age. In Chapter* ***8*** *I asked whether developmental shifts in sensory dominance are also evident in multisensory illusions (i.e. the McGurk effect). Finally, in Chapter* ***9*** *I used an exploratory, correlational approach to identify whether individual differences in sensory dominance were associated with unimodal and cross-modal interference.*

Conclusions: *Findings suggested that different mechanisms underlie unimodal and cross-modal interference. Unimodal and cross-modal interference showed different patterns of development and decline in childhood and ageing. Unimodal interference control is poor in childhood and old age, whilst cross-modal interference control is poor in childhood but spared in ageing. Cross-modal interference also occurs mainly at stimulus-encoding stages, whilst unimodal interference also occurs at the level of response selection. Following this, I found measures of sensory dominance appeared robust and modulated by age. A developmental shift from auditory to visual dominance was seen in the existing literature studying the Colavita effect and my empirical investigation using the McGurk effect. However, there was no correlation between sensory dominance (measured with the Colavita task) and cross-modal distractibility in a group of young adults. Those that were more visually dominant were more susceptible to unimodal, visual, Stroop interference. I discuss the findings of this thesis with respect to the theoretical implications of findings, the gaps in literature this experimental work addresses and directions for future research.*

Acknowledgements

First and foremostly, I would like to thank the superwomen that are my supervisors, Dr Harriet Allen and Dr Lucy Cragg, without whom this thesis would not have been possible. Thank you for always being there for support and advise (even through three pregnancies Lucy!), for encouraging me to take incredible opportunities, and for inspiring me to pursue a career in academia. I could not have asked for two better role models.

Secondly, I would like to thank all the researchers I have been honoured to collaborate with. Dr Nicholas Badcock at Macquarie University, Sydney, for supporting me through an intense 6-week research project overseas. Jammy Stacey and Dr Paula Stacey, from Nottingham Trent University, for enabling me to pursue my first peer collaboration following a conference in Nashville. These experiences have taught me that the best research originates from collaboration (and from conference dinner wine).

Of course my research, and collaborations, could not have been possible without the support of my funders, the Economic and Social Research Council and the University of Nottingham, and several organisations that have funded overseas visits and conference attendance, including the Experimental Psychology Society, Guarantors of Brain and The Nottingham Graduate School.

On a personal note, I would like to thank my family for their continuous support (despite not really knowing “what I do” or “when I finish University”). Thank you to my partner, Ben Johnson, for keeping me grounded throughout the peaks and troughs of PhD life. Thank you to all the friends that have supported me along the way (you know who you are). In particular, I would like to give a special shout out to my trusted friend and office mate Sarah Maddison. Thank you for always being my voice of reason, for putting up with long rants and irrational anxieties, and for encouraging me to have faith in myself.

Copyright permissions

Chapter **3** is adapted from the paper:

> Hirst, R.J, Kicks, E., Cragg, L., & Allen, H.A. (In Press). Cross-modal interference-control is reduced in childhood but maintained in ageing : A cohort study of stimulus- and response-interference in cross-modal and unimodal Stroop tasks. *Journal of Experimental Psychology: Human Perception and Performance.*

Chapter **7** is adapted from the paper:

> Hirst, R.J, Cragg, L., & Allen, H.A. (2018). Vision dominates audition in adults but not children: A meta-analysis of the Colavita effect. *Neuroscience and Biobehavioural reviews, 94, 286-301.*

Chapter **8** is adapted from the paper:

> Hirst, R.J, Stacey, J.E., Cragg, L., Stacey, P.C., & Allen, H.A. (2018). Vision influences auditory perception in adults but not children: Adults have a lower threshold for the McGurk effect in audio-visual noise. *Scientific Reports, 8(1), 1-12.*

Table of Content

Chapter 1: Introduction

The main aim of this thesis was to identify whether similar mechanisms underpin unimodal and cross-modal (audio-visual) interference control. The first half of this thesis asked whether unimodal and cross-modal Stroop interference manifest similar patterns of development and decline across the lifespan, and whether unimodal and cross-modal control occur at similar processing stages. The second half of this thesis investigated whether some senses are given priority over others (i.e. sensory dominance), whether this differs between age groups, and whether this is associated with susceptibility to distraction from different senses.

In this Chapter, I first demonstrate the major imbalance between unimodal and cross-modal research focus, using the example of the Stroop task. I then outline the neural mechanisms proposed to underlie unimodal and cross-modal interference control and how interference changes with development and ageing. Following this, I define sensory dominance and why it is relevant to the current thesis. Finally, I provide a roadmap of the experimental chapters that follow. The literature relating to each experimental Chapter is outlined in more depth within the respective chapters.

Unimodal and cross-modal interference control

Imagine you are reading this thesis at your favourite coffee shop. Suddenly a group sit at the next table and begin chatting, loudly. Worse, they are talking about the very topic of this thesis. The ability to ignore this distracting auditory information (conversation) whilst you focus on visual information (this thesis) is known as selective attention, or, more specifically, cross-modal interference control. Despite the every-day occurrence of this ‘cross-modal’ interference, previous research has held a unimodal focus. In this section, I will outline literature investigating cross-modal interference control, with a specific focus on studies using arguably the most popular measure of interference - the Stroop task (Stroop, 1935). I will then describe the neural pathways proposed to underlie unimodal and cross-modal interference control. Finally, I will describe how cross-modal interference control may change with development and ageing.

Studies of cross-modal interference control

Interference control is a component of executive function (Miyake et al., 2000). It is the process by which we are able to ignore distractions or interfering stimuli to maintain focus on a task or stimulus (Diamond, 2013; Hasher & Zacks, 1979). The neurocognitive processes underlying interference control have undergone much investigation. However, current knowledge is based largely on unimodal tasks, in which relevant and irrelevant information are presented to the same sense. For example, the Stroop task (Stroop, 1935)[1] is the most extensively used paradigm for the study of interference control (Dyer, 1973; Jensen & Rohwer, 1966; MacLeod, 1991, 1992). In this paradigm participants must report the ink-colour of a colour-word that is either congruent or incongruent with the ink in which it is written (e.g. “RED” in red ink or “RED” in blue ink). Participants are typically slower and less accurate during incongruent verses congruent trials, suggesting interference from irrelevant information. In its original form the Stroop task is a unimodal, visual task, requiring focus on, and suppression of, visual information (i.e. colour and written word). However, unimodal, auditory, variants of the Stroop task have also been established and report analogous interference effects (Green & Barber, 1981; Gregg & Purdy, 2007; Morgan & Brandt, 1989; Shor, 1975). A critical assumption underlying most Stroop research is that Stroop interference provides an index of every-day distractibility. Yet, in its unimodal form, the Stroop task cannot assimilate real-life, multisensory, distraction.

Literature using the Stroop task provides a prime example of the imbalance that exists between unimodal and cross-modal research. A basic

---

[1] It is recognised that a number of tasks have been used to study interference control. However, given the popularity of the Stroop task, this task forms the focus of this thesis. The current thesis also does not focus on tasks of response inhibition (e.g. Go/no-go tasks or Stop Signal Response Time Tasks). Although related, these tasks are more closely linked with behavioural inhibition rather than interference control (Nigg, 2000).

search of journal articles with the word “Stroop” in the title, abstract or keywords yields 8007 hits[2]. However, a more specific search for cross-modal Stroop studies yielded 71 hits[3]. Closer inspection of these studies indicated only 23 actually used cross-modal Stroop tasks, a careful review of the references in each study, and including additional missing articles, still only yielded 32 studies using cross-modal Stroop tasks - which I define here as simultaneously presenting conflicting information with semantically related content to two senses and asking for judgement on one sense (excluding Flanker tasks and illusions such as the McGurk). However, many of these 32 studies used complex Stroop variants, such as Emotion Stroop tasks (V. I. Müller, Cieslik, Kellermann, & Eickhoff, 2013; Weijkamp & Sadakata, 2017; Zinchenko et al., 2017), or chemosensory Stroop tasks involving gustatory (Razumiejczyk et al., 2016; Razumiejczyk, Macbeth, Marmolejo-Ramos, & Noguchi, 2015; Xiao, Dupuis-Roy, Yang, Qiu, & Zhang, 2014) and olfactory (White & Prescott, 2007) senses. Following exclusion of “complex” Stroop tasks only 24 published journal articles[4] using cross-modal (audio-visual) Stroop tasks remained. Figure 1 and Table 1 illustrate specific details of these studies. This is the extent of the literature using the Stroop task to investigate interference between the two senses we experience the most day-to-day distraction from, audition and vision. Here I will outline the existing cross-modal Stroop literature, before specifying the gaps and questions that remain in this research field.

---

[2] Search conducted on web of science (all data bases) on the 25th September 2018 using the following conditions where TI = title and TS = topic:
(TI = (Stroop*) OR TS = (Stroop*)) *AND* **LANGUAGE:** (English) *AND* **DOCUMENT TYPES:** (Article)
Note that this search is not exhaustive and does not include the original Stroop (1935) studies, in which the word Stroop was not in either the title or topic.

[3] Search conducted on web of science (all data bases) on the 25th of September 2018 using the following conditions:
(TI = (Stroop* AND (Cross-modal* OR cross-modal* OR audio-visual* OR audiovisual*)) OR TS = (Stroop* AND (Cross-modal* OR cross-modal* OR audio-visual* OR audiovisual*)))
AND **LANGUAGE:** (English) *AND* **DOCUMENT TYPES:** (Article)

[4] Two relevant conference abstracts were also identified (Ashitaka & Shimada, 2010; Shimada & Ashitaka, 2010).

Author(s)	Exp.	n	Task	Relevant stimulus	Irrelevant stimulus	Uni comp.	Resp. type	Age group(s)	Cross-modal Stroop effect?
Thackray & Jones, (1971)		10	Identify colour rectangles.	Colour of rectangle	Spoken words	Y	M	A	N
		10	Identify colour of word.	Colour ink	Spoken word +written word	Y	M	A	N
		10	Identify the colour of word.	Colour ink	Spoken numbers + written word	Y	M	A	N
Cowan & Barron (1987)		32	Sheet Stroop. Identify the colour ink of a row of Xs or the colour ink of incongruent colour-words either in silence or with one of 4 types of sound.	Colour ink	**Cond 1:** Spoken colour-words **Cond 2:** Alphabet **Cond 3:** Music **Cond 4:** the word "The" repeated	Y	V	A	Y
Miles, Madden, & Jones, (1989)	1a	12	Sheet Stroop. (As in Cowan & Barron (1987) only with spoken colour-words or silence)	Colour ink	Spoken colour-words	Y	V	A	N
	1b	8	As in 1a but whisper responses.	Colour ink	Spoken colour-words	Y	V	A	N
	2	24	**Cond 1:** identify if colour and word are same or different **Cond 2:** Identify if the colour of the current stimulus is the same as the colour of the previous (ignore the word)	Colour ink	Word (spoken or written)	Y	V	A	N
Miles & Jones, (1989)	1	30	Sheet Stroop. (As in Cowan & Barron (1987) only with spoken colour-words or silence)	Colour ink	Spoken colour-words	Y	V	A	N
Cowan, (1989)	1	32	Identify the colour of 6 dots as fast as possible in silence or with auditory stimuli.	Colour of dots	**Cond 1 / 2:** Repetition of 3 words in the response set. **Cond 3:** Repetitions of	N	V	A	Y

Author(s)	Exp.	n	Task	Relevant stimulus	Irrelevant stimulus	Uni comp.	Resp. type	Age group(s)	Cross-modal Stroop effect?
					colour neutral adjectives.				
Shimada, (1990)		10	Name colours.	Colours	Japanese colour-words or buzzer	-	-	A	Y
Stuart & Carrasco, (1993)	1	16	Name EITHER the word or picture as fast as possible.	**Cond 1:** Picture **Cond 2:** Spoken word **Cond 3:** Picture **Cond 4:** Spoken word	**Cond 1:** Spoken word **Cond 2:** Picture **Cond 3:** Noise (control) **Cond 4:** X's (Control)	N	V	A	Y
	2	16	Same as Exp. 1 except category decision made (e.g. "is the picture/word clothing?" yes or no verbal response)	**Cond 1:** Picture **Cond 2:** Spoken word **Cond 3:** Picture **Cond 4:** Spoken word	**Cond 1:** Spoken word **Cond 2:** Picture **Cond 3:** Noise (control) **Cond 4:** X's (Control)	N	V	A	N
	3	31	Same as Exp. 1 and 2 but within subjects design (i.e. all subjects perform both kinds of judgement)	**Cond 1:** Picture **Cond 2:** Spoken word **Cond 3:** Picture **Cond 4:** Spoken word	**Cond 1:** Spoken word **Cond 2:** Picture **Cond 3:** Noise (control) **Cond 4:** X's (Control)	N	V	A	Y
Elliott, Cowan, & Valle-Inclan, (1998)	1	24	Name colour patch in silence, with an incongruent spoken word (0ms SOA or -500ms SOA – word first)	Colour patch	Spoken colour-word or non-colour-word	N	V	A	Y (at 0ms)
	2	24	Same as Exp.1 except that non-colour-words, colour-words and silence were separated into different blocks (rather than intermixed)	Colour patch	Spoken colour-word or non-colour-word	N	V	A	Y (at 0ms)
Elliott & Cowan, (2001)	1	28	Either presented the same spoken colour-word repeatedly or a change was introduced part way through testing (to assess whether change made the auditory distractor more distracting). In Exp 1. auditory information was	Colour patch	**Cond 1:** Spoken colour-word (repeated) **Cond 2:** Spoken colour-word – changed to another colour-word half way. **Cond 3:** Spoken colour-	N	V	A	Y (higher with change)

Author(s)	Exp.	n	Task	Relevant stimulus	Irrelevant stimulus	Uni comp.	Resp. type	Age group(s)	Cross-modal Stroop effect?
Elliott & Cowan, (2001) *(continued)*			presented at the same time as the colour-naming task.		word changed to non-colour-word half way. Same 3 conditions with tones and non-colour-words (i.e. repeated, changed to different stimulus but of same type half way, or changed to different type half way)				
	2	24	Passive pre-exposure to auditory distractors (such as colour-words) before a critical trial presenting a colour patch and auditory stimulus to be ignored.	Colour patch	Spoken colour-word, non-colour-word or tone (habituated before test)	N	V	A	Y (higher with change)
	3	72	Same as Exp. 2 except half of participants attend to pre-exposure (press button in response to some words).	Colour patch	Spoken colour-word, non-colour-word or tone (habituated before test – half of participants actively attended)	N	V	A	Y (higher with change)
Hanauer & Brooks, (2003)		15	Identical to (Elliott et al., 1998)	Colour patch	Spoken colour-word or non-colour-word	N	V	4-5 yr	Y
		15						6 -7 yr	Y
		15						9-11yr	Y
		30						A	Y
Mayer & Kosson, (2004)		25	Target appeared in one of four locations. Spoken word presented corresponding with either the congruent or incongruent location (“UP”, “DOWN”, “LEFT”, “RIGHT).	Location of visual target	Spoken word	N	M	A	Y
Hanauer & Brooks, (2005)	1	19	Name pictures of objects whilst ignoring spoken names of objects.	Picture	Spoken word	N	V	3-5yr	Y (more robust in young

Author(s)	Exp.	n	Task	Relevant stimulus	Irrelevant stimulus	Uni comp.	Resp. type	Age group(s)	Cross-modal Stroop effect?
									children)
		19						6-7yr	Y
		20						8-11yr	Y
Hanauer & Brooks, (2005) *(continued)*		30						A	Y
	2	19	Same as Exp.1 but clothing and animal pictures intermixed (i.e. not blocked).	Picture	Spoken word	N	V	3-5yr	Y
		20						6-7yr	Y
		20						8-11yr	Y
		30						A	Y
	3	20	Same as Exp. 1 and Exp. 2 except spoken word distractors could be either in the response set or not in the response set.	Picture	Spoken word	N	V	4-7yr	Y
		18						9-12yr	Y
		30						A	Y
Elliott, Barrilleaux, & Cowan, (2006)		110	Identical to Elliott et al (1998) with the inclusion of congruent trials and investigated correlations with working memory	Colour patch	Spoken colour-word or non-colour-word	N	V	A	Y
Yuval-Greenberg & Deouell, (2009)		47	Congruent or incongruent animal sounds presented with picture followed by a categorisation question (e.g. was the sound/picture a dog?)	**Cond 1:** Animal Sound **Cond 2:** Animal picture	**Cond 1:** Animal picture **Cond 2:** Animal sound	N	M	A	Y
Morey et al (2012)		116	0ms condition of Elliott et al (1998). Investigated correlation with working memory.	Colour patch	Spoken colour-word or non-colour-word	N	V	A	Y
Donohue, Todisco, & Woldorff, (2013a)		27	Respond to spoken letter X or O and ignore visual stimuli presented bilaterally that are either fully congruent (i.e. spoken X and visual X's on both side)	Spoken letter	Visual letter	N	M	A	Y

Author(s)	Exp.	n	Task	Relevant stimulus	Irrelevant stimulus	Uni comp.	Resp. type	Age group(s)	Cross-modal Stroop effect?
			fully incongruent (i.e. spoken X and visual O's on each side) or partially congruent (i.e. spoken X and visual X one side O the other).						
Donohue, Appelbaum, Park, Roberts, & Woldorff, (2013)		15	Written and spoken colour-words presented at 8 SOAs. Half the trials participant responded to the written word and the other half participants responded to the spoken word.	**Cond 1:** Written word **Cond 2:** Spoken word	**Cond 1:** Spoken word **Cond 2:** Written word	N	M	A	Y
Appelbaum, Donohue, Park, & Woldorff, (2013)		48	Same as Donohue, Appelbaum, Park, Roberts, & Woldorff, (2013) with the added factor of visual feature combination (i.e. visual information could contain written words in black, scrambled words in colour or written words in colour ink).	**Cond 1:** Written word **Cond 2:** Spoken word **Cond 3:** Colour ink of scrambled word **Cond 4:** Colour ink of colour-word	**Cond 1:** Spoken word **Cond 2:** Written word **Cond 3:** Spoken word **Cond 4:** Spoken word + written word	N	M	A	Y
Elliott et al (2013)		200	Identify the colour of the stimulus	**Cond 1:** colour patch **Cond 2:** colour ink of @ symbols **Cond 3:** colour ink of colour-word	Spoken colour-word	Y	M	A	Y
Lutfi-proctor, Elliott, & Cowan, (2014)	1	61	Name the colour of the stimulus shown or read the written (black) colour-word as quickly as possible. Ignore auditory information.	Colour of visual stimulus (Square, @ symbols or X's)	Spoken colour-words	N	M	A	Y
Mayer, Ryman, Hanlon, Dodd, & Ling, (2017)		67	Cue indicated whether participants should "LOOK" or "HEAR" and participants then responded to a number in the attended modality. Numbers could be either congruent or incongruent.	Number (either visual or auditory depending on pre-cue)	Number in non-cued modality	N	M	A	Y
Francis, MacLeod, &	1a	50	Identify the colour of a written colour-word or row of X's.	Colour of X's or colour-word	Spoken colour-words, tone, written colour-words	Y	V	A	Y

Author(s)	Exp.	n	Task	Relevant stimulus	Irrelevant stimulus	Uni comp.	Resp. type	Age group(s)	Cross-modal Stroop effect?
Taylor, (2017)					or combination of auditory (word or tone) and written word.				
	1b	50	Identical to Exp.1					A	Y
	2	48	Identical to Exp. 1 and Exp. 2 except the visual distractors were presented as flankers.			Y*	V	A	Y
Thomas, Nardini, & Mareschal, (2017)		26	Animal sound and picture followed by a spoken animal name. Press a button if the spoken animal name matched the sound.	Animal sound	Animal picture	N	M	6-7 yrs	Y
		33						8-9 yr	Y
		17						A	Y

Table 1. Summary of all identified studies using cross-modal Stroop tasks. Uni comp = was a unimodal comparison task included (Y = Yes, N = No, Y* = yes and this was not the standard colour-word Stroop – all unimodal comparison tasks were unimodal visual tasks), Resp. type = response type used (V = Vocal, M = Manual button press), Age group = Age group tested (yrs = years old, and, because not all studies provided the exact age of adult participants, A = adult), cross-modal Stroop effect = was a cross-modal Stroop effect found (Y = yes, N = no)

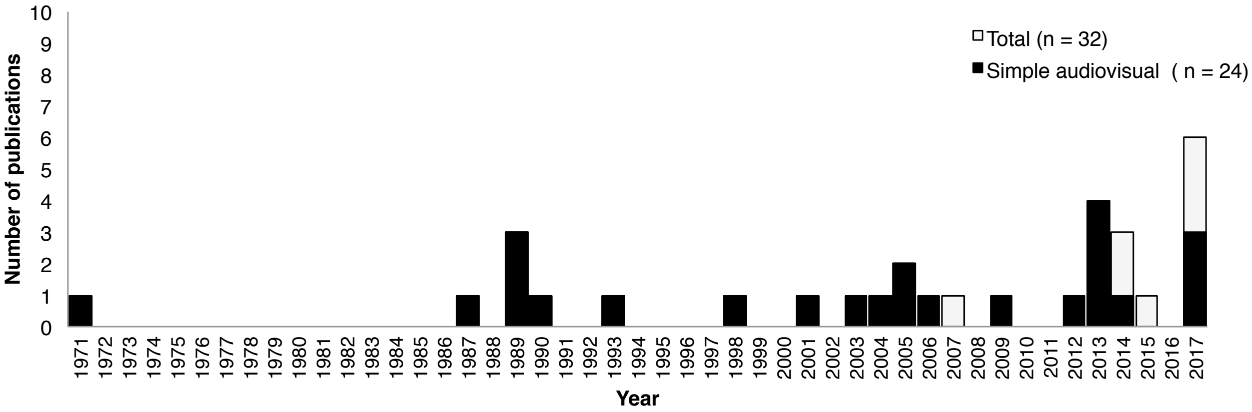


Figure 1: Number of peer-reviewed cross-modal Stroop publications by year. "Total" includes complex Stroop tasks (i.e. emotion Stroop of chemosensory Stroop) "Simple audio-visual tasks exclude these paradigms.

Thackray and Jones (1971) were the first to employ a cross-modal colour-word Stroop paradigm. In this study, one group of participants performed the standard colour-word Stroop task, however, spoken colour-words were presented in addition to the visual stimuli. Spoken words did not increase interference effects above that induced from the written word. A different group completed a cross-modal Stroop task, in which they sorted colour rectangles whilst ignoring spoken words. It was found that spoken words did not slow the time taken to sort colour rectangles. These findings suggested that Stroop interference is a unimodal phenomenon and does not occur when distracting colour-words are presented audibly. However several confounding factors may explain these null results. Firstly, prior to the Stroop task, participants sorted colour rectangles with no distraction. Thus, all subjects had experience with colour rectangle stimuli, which may have made the cross-modal task easier. Second, in the standard Stroop task colour and word are part of the same object thus the written word would likely have been more distracting than the spoken word. In addition to this, the limited sample size of this study (n = 10 per group) means statistical comparisons were likely underpowered. Given these limitations, it cannot be concluded based on these findings alone that the cross-modal Stroop effect does not exist.

Nearly 20 years later, Cowan and Barron (1987) were the first to report a cross-modal (audio-visual) Stroop effect. In this study, participants read aloud colours from a card in silence or while ignoring spoken colour-words or semantically unrelated sounds. Participants were slower and more likely to make errors when presented with colour-words than semantically unrelated sounds or silence, suggesting a cross-modal Stroop effect. This conclusion was followed by several failed replications from Miles et al (Miles & Jones, 1989; Miles et al., 1989) and criticisms of these replications from Cowan (1989a). Despite this, Cowan (1989b) did manage to replicate findings in a follow up study with new stimuli and more complex colour terms (e.g. “beige”, “turquoise”, “lavender”). Thus, discrepant findings within this early literature remain perplexing.

More recent studies have yielded replicable cross-modal Stroop effects (Donohue, Appelbaum, et al., 2013; Elliott et al., 2014, 2006, 1998; Elliott & Cowan, 2001; Francis et al., 2017; Hanauer & Brooks, 2003, 2005; Lutfi-proctor et al., 2014; Roelofs, 2005; Shimada, 1990; Stuart & Carrasco, 1993; Thomas et al., 2017), and theoretical accounts of Stroop interference have attributed cross-modal and unimodal Stroop effects to a similar cause. For example, according to the *word production architecture account (*also known as the *speed of processing account* or *the race model)*, colour naming is more demanding (and therefore slower) than word reading, requiring an additional stage to identify the associated word (Dyer, 1973; Elliott et al., 2014; Roelofs, 2005). Therefore the word, which is processed faster, interferes with colour naming. Roelofs (2005) reported this asymmetry also exists for spoken word repetition and colour naming, thus resulting in cross-modal Stroop effects. Nevertheless, several important differences seem to exist between unimodal and cross-modal effects.

First, cross-modal Stroop effects appear smaller than unimodal effects (Elliott et al., 2014; Francis et al., 2017). Second, whilst the traditional Stroop effect increases with response time latency, cross-modal interference appears equivalent across the response time distribution (Elliott et al., 2014), suggesting different stages of processing. Third, cross-modal effects produce stronger facilitation, in contrast to unimodal results, which show equal facilitation and interference (Donohue, Appelbaum, et al., 2013). Fourth, the traditional print version of the Stroop task appears correlated with working memory capacity, whilst the cross-modal Stroop task does not (Morey et al., 2012). Together, these differences suggest that unimodal and cross-modal Stroop effects manifest themselves differently, perhaps indicating different underlying processes.

Further to the differences between unimodal and cross-modal effects, numerous findings remain unexplained. For example, Appelbaum et al (2013) found that when audio-visual colour-word stimuli are presented, the spoken word did not have an additive effect on the written word (i.e. written and spoken words together do not produce stronger slowing and accuracy decrements than either distractor alone; as in Thackray & Jones, 1971). This could be due to the combined nature of colour and word in visual stimuli or this could indicate a special influence of vision (i.e. sensory dominance). As will be discussed later in this introduction, it has been suggested that adults are more strongly influenced by vision (Colavita, 1974). As such, if auditory and visual distractors are present simultaneously, interference effects may be attributed solely to the visual distractor. In line with this, Donohue et al (2013) found visual distractors had a stronger effect on audition than vice versa in young adults. Furthermore, through manipulating the stimulus-onset-asynchrony it was found that visual distractors produced longer lasting incongruency effects than auditory distractors. Together these findings suggest that adults may be more susceptible to visual distraction.

There also remain several gaps in the cross-modal Stroop literature. First, literature using cross-modal Stroop tasks with participants of different ages is severely limited (three developmental studies (Hanauer & Brooks, 2003, 2005; Thomas et al., 2017) and no studies in older adults). Second, the majority of studies have investigated the cross-modal effects of audition on visual attention but not vice versa. Third, very few studies include a comparable unimodal Stroop task to compare unimodal and cross-modal effects. Most studies have used the standard unimodal visual colour-word Stroop for comparison, in which the colour and word are part of the same object, thus resulting in stronger interference. Although this final point is not directly addressed in this thesis, I aim to use a fairer unimodal comparison task.

In sum, there exists major imbalance in the use of unimodal and cross-modal paradigms to assess interference control. A prime example of this is the Stroop task. In comparison to the widespread use of the unimodal Stroop task, a limited number of studies (n = 24) have used cross-modal Stroop tasks. Of these studies, some have presented difficulties in replication, very few have investigated cross-modal Stroop effects in specific age groups (children n = 3, older adults n = 0), few have explored the effect of audition on vision *and* vice versa and only one used a fair unimodal comparison task (Francis et al., 2017). Furthermore, several differences appear to exist between unimodal and cross-modal interference effects, which might indicate different underlying mechanisms. In this thesis I use two approaches to further existing literature and identify whether similar mechanisms underpin unimodal and cross-modal interference. First I examine how unimodal vs. cross-modal Stroop effects change with development and ageing. Second, I investigate the level of processing at which unimodal and cross-modal interference occur.

Neural mechanisms of unimodal and cross-modal interference control

A fronto-parietal network of brain regions, including the lateral prefrontal cortex (LPFC), anterior cingulate cortex (ACC), and the parietal cortex, have been proposed to underlie interference control (Blasi et al., 2006). Within this network, the ACC has been associated with detecting conflict between sensory inputs (Botvinick, Cohen, & Carter, 2004; Macdonald, 2000; Milham et al., 2001) and focusing attention towards relevant information (Weissman, Warner, & Woldorff, 2004). Following this, sensory processing must be directed towards task relevant features and away from task irrelevant features. In their *conflict monitoring hypothesis*, Botvinick et al (2004) propose that this is achieved as the ACC feeds information regarding conflict to the LPFC, which maintains goal-relevant information (Banich et al., 2000), and feeds this information to other frontal regions, such as the frontal eye fields (FEF) and posterior regions, such as the parietal cortex, which, in turn, modulate activity in early sensory cortices (discussed below; Paneri & Gregoriou, 2017). This is in line with the proposed role of the LPFC in maintaining goal-related information to bias sensory representations in working memory (Sreenivasan, Curtis, & D'Esposito, 2014). However, as the FEF are associated with occulomotor behaviour and show retinotopic (not tonotopic) organisation (Schall, 2009), it is likely that this pathway is more intrinsically linked with visual, rather than auditory and cross-modal, attention. Nevertheless, it is generally agreed that a frontal network of structures, including the LPFC and ACC acts to guide selective attention.

Some theories suggest that the prefrontal attention network is supramodal, acting to coordinate attention within and across sensory modalities (Fritz, Elhilali, David, & Shamma, 2007; Wu, Weissman, Roberts, & Woldorff, 2007). As such, the LPFC and ACC have been implicated in both unimodal (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Silton et al., 2010) and cross-modal (J. A. Johnson, Strafella, & Zatorre, 2007; Weissman et al., 2004) control. However, some researchers have argued for the presence of independent pathways, dependent on the modality of the task. Braga, Wilson, Sharp, Wise and Leech (2013) found a superior fronto-parietal network (encompassing the FEF) in visual attention tasks but an inferior fronto-temporal network supporting auditory attention (also seen in Christensen, Lockwood, Almryde, & Plante, 2011). However, this study compared unimodal visual with unimodal auditory attention and did not include a cross-modal task. More recent studies using cross-modal paradigms have shown modality dependent stratification of the LPFC. In a large fMRI study (n = 64) Mayer, Ryman, Hanlon, Dodd and Ling (2017) found that rostral LPFC was associated with directing attention towards audition whilst caudal LPFC was associated with visual attention in an audio-visual Stroop task. However, they also reported that caudal LPFC functioned in a supramodal capacity, resolving multisensory conflict when necessary. Thus the LPFC appears to hold regions associated with sensory specific and supramodal attention. Together, this suggests cross-modal attention likely involves some neural structures recruited for unimodal attention tasks, as well as independent pathways.

It is well established that top-down attention modulates the tuning of neurones in earlier sensory cortices (for review of audition see Fritz et al., 2007 and for vision see Paneri & Gregoriou, 2017). Much of this evidence comes from single cell recordings with cats, and non-human primates. In terms of visual processing, directing attention towards one region of space is associated with increased activity of neurones in early visual cortex with receptive fields in the corresponding locations (Moran & Desimone, 1985; Motter, 1993; for review see Treue, 2003). It has been suggested that this results from direct connections between FEF and visual cortices found in primates (Moore & Armstrong, 2003) and humans (Ruff et al., 2006). Very recent evidence suggests the FEF has a role in visual distractor suppression. Cosman, Lowe, Woodman and Schall (2018) found that when monkeys ignored stimuli, activity in FEF neurones associated with the stimulus was suppressed, and this preceded suppression of occipital visual evoked responses (VEPs).

In the same sense that visual attention modulates visual cortices, it has been shown that attention towards auditory information also modulates the tuning of neurones in auditory cortex in animals (Fritz et al., 2007; Hubel, Henson, Rupert, & Galambos, 1959) and EEG research supports this in humans (Hillyard et al., 1973; Woldorff & Hillyard, 1991). Evidence has also shown modulation of auditory cortex arises from prefrontal sites, for example, the medial prefrontal cortex has been linked with suppressing auditory cortex activity associated with ones own voice (a usually irrelevant auditory stimulus; N. Müller, Leske, Hartmann, Szebényi, & Weisz, 2015). Together, these findings show that the prefrontal cortex elicits top-down control over unimodal visual and auditory processing, enhancing processing of relevant information and suppressing processing of irrelevant information.

It is likely that the mechanisms underlying cross-modal attention reflect unisensory tuning on a larger, cross-cortical, level. Studies using fMRI have shown that focusing on one sensory modality whilst ignoring another is associated with increased activation in the relevant sensory cortex and down-regulation of irrelevant sensory cortices (Baier, Kleinschmidt, & Müller, 2006; J. A. Johnson & Zatorre, 2005, 2006; Kawashima, O'Sullivan, & Roland, 1995; Mayer, Franco, Canive, & Harrington, 2009). If the relevant modality is pre-cued it has also been shown that this induces preparatory reweighting of cortical activation between sensory modalities (Baier et al., 2006). However, unlike unimodal tuning, some findings suggest cross-cortical effects may not require prefrontal suppression to the same extent. Lewis, Beauchamp and DeYoe (2000) presented subjects with simultaneous auditory and visual motion signals and asked them to either respond only to the auditory or only to the visual signal. Results showed down regulation of sensory regions associated with the to-be-ignored stimulus. However, in this study, directing attention towards one modality was not associated with prefrontal activity. However, this task was a perceptual motion detection task and it is therefore possible that these effects may not generalise to more cognitive, interference control, paradigms. Nevertheless, in a combined EEG-fMRI paradigm Wang, Viswanathan, Lee, and Grafton (2016) found that focusing on vision whilst ignoring audition was associated with increased theta[5] in the fronto-parietal network, whilst cross-modal auditory attention was associated with theta increases in relevant, and decreases in irrelevant, sensory cortices (and no frontal change). Thus, different networks may underlie cross-modal visual vs. cross-modal auditory attention and some cross-modal contexts may rely less on prefrontal mechanisms.

Two findings might explain why suppression of irrelevant sensory cortices in cross-modal tasks may not require prefrontal recruitment to the same extent. Firstly, Shulman et al. (1997) refer to cortical modulation as either selective or non-selective. Selective modulation occurs when a set of neurones encoding task relevant features are modulated (i.e. if participants were instructed to "ignore auditory and focus on visual" neurones in both auditory and visual cortices are relevant to the task). Non-selective

---

[5] Theta: 4-8Hz oscillatory activity associated with cognitive control and attention.

modulation occurs when neurones are modulated irrespective of the task, these modulations are associated with tonic state of arousal. Due to non-selective modulation, if participants perform a visual task with no auditory stimulation, and no instructions regarding the auditory modality, a decrease in auditory cortex activity might still be observed. In an analysis of nine Positron Emission Tomography (PET) studies Shulman et al. (1997) observed just this, that visual tasks were associated not only with increased activity in visual regions but also a decreased activity in auditory cortex. Thus, it is possible that cross-modal control may, in part, be facilitated by non-specific suppression of task irrelevant cortices. However, Laurienti et al, (2002) note that the concept of cross-modal deactivation induced from a single sensory modality remains controversial, as this has not been observed in all studies.

Another possible explanation of reduced activity in irrelevant sensory cortices is that sensory input is modulated at peripheral, pre-cortical stages. Haxby et al, (1994) suggested that decreased activity in task irrelevant cortices might result from pre-cortical gating. Pre-cortical gating refers to the dampening of a modality entirely through limiting access of sensory information to the cortex and this process is typically attributed to the thalamus (McCormick & Bal, 1994). It is possible that this gating is partly responsible for cross-modal control. Other evidence suggests that the sensitivity of peripheral sensory organs can also be modulated by attention, reducing the likelihood of sensory information reaching the cortex (I also refer to this as “gating”). Early electrophysiological findings showed that sensitivity of the cochlea in cats could be increased by auditory attention to clicks and decreased by visual attention to a mouse or olfactory attention to a fish odour. However, in contrast to this hypothesis, Shulman et al, (1997) found that suppression of auditory cortex differed across different visual tasks, suggesting results did not reflect a general gating but task-specific inhibition. However, Shulman et al, (1997) do suggest that variability may be because gating only occurs in very demanding unimodal tasks. In line with this, more recent work has also demonstrated that the decreases in cochlear sensitivity correlates with the visual demands of the task (Delano, Elgueda, Hamame, & Robles, 2007). Thus, it appears that cross-modal control over auditory influence may, in part, be modulated by pre-cortical mechanisms.

In sum, current evidence suggests interference control is guided by a fronto-parietal network that exerts control over early sensory regions to enhance processing of relevant, and suppress processing of irrelevant, information. Evidence suggests the nodes of the fronto-parietal network such as the LPFC and ACC may be active in both unimodal and cross-modal contexts. However, the exact brain regions recruited likely depend on the modality of the relevant stimulus. Furthermore, some evidence suggests functionally specialised modules in LPFC that may be recruited for unimodal and cross-modal tasks, respectively. Evidence also suggests that unimodal and cross-modal attention are guided by enhancement of task relevant, suppression of task irrelevant, neural activity. However, some findings suggest cross-modal control may also be supported by pre-cortical mechanisms or non-selective cortical modulation. Thus, although both unimodal and cross-modal control may be associated with similar systems, it is possible that cross-modal control may also be achieved via different mechanisms.

Unimodal and cross-modal interference control in development and ageing

Our ability to suppress interference from irrelevant information is proposed to improve with development and deteriorate with age (Comalli, Wapner, & Werner, 1962). In a lifespan study of the unimodal Stroop effect, Comalli and collegues (1962) found interference was greatest in children aged 7-8 years and adults aged over 60 years. According to the *frontal lobe hypothesis* of ageing (West, 1996b) poorer inhibitory control in ageing is attributable to deficient frontal lobe function. Dempster and Vegas (1992) extend this hypothesis to development, and suggest that the development and decline of executive functions such as interference control are underpinned by change in the prefrontal cortex. In line with this, structural imaging studies show the prefrontal cortex manifests protracted development across childhood into adolescence, with development of myelination (Pfefferbaum et al., 1994; Yakovlev & Lecours, 1967), grey matter density (Pfefferbaum et al., 1994; Sowell, Delis, Stiles, & Jernigan, 2001) and functional (Solé-Padullés et al., 2016) and structural (Barnea-Goraly et al., 2005) connectivity with other brain regions (for review see Fuster, 2002; Tsujimoto, 2008). Similarly, in ageing, the prefrontal cortex appears most susceptible to age-related atrophy (Raz, 1997, 2000, however see Greenwood, 2000). Together, this suggests that poorer interference control in childhood and older age is due to the protracted development and early decline of frontal attention networks. If similar brain regions and processes underpin unimodal and cross-modal control, it would be expected that this pattern of development and decline extend to cross-modal tasks.

To my knowledge, only three studies have compared the cross-modal Stroop performance between children and adults (Hanauer & Brooks, 2003, 2005; Thomas et al., 2017). Hanauer and Brookes tested a group of 4-11 year olds and a group of 17-34 year-olds. Participants named the colour of a patch while ignoring a spoken colour-word or a spoken non colour-word. Findings showed although cross-modal Stroop effects occurred in all age groups the youngest children, aged 4-5, were slowed the most by cross-modal, auditory, distraction. Presenting the distractor prior to the target also reduced cross-modal effects in older children and adults (in line with the finding that auditory distractors have shorter lasting incongruency effects ;Donohue, Appelbaum, et al., 2013). However, younger children showed strong cross-modal effects even when the distractor was presented 500ms before the target. In a second set of studies, Hanauer and Brooks (2005) found that adults, but not children, were more slowed by cross-modal distractors if they were part of the same response set (i.e. if the distractor was an item of clothing whilst making judgements about clothes). Together, Hanauer and Brooks' (2003;2005) studies suggest that a) children are more susceptible to cross-modal interference than adults, b) that these effects take place over a wider temporal window in young children and c) that adults are more affected by distractors interfering at the response selection stages.

More recently, Thomas et al, (2017) used a picture based audio-visual Stroop task with children aged 6 and 8 years (and adults). They found that children 8 years and older were faster for auditory targets presented with congruent visual information and slower for auditory targets presented with incongruent visual information. Children aged 6 years were faster for auditory stimuli paired with congruent visual information but were, counter intuitively, not disadvantaged by incongruent visual information. Interestingly, the opposite of this effect has been reported in adults. Donohue, Appelbaum, et al (2013) and Yuval-Greenberg and Deouell (2009) both found that visual information had a stronger, negative, effect on auditory detection than vice versa in young adults. Thus, it appears that the relationship between audition and vision may change from childhood to adulthood, and this may, in turn, influence the cross-modal Stroop effects observed across development. The findings from Hanauer and Brooks (2003) imply that a similar pattern of immature interference control is seen in both unimodal and cross-modal tasks. The findings from Thomas, Nardini, and Mareschal (2017) imply subtle differences in cross-modal interference may exist between childhood and adulthood. However, as neither of these studies included a unimodal comparison task it is difficult to directly compare the developmental trajectory of unimodal and cross-modal interference.

In terms of ageing, a mounting body of evidence suggests some aspects of cross-modal interference control may remain intact (Guerreiro, Anguera, Mishra, Van Gerven, & Gazzaley, 2014; Guerreiro, Murphy, & Van Gerven, 2010, 2013). In 2010, Guerreiro et al provided an extensive meta-analytic review of studies investigating distractibility in ageing. In summary, the results of this review indicated that older adults show deteriorations in unimodal interference control, however, the ability to focus on vision whilst ignore audition may remain intact. Whether older adults are also able to focus on audition whilst ignoring vision remains unclear. In a series of studies, Guerreiro and colleagues used tasks in which participants had to remember either visual or auditory information whilst ignoring distracting information in the other modality. Some findings from these studies suggested asymmetry, whereby auditory processing was disrupted by visual distraction but not vice versa (Guerreiro & Van Gerven, 2011). Similarly, using these tasks alongside EEG, Guerreiro, Anguera, et al (2014) found intact top-down modulation of auditory brain regions whilst processing vision, but not vice versa. On the other hand, in an fMRI study Guerreiro, Eck, Moerel, Evers, and Van Gerven (2015) found age-equivalent suppression of activity in visual cortices during auditory attention. As previously discussed, evidence has suggested visual attention during cross-modal tasks requires frontal regions, whilst auditory attention may not (W. Wang et al., 2016). Thus, based on the frontal-lobe hypothesis of ageing, older adults should be impaired for ignoring audition but not vision in cross-modal tasks - the reverse of what has been observed. To my knowledge no previous study has employed a cross-modal Stroop task with older adults. Furthermore, no study has used both unimodal and cross-modal Stroop tasks within the same older adult group. As such the current thesis extends the findings covered by Guerreiro et al (2010) to the use of a cross-modal Stroop paradigm.

In sum, the literature using cross-modal Stroop paradigms is limited in childhood and, to my knowledge, non-existent in ageing. Findings suggest young children are more suceptible to cross-modal distraction. However, in the youngest children, vision may have less negative influence on audition (whilst the opposite effect is observed in adults). The literature seems to suggest older adults might be able to ignore cross-modal distraction. However whether older adults can focus on audition and ignore vision remains unknown. Comparing lifespan trajectories of unimodal and cross-modal Stroop effects also enables inference as to whether similar or different mechanisms underpin unimodal and cross-modal interference control. If similar neural networks are recruited for unimodal and cross-modal control it would be expected that a similar pattern of developmental maturation and age-related decline would be observed between the two conditions.

The role of “sensory dominance”

Our perception of the world results from the integration of multiple sensory modalities. However, whether all senses are treated equally is unknown. Answering this question may be critical to understanding individual differences in cross-modal interference. This is because it is possible that more dominant sensory modalities are allocated more attentional resources. As such, it is more likely one would be distracted by these senses. Despite the apparent logic in this theory, no research to date has explicitly investigated the role of sensory dominance in cross-modal interference. In this section, I define what I mean by sensory dominance, why it is important to the current thesis, and how I will measure sensory dominance.

Defining sensory dominance

“Sensory dominance”, in the strictest sense, refers to a hierarchy of sensory processing in which one sense is given precedence. For example, the Colavita effect (Colavita, 1974) is an experimental phenomenon interpreted to reflect visual dominance in humans. In the Colavita task participants respond to visual, auditory and audio-visual targets. It is typically found that on audio-visual trials participants respond only to the visual element of the audio-visual target. Thus, vision appears to “dominate” audition. Despite this finding, the idea of a “vision dominates all” approach remains controversial. This is because the weighting between senses likely fluctuates within and between individuals depending on context, sensory ability, and age.

Multisensory illusions demonstrate how vision and audition can each dominate perception, depending on the context. Examples in which vision can dominate include the McGurk effect (McGurk & MacDonald, 1976), in which pairing an auditory “Ba” with the mouth movement “Ga” typically results in the perception of a third sound (“Da” or “Tha”), and the ventriloquist effect (Thurlow & Jack, 1973), in which the perceived location of a sound is shifted towards the location of a simultaneously presented light. Examples in which audition can dominate include the sound-induced-flash illusion (Shams, Kamitani, & Shimojo, 2000), in which participants report *seeing* two flashes when one flash is paired with two concurrent auditory tones, and the cross-bounce illusion (Sekuler, Sekuler, & Lau, 1997), in which presenting a tone at the point in time where two disks cross paths results in the perception of "bouncing" off each other. These examples show how vision and audition can both modulate perception depending on the context. One traditional explanation of this is the *modality appropriateness hypothesis* (Welch & Warren, 1980). According to this hypothesis, the sense most appropriate to the current task will drive perception. As such, vision and audition may drive perception for spatial and temporal tasks respectively, given their respective acuity in these domains. This evidence shows that sensory dominance is likely flexible, as such, I use the term "sensory dominance" in this flexible sense rather than inferring one sense dominates all contexts.

Sensory dominance in the current thesis

The cross-modal interference literature discussed so far highlights a potential shift in sensory dominance across the lifespan. Children appear more susceptible to auditory distraction when focusing on vision (Hanauer & Brooks, 2003, 2005), yet vision does not interfere with audition in younger children (Thomas et al., 2017). In contrast, young adults show asymmetrical interference, such that vision influences audition more than vice versa (Donohue, Appelbaum, et al., 2013; Yuval-Greenberg & Deouell, 2009). Furthermore, older adults appear able to ignore audition whilst focusing on vision (Guerreiro et al., 2010; Guerreiro & Van Gerven, 2011). In parallel to this, it has been shown that children might show less visual dominance (Nava & Pavani, 2013; Wille & Ebersbach, 2016) whilst adults (Colavita, 1974) and older adults (Diaconescu, Hasher, & McIntosh, 2013) may be visually dominant. Given these parallels, I speculate that sensory dominance and cross-modal distractibility might be related. However, to date, no study has investigated this link.

Before exploring the relationship between sensory dominance and distractibility in this thesis, I assess the robustness of two paradigms for assessing sensory dominance and testing the theory that sensory dominance may shift across development. First, I use a meta-analytic approach to investigate the robustness of the Colavita effect. Second, I use the McGurk effect, in which the fusion of visual and auditory information results in a third percept. However, whilst some participants respond to incongruent McGurk stimuli with “Da” or “Tha”, some respond to the visually presented mouth movement (“Ga”) whilst others respond to the auditory sound (“Ba”). Thus, it is possible to index the weighting between auditory and visual modalities. The details of the approach I adopt will be outlined further in the relevant experimental Chapter (Chapter 8)

Thesis structure

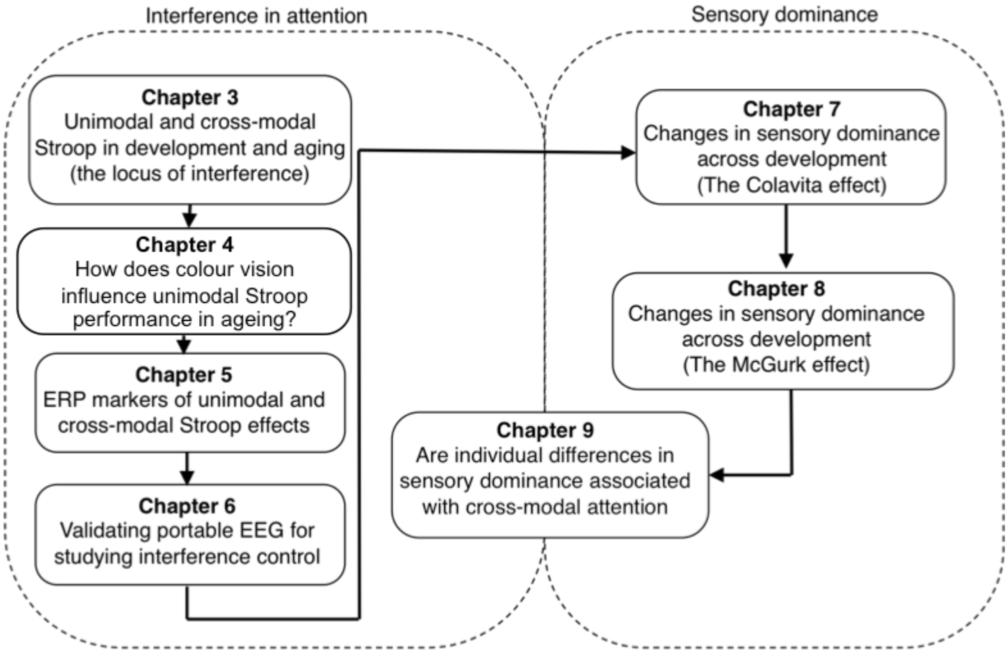


Figure 2 Roadmap of experimental chapters. Note: this does not include introduction (Chapter 1), general methods (chapter 2) and general discussion (Chapter 10).

Figure 2 provides a roadmap of the experimental chapters that follow. In this introduction I first summarised existing cross-modal Stroop literature and several caveats. The first half of this thesis will attempt to address some of these caveats. In the first experimental chapter, Chapter 3, I will expand upon the limited cross-modal Stroop literature in development and ageing by examining cross-modal Stroop effects in three age cohorts. In this Chapter I also introduce a cross-modal Stroop paradigm that advances previous approaches in two ways a) this paradigm separates interference occurring at stimulus and response levels to identify whether unimodal and cross-modal effects occur at similar processing stages and b) this paradigm has a fairer unimodal comparison task. In Chapter 4 I investigate the effect of limiting colour quality in the unimodal Stroop task. Although not cross-modal, this experiment aims to answer important questions arising from Chapter 3 (i.e. can some ageing effects in Stroop performance be attributed to changes in vision). In chapters 5 and 6, I implement EEG to more closely inspect whether similar neural processes underlie unimodal and cross-modal interference. The core difference between these chapters is that in Chapter 6 I attempt to validate portable EEG for studying Stroop effects, so that future research may use these methods to study real-world distraction.

In the second half of this thesis, I address whether a shift in sensory dominance may account for changes in cross-modal interference effects. In Chapter 7, I conduct a meta-analysis of the Colavita literature, to explore how robust sensory dominance effects are. In Chapter 8, I use the McGurk effect in children aged 4 – 12 and young adults to investigate whether a shift in sensory dominance occurs across development. Finally, in Chapter 9, I conduct an exploratory, correlational study to investigate whether sensory dominance (measured using the Colavita effect) is associated with unimodal and cross-modal Stroop performance.

In sum, by the end of this thesis I intend to address the following questions:

1. Do unimodal and cross-modal interference arise from the same or different mechanisms?
2. Does a shift in sensory dominance occur across development?
3. Is sensory dominance associated with cross-modal interference?

Chapter 2: General Methods

The equipment used to present visual and auditory stimuli differed between studies and are therefore described in the relevant experimental chapters. However, across studies, to accurately control the presentation of visual and auditory stimuli, a calibration procedure was required prior to each experiment. Furthermore, as I intended to create an inexpensive "portable" set up that could be used to collect data at public engagement events, new methodology was developed to control the luminance of monitors without the need for additional hardware costing around £3000 (i.e. a "Bits" stimulus processor; https://www.crsltd.com). This Chapter details the calibration protocol and adopted approach for controlling the luminance of images across experiments.

Calibration protocol

In the experimental chapters that follow, participants' thresholds for seeing visual stimuli and hearing auditory stimuli are measured. This requires knowing the relationship between requested stimulus intensity (luminance or volume) and actual intensity. For example, if I set the intensity of an auditory or visual stimulus to be 10dB or 100cd/m^2 respectively, I expect that the output will be 10dB and 100cd/m^2. However, as explained below, this is not necessarily the case. In this section, I will outline how the relationship between requested intensity and actual intensity was assessed, and corrected, for visual and auditory stimuli.

Monitor calibration

Ideally, the voltage input provided to a monitor should be proportional to the luminance output of the monitor. However, the actual relationship between voltage input and luminance output for most monitors is non-linear. To correct for this, a gamma-correction function, inversely related to the measured output of the non-corrected monitor, must be implemented (Pelli & Zhang, 1991). This function is illustrated in Equation 1.

Equation 1

$$y = ax^{g} + b$$

Where y is the luminance, a is a constant, x is the gun intensity requested, g is gamma and b is the minimum luminance. When setting the luminance of patches in PsychoPy (Peirce, 2007, 2009) x represents a decimal value ranging between -1 and 1 (which is equivalent to the monitor's lowest, 0, and highest luminance level, 255).

A 500x500 pixel patch was presented in the centre of the screen, via PsychoPy version 1.85.2 (Peirce, 2007, 2009). The luminance of the patch was measured at 21 evenly distributed sampling points using a PR655 spectrascan positioned at a viewing distance of 57cm[6]. The luminance of each sampling level was determined via the median of three measurements. Measurements were conducted for overall luminance, through simultaneously setting the three colour channels to each sampling level, and for each colour channel, through independently varying either the red, green or blue channel while all other channels were set to minimum luminance. Pre- and post-gamma-corrected measurements for each monitor used in experiments are represented in Figure 3. Following application of the gamma-correction function, implemented via PsychoPy, a normalised linear input-output relationship was observed (as can be seen in the right hand column of Figure 3).

---

[6] This is with the exception of Chapters 5 and 6, in which 8 sampling points were used and measurements were taken with a ColorCAL MKII Cambridge research systems Colorimeter.

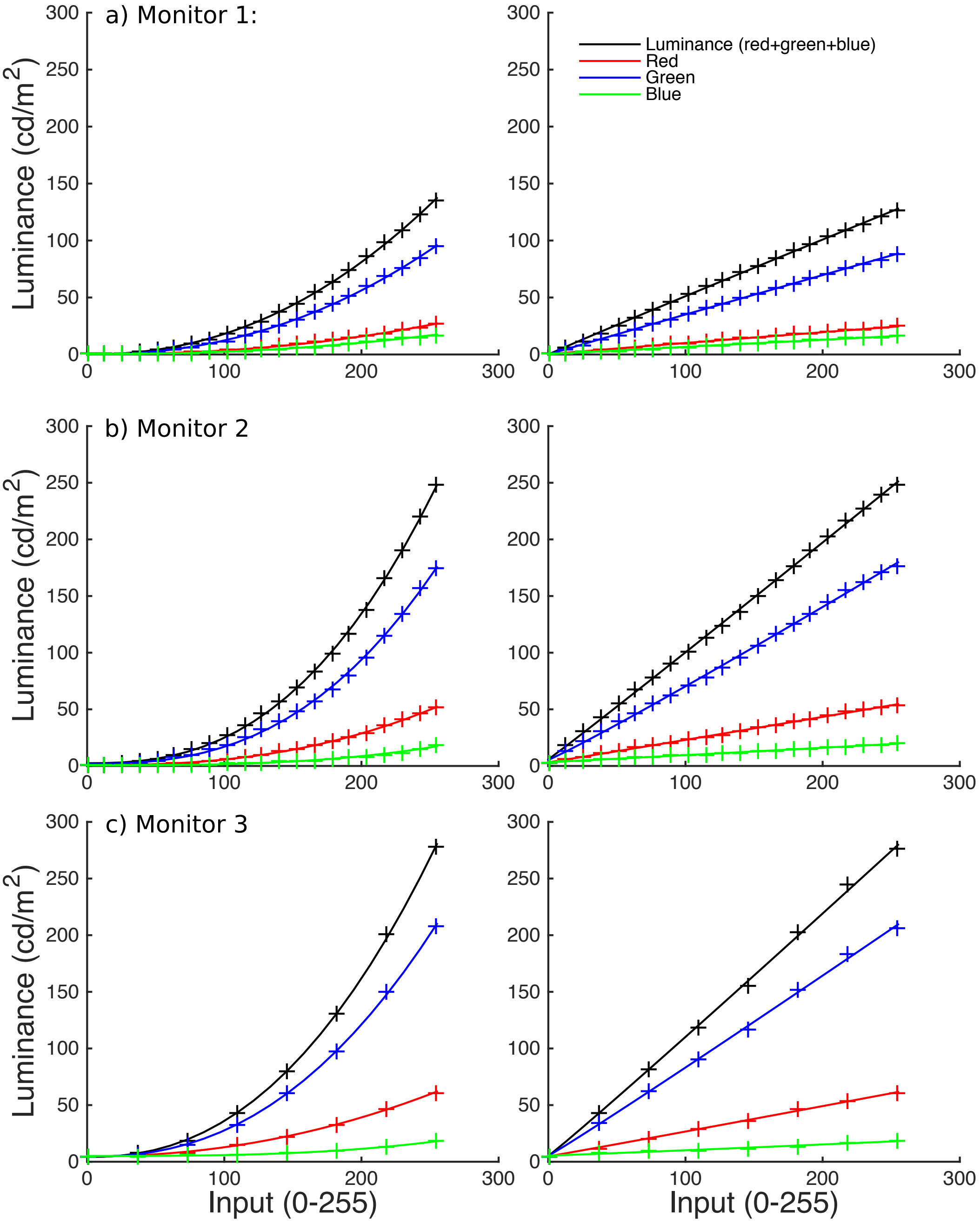


Figure 3. Pre- and post-calibration measures of luminance in each monitor used. “Monitor 1” was used in experimental Chapters 3 and 9. “Monitor 2” was used in experimental Chapter 4. “Monitor 3” was used in experimental Chapters 5 and 6. Lines indicate fitted Gamma functions pre and post correction.

Headphone calibration

Auditory stimuli were all presented via PsychoPy, which controls the volume of stimuli using a decimal value between 0 and 1. Equation 2 was used to translate the desired output in decibels (dB) to the required PsychoPy input.

Equation 2

$$c = \left(\frac{a / 0.1}{10^{\frac{b}{20}}}\right) * \left(10^{\frac{z}{20}}\right)$$

Where c is the "PsychoPy unit" required to achieve the required dB unit output, a is the PsychoPy unit level required to output 0.1vRMS (measured using Multimeter M2005 voltmeter and validated with a Tektronix TDS2014C oscilloscope), b is the dBSPL output from 0.1vRMS and z is the output required in dB SPL.

To ensure this algorithm permitted accurate control over auditory stimuli, the dB output of headphones were measured using Samuri v2.26 software and a Bruel and Kjaer 4153 (Naerum, Denmark) artificial ear, with 4134 ½" microphone (to BS EN 60318-1:2009), and Bruel and Kjaer 4157 Ear simulator, with 4134 ½" microphone (to IEC 711-1981, ANSI S3.25-1979 (R 1986)). As this equipment was only available in a location outside of the lab, an Optimus sound pressure level meter (Cirrus Research plc) was used to validate measurements in the lab and conduct further calibration (i.e. the measurements shown in Figure 4b). Initial measurements indicated that stimuli could not be accurately presented at levels below 30dB, likely due to the soundcard of the computer and ambient noise levels. To counter this problem, all auditory stimuli were presented with noise to elevate detection thresholds above 30dB. As speech stimuli can be more effectively masked through low frequency noise (Miller, 1947) Brown noise was used to mask stimuli.

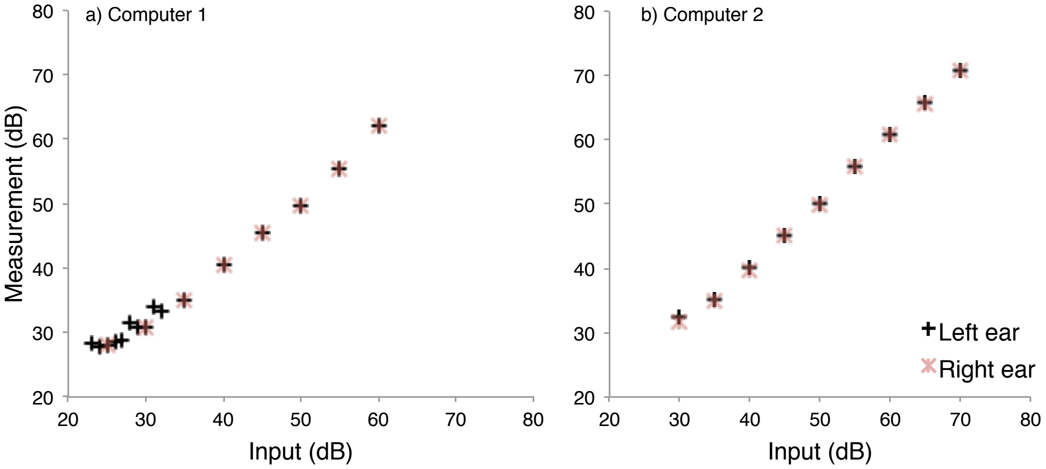


Figure 4. Measured dB output relative to requested intensity measured from the left and right ear. The first set of measurements taken from computer 1 show instable measurements of stimuli below 30dB. A set of Senheiser HMD280 PRO headphones was used for both measurements. “Computer 1” was used in Experimental chapters 3 and 9. “Computer 2” was used in experimental Chapter 4. Note that calibration in all other chapters was conducted by presenting sounds at 60dB and adjusting with a scaling factor until the measured output matched the requested volume.

To ensure all experimental stimuli were presented at the correct volume, 30-second samples of concatenated stimuli were created (spoken colour-words, “Babble” and Brown noise used in the Stroop tasks) and presented via Senheiser HMD280 PRO headphones (used in experimental chapters 3, 4 and 9). Using Samuri v2.26 software the average peak-to-peak amplitude of concatenated auditory stimuli were measured whilst the volume of the computer was set to maximum. Stimuli were then scaled to the desired volume using Praat (Boersma & Weenink, 2013). Note that calibration in chapters 5 and 6 was conducted by presenting sounds at 60dB and applying a scaling factor until the measured output matched the requested volume (i.e. the scaling factor would be -5 if the sound was consistently presented 5dB above the desired output). For this reason these values are not shown in Figure 4.

Achieving a continuous scale of luminance values using "Noisy bit"[7]

A limitation of standard 8-bit CRT monitors for deriving detection thresholds is the limited number of luminance intensities that can be presented (typically 255; Pelli & Zhang, 1991). This places limits on the accurate measurement of visual thresholds, as veridical thresholds can fall at a mid-range value between two levels. Three methods were considered to counter this limitation; brute force solutions, hardware solutions and programmable solutions (bit-stealing and noisy-bit).

Brute force solutions entail lowering the monitors' contrast - either through turning down the contrast setting or through adding background luminance. The latter can be implemented through projecting light onto the monitor or through using a half silvered mirror to combine the monitor's image with a uniform field of light (Pelli & Zhang, 1991; Savoy, 1986). However, such methods introduce difficulties with controlling the stability of background luminance. For example shining a light onto the monitor requires the luminance of the light is consistent, when, actually, the luminance of the bulb may fade over time. A second, more popular, solution is to utilise hardware solutions. This typically involves utilising a programmable attenuator, which modulates the video card input to the monitor (Pelli & Zhang, 1991). Although effective, this method involves expensive hardware that can be susceptible to damage and variability if moved. As such the monitor should be recalibrated if hardware has been moved or altered. The experiments in the current thesis involved testing at public engagement events outside of the laboratory. As such, it was viewed as impractical to employ an expensive, damageable, hardware solution.

An alternative option is to artificially create mid-range luminance values not usually permitted by the monitor; this includes the "bit stealing" (Tyler, 1997) and the "noisy bit" method (Allard & Faubert, 2008). The former of these methods relies upon the differential contribution of each colour gun

---

[7] All code relating to this method are available via the github repository https://github.com/RebeccaHirst/NoisyBit_RJHirst

towards a monitors overall luminance. For example, as can be seen in Figure 3 the blue and red guns contribute to relatively small increments in overall luminance while the green gun contributes to large increments in luminance. The bit stealing method individually varies the level of each colour gun in order to create new luminance levels. For example, we might have two levels of luminance through varying overall luminance (i.e. [100, 100, 100] and [101, 101, 101]), however, if we were to increment only the red gun (i.e. [101, 100, 100]) we would expect to produce some mid-way luminance level. Importantly this allows for mid levels of luminance without altering the overall perceivable colour of an image (Tyler, 1997).

The noisy bit method is similar to bit stealing in that it allows us to create a continuous scale of luminance from discontinuous steps. However, it does so via a different, arguably easier to implement, technique. This method produces an image with the desired mean luminance through randomly setting a proportion of the pixels in an image to the lower integer and a proportion to the higher integer. For example if we wanted an image with a mean luminance of 100.5 then we would set 50% of the pixels in the image to the lower integer (100) and 50% to the higher integer (101) (Allard & Faubert, 2008). This thesis uses the noisy bit method to increase control over luminance levels. This was selected, firstly, as the intention of this research was to create a portable set up that could be moved to new environments, such as schools and public engagement events. Thus, it appeared impractical to employ hardware solutions subject to variable output and damage if moved excessively. Secondly, this research aimed to assess contrast sensitivity for stimuli presented on different colour backgrounds. This would complicate the bit stealing method, as the colour of the background would have to be considered when implementing the desired algorithm for altering luminance. Given these issues, the noisy bit method was deemed the most practical and easily implemented method for increasing the luminance levels permitted by the monitor.

The noisy bit method was implemented through manipulating the alpha channel (responsible for opacity) of each pixel in an image (Hirst & Allen, 2018). This was a novel approach that differed from varying all three colour channels simultaneously (Allard & Faubert, 2008). To illustrate the effectiveness of this method the luminance of a 500x500 pixel grey-scale patch on a green background was measured at 81 evenly distributed, low luminance, sampling points. Each measurement reflects the median of three measurements. Notably, measurements were taken using a high luminance background in order to control for the variability of measurements produced from low luminance images. Importantly, in 5 luminance decreases as alpha (opacity) increases. These measurements illustrate a relatively continuous scale of luminance values produced via the noisy bit method.

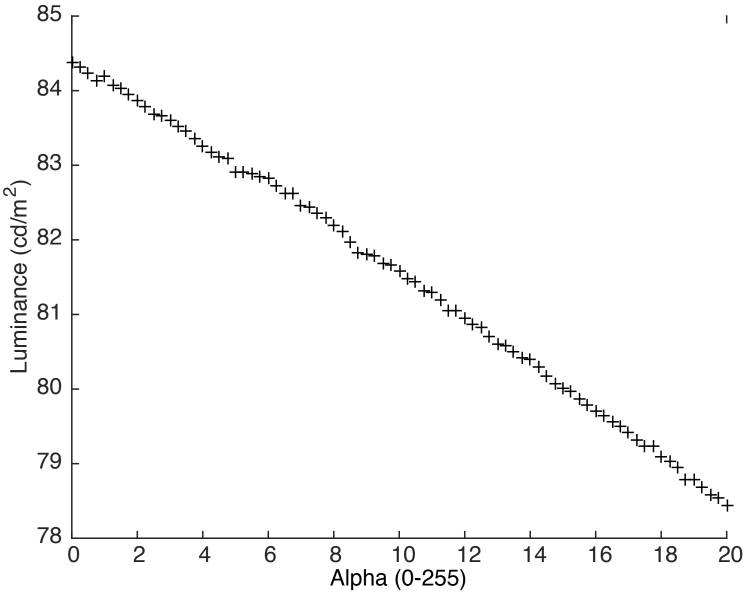


Figure 5. Luminance measurements of a grey scale patch presented on a green background at 81 evenly distributed sampled alpha channel values between 0 and 20. Each measurement reflects the median of three measurements.

Chapter 3: Unimodal and Cross-modal Interference in Development and Ageing

The main aim of this thesis is to identify whether unimodal and cross-modal interference arise from similar mechanisms. To address this, in this Chapter I compared the lifespan trajectories of unimodal and cross-modal interference, and compared the level of processing at which interference occurred.

In Experiment 1, 42 children (6-11 years), 31 younger adults (18-25 years) and 32 older adults (60-84 years) identified colour rectangles with either written (unimodal) or spoken (cross-modal) distractor-words. Stimuli could be congruent, incongruent but mapped to the same response (stimulus-incongruent), or incongruent and mapped to different responses (response-incongruent), thus separating interference occurring at early (sensory) and late (response) processing levels. Unimodal interference was worst in childhood and old age; however, older adults maintained the ability to ignore cross-modal distraction. Unimodal but not cross-modal response-interference also reduced accuracy. In Experiment 2 the effect of audition on vision and vice versa were compared in 52 children (6-11 years), 30 young adults (22-33 years) and 30 older adults (60-84 years). As in Experiment 1, older adults maintained the ability to ignore cross-modal distraction arising from either modality, and neither type of cross-modal distraction limited accuracy in adults. However cross-modal distraction still reduced accuracy in children, and children were more slowed by stimulus-interference compared with adults. It is concluded that; unimodal and cross-modal interference follow different lifespan trajectories and differences in stimulus- and response-interference may increase cross-modal distractibility in childhood.

Introduction

The introduction to this thesis outlined the current state-of-play in the cross-modal Stroop literature. In short, this literature is limited in comparison to unimodal Stroop tasks. In particular, only three studies to date have investigated cross-modal Stroop tasks in development (Hanauer & Brooks, 2003, 2005; Thomas et al., 2017) and none, to my knowledge, have assessed cross-modal Stroop interference in older adults. Given that the existing literature in development and ageing has already been outlined, I will first briefly recap why it cannot be assumed that unimodal and cross-modal interference follow the same lifespan trajectories. I will then follow this with discussion of how sensory processing and stage of processing may influence interference in development and ageing. Finally, the aims of the current study will be specified.

Interference in Development and Ageing

Our ability to suppress interference from irrelevant information improves with development and deteriorates with age (Comalli et al., 1962). However, it cannot be assumed that cross-modal interference also follows this trajectory. First, it is likely different factors contribute towards unimodal and cross-modal control. Whilst unimodal Stroop tasks entail inhibition of written words, cross-modal tasks entail inhibition of speech processing. It is likely that these facets of cognition are differentially affected by development and ageing. Second, data suggest that although unimodal and cross-modal control use similar frontal brain regions (Weissman et al., 2004), cross-modal interference control may be achieved via different, pre-cortical mechanisms (Haxby et al., 1994). Third, it has been shown that poorer attentional resources in early childhood can paradoxically reduce distractibility to multi-sensory stimuli (Matusz et al., 2015). It is possible that similar limited resources may have paradoxical effect in ageing. However, this has yet to be investigated.

The review of age-related distractibility from Guerreiro, Murphy, and Van Gerven (2010) suggested that although older adults typically show enhanced interference in unimodal tasks, cross-modal interference appears equivalent across older and younger adults, particularly if irrelevant information is auditory (Guerreiro, Adam, & Van Gerven, 2014; Guerreiro, Anguera, et al., 2014; Guerreiro et al., 2013). However, it is not clear whether this extends to all experimental paradigms. Particularly, older adults may still be susceptible to cross-modal Stroop effects. Laurienti, Burdette, Maldjian, and Wallace (2006) found that older adults responded faster to coloured circles when presented alongside a congruent spoken colour-word. This benefit was greater than presenting circles with a written colour-word and greater in older vs. younger adults. Importantly, this study only assessed facilitation effects and did not investigate cross-modal interference (thus, this study is not included as a cross-modal Stroop paradigm). In contrast to findings showing older adults effectively “filter out” auditory information whilst focusing on vision (Bell & Buchner, 2007; Belleville, Rouleau, Linden, & Collette, 2003; Proctor, Pick, Vu, & Anderson, 2005; Murphy, McDowd, & Wilcox, 1999) this suggests stronger, facilitatory, cross-modal Stroop effects in older adults. Thus, it cannot be assumed based on the review of Guerreiro et al (2010) that an ability to ignore cross-modal distraction in ageing extends to cross-modal Stroop paradigms.

Sensory Processing in Development and Ageing

When considering changes in unimodal compared with cross-modal Stroop performance for participants of different ages it is important to also consider the effects of development and ageing on the auditory and visual sensory systems and how information from different senses is integrated. Stroop interference has been attributed to an asymmetry in the ease of access to word and colour information, whereby colour naming is more difficult than word reading, as it requires intermediate processes to retrieve the word to be spoken (Melara & Algom, 2003; Roelofs, 2005). Many factors change the accessibility of colour and word information in ageing. Deteriorations in colour vision may limit the accessibility of colour, thus increasing Stroop interference in older adults (this is explored in Chapter 4; Anstey, Dain, Andrews, & Drobny, 2002; Ben-David & Schneider, 2009, 2010; Cooper, Ward, Gowland, & McIntosh, 1991). Alternatively, age-related hearing loss might make auditory distractors in cross-modal Stroop tasks less salient. Thus, to conclude that older adults have maintained ability to ignore cross-modal distractions, it is essential to ensure distractors are presented well above perceptual thresholds for all age groups being tested. Some studies using simple stimuli have attempted to control for the intensity of auditory distractors by presenting irrelevant sounds based on participants' thresholds (Belleville et al., 2003). Studies using spatial cueing tasks have also matched response times to auditory and visual stimuli across age groups (Guerreiro, Adam, & Van Gerven, 2012) and studies assessing cross-modal interference in memory for faces and voices have adjusted stimuli to a "comfortable" level to control for individual differences (Guerreiro, Anguera, et al., 2014; Guerreiro et al., 2015). However, to my knowledge, no study has yet attempted to equate visual and auditory stimulus intensity based on individual sensory thresholds, which may provide a more precise approach to equating stimuli.

It is also important to consider the role that age-related changes in multisensory integration may have on cross-modal interference. The neural processes underlying multisensory integration are thought to develop late in humans (Burr & Gori, 2012; Ernst, 2008) and are susceptible to plasticity depending on early sensory experience (Carriere et al., 2007; Polly et al., 2008). In line with the protracted development of the visual cortex (Graven & Browne, 2008b) relative to the auditory cortex (Graven & Browne, 2008a), children under 10 years of age appear less susceptible to multisensory illusions in which vision alters auditory perception (Tremblay et al., 2007) and more susceptible to illusions in which audition changes vision (Innes-Brown et al., 2011). However, these effects are modulated by early sensory experience (Narinesingh, Goltz, Raashid, & Wong, 2015; Schorr, Fox, van Wassenhove, & Knudsen, 2005). Given this, children might be more susceptible to interference from audition when focusing on vision than vice versa and this may be influenced by experience. Furthermore, children under 11 years of age show lower audio-visual facilitation of response times (Barutchu et al., 2010; Barutchu, Crewther, & Crewther, 2009) and the time window over which auditory and visual information are integrated narrows between the ages of 4 and 6 and continues to narrow into adulthood (Lewkowicz & Flom, 2014; Noel, Niear, Burg, & Wallace, 2016). As such, general developmental differences would be expected in response times to audio-visual stimuli and the extent to which auditory and visual stimuli are attributed to the same object.

Findings regarding the effect of ageing on multisensory integration have been mixed (Brooks, Chan, Anderson, & Mckendrick, 2018). Some findings suggest enhanced multisensory integration in ageing: Older adults appear more susceptible to multisensory illusions in which vision modulates audition (Sekiyama, Soshi, & Sakamoto, 2014) and vice versa (Noel et al., 2016; Parker & Robinson, 2018), and manifest greater multisensory enhancement of response times (Laurienti et al., 2006). Older adults have been shown to integrate information over wider (Bedard & Barnett-Cohen, 2016; Chan, Pianta, & Mckendrick, 2014; Noel et al., 2016) and similar (Bedard & Barnett-Cohen, 2016) time windows compared with younger adults depending on the task. A recent review of multisensory processing in ageing highlighted the need to consider unisensory change when assessing multisensory integration in ageing (Brooks et al., 2018). As such, within the current study I controlled for differences in sensory ability whilst measuring cross-modal effects.

The Locus of Interference

Irrelevant information can interfere at all stages of the information-processing stream. This includes early stages of processing at the level of encoding; stimulus-interference, and later processing at the level of response selection; response-interference (Chen, Bailey, Tiernan, & West, 2011; Cragg, 2016; De Houwer, 2003; Jongen & Jonkman, 2008; Killikelly & Szűcs, 2010, 2013; Zhang & Kornblum, 1998). In traditional interference tasks, however, stimulus- and response-interference are confounded. Incongruent conditions require participants to encode two conflicting perceptual representations and select from two competing responses, whilst congruent conditions prime complementary perceptual representations and require the same response.

To separate these processes, De Houwer (2003) presented participants with a Stroop task in which two colours were mapped to a left button and two to a right button. Thus, the colour-word and ink-colour could be congruent, incongruent but mapped to the same response (stimulus-incongruent), or incongruent and mapped to different responses (response-incongruent). Using the De Houwer (2003) paradigm, it is possible to separate three types of interference. General interference, encompassing both stimulus- and response-interference, which can be calculated as:

Equation 3

$$General\ Interference = Response\ incongruent/Congruent$$

Where “Response-incongruent” reflects response time, or accuracy, under response-incongruent conditions and “Congruent” reflects response time, or accuracy under congruent conditions.

Following this, stimulus- and response-interference can be isolated as:

Equation 4

$$Stimulus\text{-}Interference = Stimulus\ Incongruent\ /\ Congruent$$

Equation 5

$$Response\text{-}Interference = Response\ Incongruent\ /\ Stimulus\ Incongruent$$

Thus, response-interference reflects *additional* interference occurring due to the response demands of the task (i.e. over and above interference arising from stimulus level competition). General interference thus reflects the sum of stimulus- and response-interference.

Stimulus- and response-interference are candidate measures to tease apart the mechanisms underlying unimodal and cross-modal interference. For instance, Chen et al. (2011) used a variant of the De Houwer (2003) paradigm in which participants were shown six colour-words in the same or different coloured ink. Three colours were mapped to one button whilst three were mapped to the other (thus producing congruent, stimulus-incongruent and response-incongruent conditions). To map when different types of interference occurred, stimulus- and response-interference were plotted as a function of response time. Response-interference was found to occur at longer response latencies whilst stimulus-interference remained relatively constant across the response time distribution (Chen et al., 2011). These findings parallel the differences between unimodal and cross-modal interference reported by Elliott et al (2014) and suggest unimodal and cross-modal interference may arise from different types of interference (stimulus- compared with response-interference).

The balance of stimulus- and response-interference has also been shown to shift in mid-childhood, between the ages of 7 and 10 years (Cragg, 2016), and shift from young- to mid-adulthood between the ages of 30 and 45 years (Killikelly & Szűcs, 2013). Cragg (2016) found that 7-year-old children showed greater stimulus-interference while 10-year-olds and adults showed greater response-interference. Killikelly and Szűcs (2013) report increased stimulus-interference in adults aged >40 years. If unimodal and cross-modal interference share associated mechanisms this shift would be expected to also occur under cross-modal conditions.

The Current Study

Comparing developmental trajectories in childhood and old age and examining the point in the information-processing stream at which interference occurs can help to elucidate whether similar mechanisms underpin unimodal and cross-modal interference-control. However, to my knowledge no previous study has attempted to separate stimulus- and response-interference within a cross-modal paradigm and explore these processes in both development and ageing. This study used two Experiments which implemented adapted versions of the colour-word Stroop paradigm to investigate a) whether unimodal and cross-modal interference follow similar patterns of development and deterioration b) if stimulus- and response-interference contribute to cross-modal, as well as unimodal, interference and c) whether the relative contribution of stimulus- and response-interference under unimodal and cross-modal conditions changes across the lifespan. Experiment 1 compared the ability to focus on vision and ignore either visual (unimodal) or auditory (cross-modal) information. Experiment 2 compared whether a similar pattern of effects occur when focusing on audition and ignoring vision in order to ensure that the pattern of findings generalised across cross-modal conditions. Both experiments focus on age groups in which multisensory and interference control processes are known to be immature; below 11 years (Noel et al., 2016), and susceptible to age-related decline; above 64 years (Comalli et al., 1962; Noel et al., 2016). Importantly, both experiments also controlled for sensory differences that may differentially affect the balance of relevant and irrelevant information between age groups. This was achieved by ensuring distractors were presented at equivalent levels above perceptual thresholds for all participants. The findings from these Experiments will help to understand whether similar mechanisms underpin unimodal and cross-modal interference-control and whether such mechanisms are equally susceptible to developmental maturation and ageing.

Experiment 1

Experiment 1 compared unimodal and cross-modal interference in children, young adults and older adults using a modified version of the colour-word Stroop paradigm. If similar processes underlie unimodal and cross-modal interference-control we would expect unimodal and cross-modal interference to be higher in childhood and old age relative to young adulthood. Furthermore, we would predict stimulus- and response-interference to be evident under unimodal and cross-modal conditions, and manifest similar patterns across the response time distribution (Chen et al., 2011; Killikelly & Szűcs, 2010).

Method

Participants

Appropriate sample sizes for adult samples were estimated a priori via a power analysis in G*power 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007). Given the exploratory nature of the study a conservative, medium Cohens *d* effect size of 0.5 was assumed based upon well documented unimodal Stroop effects in young adults, and greater unimodal Stroop effects in children and older adults (Comalli et al., 1962; MacLeod, 1991) whilst considering the limited cross-modal Stroop effects literature. The sample size required to detect a within * between interaction with a Cohens *d* of 0.5

with 3 groups and 4 measurements was calculated. Thus the sample size was large enough to detect a difference between unimodal and cross-modal stimulus and response-interference between age groups (i.e. a 2 (sensory condition: unimodal vs. cross-modal) x 2 (interference type: stimulus vs. response-interference) x 3 (age group) mixed analysis of variance ANOVA). The exact parameters used within the power analysis were therefore; a "Repeated measures within-between interaction", alpha = .05, power= .95, number of groups = 3, number of measures = 4, Nonsphericity correction = 1. This analysis indicated a need for a minimum of 29 participants per age group (87 in total), a criterion that was met by all three samples. Sample size for children was opportunistic, data were gathered at a public engagement event and all children attending the event had the opportunity to participate.

Thirty-three young adults (mean age = 22.4 years, range = 18 - 25, 23 female), 39 older adults (mean age 71.3 years, range 61 -85, 23 female) and 49 children (mean age 9.03 years, range 6 -11, 21 female) took part. Young adults were staff and students at the University of Nottingham that were known to the researchers or recruited via the university's research participation scheme. Older adults were healthy participants recruited via the university's volunteer register and children were recruited via Summer Scientist Week, a public engagement event at the university (www.summerscientist.org). In exchange for participation young adults were offered credit as part of their degree, older adults were paid £7 and children received "tokens" to be spent on games at the event.

Two older adults were excluded from later analysis, one due to hearing aid use and one due to red-green colour blindness. Four children were excluded from later analysis due to parents reporting diagnosed developmental disorders. Following these exclusion criteria and the exclusion of outliers (see analysis section) this left a final sample of 42 children, 31 young adults and 32 older adults for analysis.

Equipment

Visual stimuli were presented via a Mac mini 3.1 on a 16" KFC Smile CA6748SL cathode ray tube (CRT) monitor (resolution 1024x768@85Hz).

Stimuli were always presented at a viewing distance of ~57cm maintained via a chinrest. Auditory stimuli were presented via Senheiser HMD280 PRO headphones. For details regarding calibration of the monitor and headphones please see Chapter 2 (page 27).

Stimuli

Notably, in the traditional Stroop task, relevant (colour) and irrelevant (word) dimensions are part of the same object (a colour-word written in coloured ink). However, under cross-modal conditions the visibly seen colour is not an attribute of the written word, rather the spoken word is separate from the visually presented colour. This may result in reduced Stroop interference, as colour and word information are not processed as part of the same object and Stroop interference has been shown to be larger when colour and word information are integrated (Macleod, 1998). To alleviate this imbalance between unimodal and cross-modal conditions the current research used a colour patch Stroop to separate colour and word information. Visual stimuli consisted of 4 coloured rectangles (initial luminance without overlaid word stimuli: red=24.47cd/m^2, green=87.45cd/m^2, blue=16cd/m^2, yellow=109.3cd/m^2; RGB colour space: red=(255, 0, 0), green=(0, 255, 0), blue= (0, 0, 255), yellow=(255, 255, 0)). Rectangles were presented with one of five written or spoken colour-words “RED”, “GREEN”, “BLUE” “YELLOW” or “BROWN”. Written colour-words were presented in black font.

To prevent participants looking at “blank” coloured areas of stimuli (thereby allowing the task to be performed without reading the words), the dimensions of the rectangle varied depending on the word it was presented with (“RED”= 0.40x0.15°, “YELLOW”=0.65x0.15°, “BLUE”=0.45x0.15°, “GREEN”=0.55x0.15°, “BROWN”=0.60x0.15°). When deriving participants’ thresholds prior to Stroop performance the dimensions of the rectangle were always set to 0.65 x 0.15° to avoid associative learning between rectangle size and word to be identified.

Spoken colour-words were spoken in a female voice with an average duration of 478.2ms (“RED” = 441ms, “GREEN” = 501ms, “BLUE” = 409ms, “YELLOW” = 485ms, “BROWN” = 555ms). By presenting auditory stimuli binaurally whilst participants fixated upon stimuli it is assumed visual and auditory stimuli were co-localised to the same location (Stern, Brown, & Wang, 2006). To control for extraneous noise (and to ensure that thresholds were pushed higher than 30dB – see Headphone Calibration section, page 27), Brown noise was presented alongside all auditory stimuli. Brown noise was created via Audacity (version 2.0.6.0) and set to 60dB throughout threshold and Stroop tasks.

Prior to performance of the Stroop task, thresholds (contrast/volume required for participants to identify visual/spoken words on 79% of trials) were measured. Written words and spoken words were then presented 10x (20dB) above threshold. For visual stimuli, if this value fell above 100% opacity then stimuli were presented at 100% opacity. For auditory stimuli, if this value was higher than 65dB then stimuli were presented at 65dB. Auditory stimuli had to be presented at maximum for 5 young adults, 17 older adults (4 of which were later removed as outliers and one of which was removed due to difficulty deriving an auditory threshold) and 11 children (2 of which were later removed as outliers). Notably, auditory stimuli were still set well above threshold, and were therefore audible, for all of these participants (M = 13.67dB above auditory threshold; range = 5.5-19.5dB above auditory threshold). No participants required visual stimuli to be set to maximum.

Balancing sensory information

The mere presence of a stimulus in one sensory modality can affect participants' detection of stimuli in another (e.g. Frassinetti, Bolognini, & Làdavas, 2002), response times to stimuli (Gondan, Niederhaus, Rösler, & Röder, 2005) and influence response times differently across age groups (Diederich, Colonius, & Schomburg, 2008; Laurienti et al., 2006). Thus, sensory input was balanced across unimodal conditions and cross-modal conditions by presenting auditory "babble" (multiple speakers saying different words at once) during unimodal conditions and visual babble (multiple words overlaying one another) during cross-modal conditions. Thus, auditory and visual information was presented in both unimodal and cross-modal conditions.

Auditory babble consisted of 96 unique samples of 3-speaker babble, created using three words from different speakers and randomly jittering word onset and offset. Words were non colour-word nouns, speakers were selected from 12 speakers (6 female) and babble duration was matched to the average duration of spoken word stimuli.

Visual babble consisted of 96 unique samples of 3-word babble, created from 3 overlaid words with randomly jittered onset/offsets. The same words were used to create visual and auditory babble. Visual babble varied in length to approximately match the length of written colour-words (19 samples for the length of rectangles used with the words “Red”, “Green”, “Blue” and “Yellow” and 20 samples for the length of the word “Brown”). When deriving thresholds, visual and auditory babble were restricted to the same length/duration to prevent associative learning between auditory/visual babble-length and colour-word to be identified.

In the Stroop task babble stimuli were set to appear 10x (20dB) above each participant’s visual or auditory threshold (with a maximum of 100% opacity/65dB) as above.

Procedure

Threshold task

To control for the detectability of distracting information in the Stroop task, thresholds for reading colour-words, on each colour background, and hearing spoken colour-words were measured. Due to time constraints for testing with children, adults and children completed two different procedures. Adults took part in a staircase procedure, whilst a method-of-adjustment was used with children. These are outlined below.

Staircase procedure

The protocol implemented with adult participants is illustrated in Figure 6. Participants identified the colour-word they saw/heard under visual and auditory conditions respectively by pressing one of four buttons (on a QWERTY keyboard). Two colour-words were mapped to two buttons on the right hand side (press the K key for “RED” and the L key for “GREEN”) and two to the left hand side (press the A key for ”BLUE” and the S key for “YELLOW”). The side to which colour-words were mapped was counterbalanced across participants and remained the same across threshold and Stroop tasks in each participant. A 1-up-3-down staircase converged upon the level at which participants correctly identified words on 79% of trials. Thresholds for each staircase were taken as the average of the final 6 out of 8 reversals.

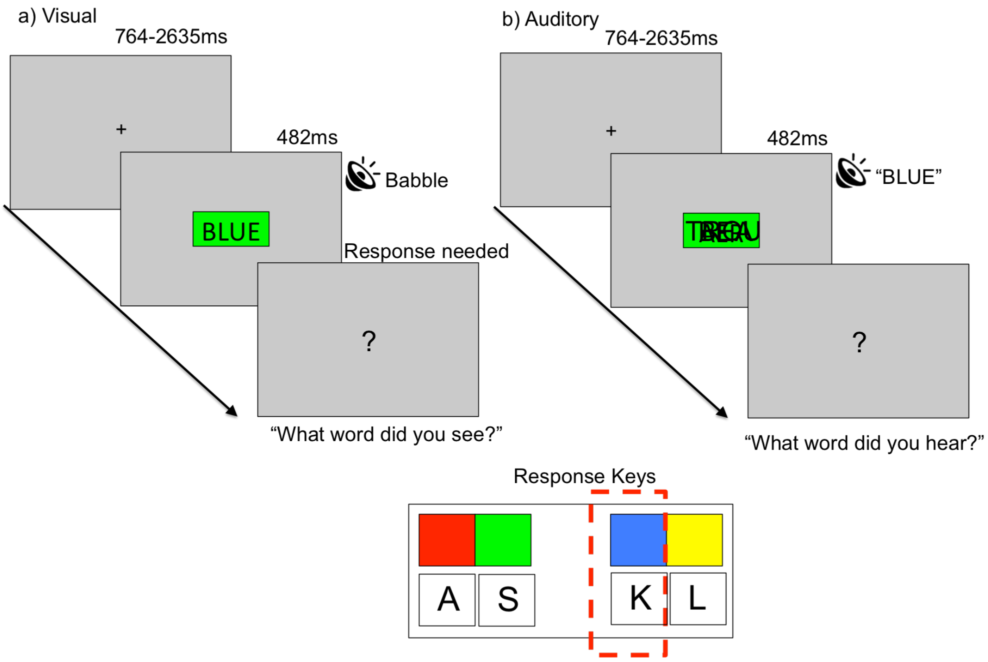


Figure 6. Protocol for isolating thresholds with adults for reading (a) and hearing (b) written and spoken words presented with different colour rectangles. Under visual conditions participants saw a written word on a coloured rectangle and heard 60dB auditory babble. Under auditory conditions participants saw a coloured rectangle containing written babble and heard a spoken colour-word. Participants were asked to wait until the question mark to identify what word they saw/heard using four possible response options. The intensity (contrast or volume) of the word decreased or increased depending on responses. Participants completed 12 visual and 12 auditory staircases (three for words/spoken words presented with each colour rectangle; red, green, blue and yellow).

To derive thresholds for *reading* words, participants were presented with a coloured rectangle overlaid by a randomly selected written colour-word

(“RED”, “GREEN”,”BLUE”, or “YELLOW”). To ensure conditions were similar to the Stroop task, stimuli were presented for 482ms alongside a randomly selected sample of auditory babble and Brown noise (both set to 60dB). To reduce reverse Stroop effects (i.e. incorrectly responding to the colour of the rectangle rather than the written word), the colour of the rectangle presented remained constant throughout each block. Furthermore, participants were instructed to withhold their response until a question mark was presented following stimulus offset. The opacity of the word adjusted in a 1-up-3-down staircase (step size (% opacity) = [11.76, 7.84, 3.92, 3.92, 1.96, 0.39, 0.10, 0.04], start value = 39.21% opacity) until 8 reversals had been reached. The threshold was defined as the average of the final 6 reversals. Participants completed 12 visual staircases each (three per coloured rectangle) in a random order. Final thresholds for reading words on each colour were taken as the average of the three derived thresholds.

The protocol for deriving *hearing* thresholds was identical to that used for reading words, except spoken colour-words (embedded in 60dB Brown noise) were presented instead of written words, and coloured rectangles contained randomly selected visual babble instead of a written word (instead of auditory babble). Participants identified the spoken word using the same 4 response keys within a 1-up-3-down procedure (step sizes (dB) = [20, 15, 10, 5, 3, 2, 1, 1], auditory start value = 60dB). It was expected that the colour of the rectangle would influence the ability to read overlaying words (i.e. due to differences in contrast). However, the effect of visually presented colour/luminance on auditory thresholds had not yet been documented. To exclude this as a moderating factor, adults in Experiment 1 also completed 12 auditory staircases (three per colour rectangle), thus enabling me to identify whether the colour/luminance of the rectangle influenced auditory threshold.

Method-of-adjustment

The method-of-adjustment (MOA) protocol is illustrated in Figure 7. This approach was used with children in Experiment 1 and all participants in Experiment 2. Participants completed 4 visual threshold assessments (1 per colour rectangle). Given that data from adult participants in Experiment 1 revealed no effect of visually seen colour on auditory threshold (see results) 1 auditory assessment was made with a randomly selected colour rectangle.

To identify threshold for *reading*, a written word on a coloured rectangle was presented alongside a randomly selected sample of auditory babble and Brown noise (both set to 60dB). Within each block the colour of the rectangle remained the same. The starting contrast of the written word varied depending on its coloured background; yellow = 3.92%, green = 3.92%, red = 5.88%, blue = 7.84% opacity. These starting values were judged based on adult data (see results) showing written words were easier to identify in higher contrast conditions (i.e. written words on yellow and green) compared with lower contrast conditions (i.e. written words on blue and red). The experimenter then decreased the contrast of the written word (using 0.04% opacity steps). Each time the experimenter decreased the contrast of the word a new, randomly selected colour-word and sample of auditory babble was presented. Participants then told the experimenter when they could “only just read the word”. Five test stimuli were then presented on the coloured rectangle that had just been presented at the reported threshold and participants were asked to read the word presented in each test stimulus. A minimum of 4/5 (80%) correct answers was required to move

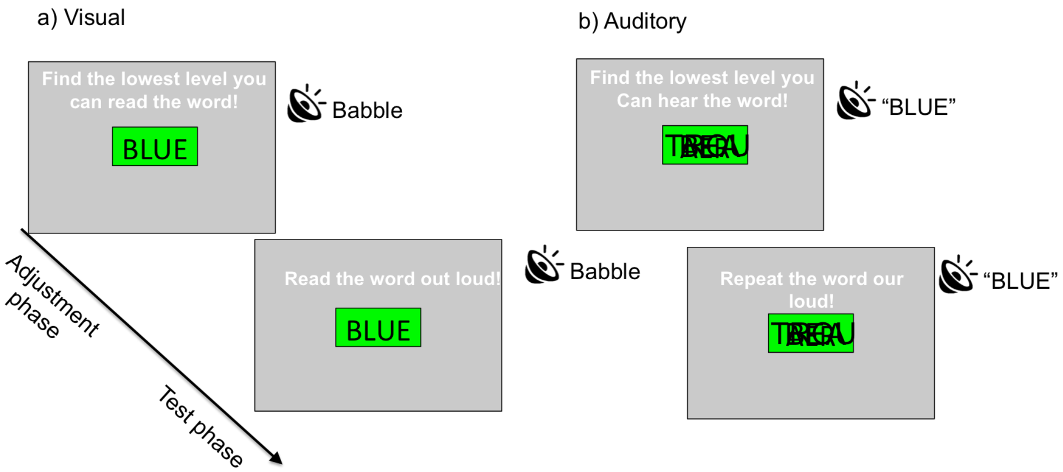


Figure 7. Method of adjustment protocol for reading (a) and hearing (b) colour-words. In the “adjustment phase” the experimenter adjusted the contrast or volume of the word until participants reported they could “only just read/hear the word”. In the “test phase” participants completed 5 test trials in which the word was presented at threshold. A minimum of 80% correct was required otherwise the protocol was repeated to derive a new threshold.

onto the next coloured rectangle, otherwise thresholds for words on that colour were re-measured.

The protocol for deriving threshold for *hearing* words was identical to the method-of-adjustment used to derive threshold for reading, except spoken words were presented instead of written words and visual babble was presented over the rectangle (instead of auditory babble). The spoken word was presented with a starting value of 65dB and the experimenter then decreased the volume of the spoken word (using 1dB steps).

Stroop task

The Stroop task used to measure unimodal and cross-modal interference is shown in Figure 8. Participants were instructed to sort coloured "tickets" (rectangles) into two boxes using two buttons (the "A" and "L" keys on a QWERTY keyboard), one for red/green rectangles and the other for blue/yellow rectangles (response mappings were counterbalanced across participants).

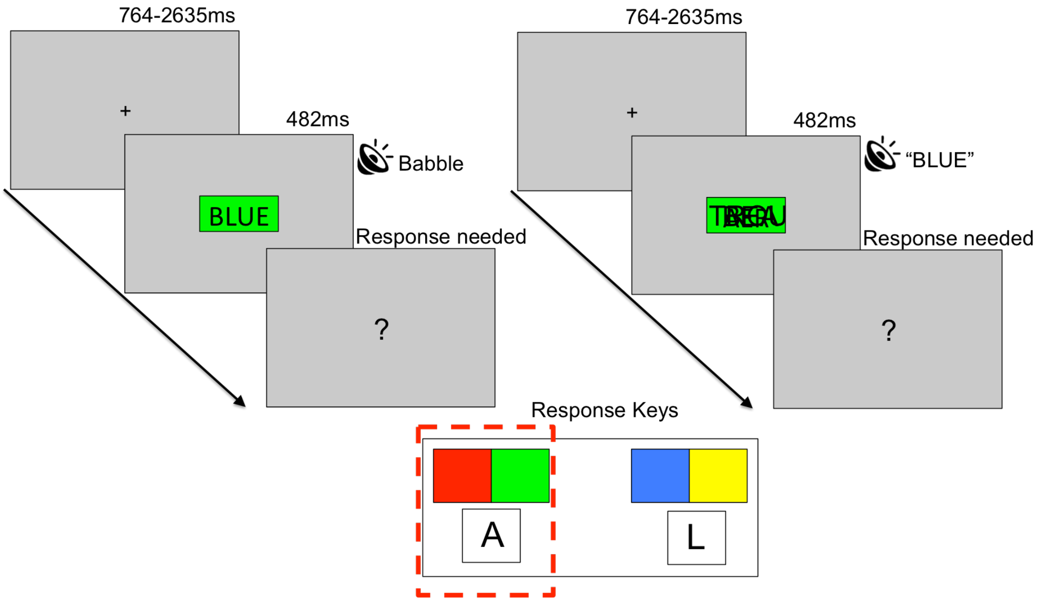


Figure 8. Unimodal (left) and cross-modal (right) Stroop protocol. Red dashed box indicates the correct response selection.

On each trial a fixation point was presented for 764-2635ms. Following this a coloured rectangle overlaid by a written word (in unimodal conditions) or visual babble (in cross-modal conditions) was presented for 482ms. Visual stimuli were presented simultaneously with auditory babble (in unimodal conditions) or a spoken colour-word (in cross-modal conditions) embedded in 60dB Brown noise. Participants identified the *colour of the rectangle* by pressing the left or right key and were instructed to ignore any written or spoken information. Participants were told they could respond as soon as they saw the ticket and to be as fast and accurate as possible. If no response had been made following stimulus offset a question mark was presented, signalling the need for a response.

In addition to sensory condition (unimodal vs. cross-modal) there were four levels of congruency; congruent (i.e. “RED” with a red rectangle), incongruent but mapped to the same response; stimulus-incongruent (i.e. “GREEN” with a red rectangle), incongruent and mapped to a different response; response-incongruent (i.e. “BLUE” with a red rectangle), or incongruent and mapped to no response; “neutral” (i.e. “BROWN” with a red rectangle).

Participants completed the unimodal and cross-modal conditions in separate blocks (counterbalanced across participants) each containing 96 trials. An optional break was offered after 48 trials. A 20 trial practice block was given prior to the first block and a 10 trial practice block prior to the second block to accustom participants to the new stimuli. Each 96 trial block contained 24 trials of each congruency condition presented in a pseudorandomised order for each participant such that no immediate repeats in the colour rectangle or colour-word would occur and congruency condition would not be repeated more than twice.

Analysis and Results

In this section, I first provide an overview of the derived threshold data followed by analysis of Stroop performance. It should be noted that, to remain conservative, throughout this analysis I discuss *p* values =< .01 as significant, however, all *p* values are reported throughout.

Thresholds

Mean visual and auditory thresholds for identifying words with different coloured backgrounds are shown in Table 2. To justify methods used with children (i.e. the selected starting values in visual MOA and the use of a single auditory MOA) I first provide analyses of younger and older adults (showing lower visual thresholds on some colour backgrounds compared with others and showing colour background had no effect on auditory threshold). This is followed by a full analysis comparing thresholds across age groups.

Visual Thresholds

A 4 (colour background) x 2 (age group; young adults vs. older adults) ANOVA comparing visual thresholds in adults showed a main effect of colour background ($F(3, 183) = 38.38$, $p < .001$, $\eta_p^2 = .39$). Thresholds for reading words on yellow ($M = 4.04\%$ opacity, $SE = 0.48\%$) and green ($M = 3.75\%$ opacity, $SE = 0.22\%$) were significantly lower than thresholds for reading words on red ($M = 6.63\%$, $SE = 0.47\%$) and blue ($M = 7.63\%$, $SE = 0.51\%$) ($p < .001$ for all comparisons). Thresholds did not significantly differ between green and yellow backgrounds ($p = 1$) or blue and red backgrounds ($p = .27$). This justified the use of lower starting values for yellow and green compared with red and blue for the method-of-adjustment in children.

A significant interaction was also seen between colour background and age group ($F(3, 183) = 7.73$, $p < .001$, $\eta_p^2 = .07$) this occurred because although there was a main effect of age group ($F(1, 61) = 7.91$, $p = .007$, $\eta_p^2 = .12$), in which thresholds were higher for older adults, this was significant with green ($p = .008$), red ($p = .004$) and blue ($p < .001$) backgrounds but not yellow backgrounds ($p = .93$) in which thresholds were persistently low across younger and older adults.

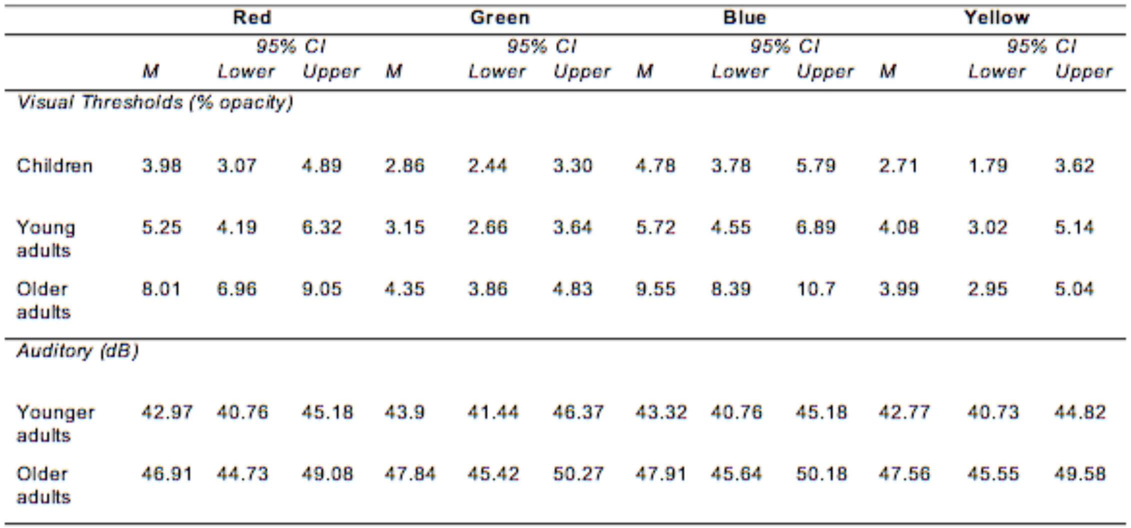
A 4 (colour background) x 3 (age group) ANOVA comparing visual thresholds across all age groups showed a main effect of age group ($F(2, 102) = 16.48$, $p < .001$, $\eta_p^2 = .24$), a main effect of colour background ($F(2.78, 283.1) = 67.08$, $p < .001$, $\eta_p^2 = .34$) and an interaction between age group and colour background ($F(5.55, 283.1) = 9.011$, $p < .001$, $\eta_p^2 = .1$). The main effect of age occurred because older adults' visual thresholds were significantly higher than children ($p < .001$) and young adults ($p = .002$). However children and young adults did not differ ($p = .18$). The main effect of colour background occurred as thresholds for detecting words written on green and yellow were significantly lower than thresholds for detecting words on blue and red ($p < .001$ for all comparisons). However thresholds for detecting words on yellow and green did not significantly differ from each other ($p = 1$), as did thresholds for blue and red ($p = .022$). Simple main effects showed the interaction between age group and colour background occurred because the main effect of age group was significant for all colours apart from yellow ($p < .001$ for all compared with $p = .09$), in which thresholds were comparable low across age groups.

Auditory Thresholds

A 4 (colour rectangle; red, green, blue, yellow) x 2 (age group; older adults, younger adults) ANOVA showed a main effect of age group ($F(1, 61) = 8.75$, $p = .004$, $\eta_p^2 = .13$) in which auditory thresholds were higher for older adults. Importantly, there was no significant effect of colour on auditory threshold ($F(2.65, 161.89) = 1.37$, $p = .257$, $\eta_p^2 = .02$) and this did not interact with age group ($F(2.65, 161.9) = .37$, $p = .75$, $\eta_p^2 =.01$). Bayesian analyses showed strong support for the conclusion that auditory threshold did not differ between coloured backgrounds as these results were 9.255 times more likely to occur under the null ($BF_{01} = 9.255$, $BF_{10} = .108$)– thus verifying the use of deriving auditory thresholds in the presence of a single coloured background in children.

To compare auditory thresholds across age groups the mean of the auditory thresholds derived in each adult participant was used to compare with the single auditory threshold derived in children. A 3-way ANOVA comparing this single auditory threshold value across age groups showed a significant effect of age group ($F(2, 102) = 6.73$, $p = .002$, $\eta_p^2 =.12$). Older adults' auditory thresholds ($M = 47.56$dB, $SE = .87$) were higher than younger adults' ($M = 43.24$dB, $SE = .88$, $p = .002$) and childrens' ($M =$

44.29dB, *SE* = .76, *p* = .017). Auditory threshold did not significantly differ between children and young adults (*p* = 1).



	Red			Green			Blue			Yellow		
		95% CI			95% CI			95% CI			95% CI	
	M	*Lower*	*Upper*	*M*	*Lower*	*Upper*	*M*	*Lower*	*Upper*	*M*	*Lower*	*Upper*
Visual Thresholds (% opacity)												
Children	3.98	3.07	4.89	2.86	2.44	3.30	4.78	3.78	5.79	2.71	1.79	3.62
Young adults	5.25	4.19	6.32	3.15	2.66	3.64	5.72	4.55	6.89	4.08	3.02	5.14
Older adults	8.01	6.96	9.05	4.35	3.86	4.83	9.55	8.39	10.7	3.99	2.95	5.04
Auditory (dB)												
Younger adults	42.97	40.76	45.18	43.9	41.44	46.37	43.32	40.76	45.18	42.77	40.73	44.82
Older adults	46.91	44.73	49.08	47.84	45.42	50.27	47.91	45.64	50.18	47.56	45.55	49.58

Table 2. Mean (*M*) visual and auditory thresholds and 95% confidence intervals (CI) for reading and hearing words presented with different coloured backgrounds in each age group in experiment 1.

Stroop performance

This study aimed to answer the following two questions:

1. Does general interference differ between unimodal and cross-modal conditions, and is this different between age groups?
2. Can any differences between unimodal and cross-modal interference be explained by comparing stimulus- and response-interference?

To answer these questions general interference, stimulus- and response-interference ratios were defined as noted in Equation 3 and Equation 4. The use of ratio scores allows for comparison of interference whilst controlling for general differences in speed across age groups. However, a full analysis of raw data (prior to calculation of ratios) is provided within the Appendix (page 300). These supplementary analyses are consistent with the findings reported here, in which ratio scores were used throughout. Notably, analysis of raw data showed the “neutral” condition, in which the word “BROWN” was presented actually introduced stimulus-interference. For this reason, neutral trials were not considered within the main analysis.

Outliers were removed from each participants’ data by identifying response times that fell outside the range of the absolute deviation around the median (Leys, Ley, Klein, Bernard, & Licata, 2013). Outlying participants were identified and removed by calculating the mahalanobis distance of each subject from the chi-squared distribution of their age group based upon response time and accuracy during congruent, stimulus-incongruent and response-incongruent trials under unimodal and cross-modal conditions. Cases holding a probability of < .001 of belonging to the population were removed from analyses. This resulted in 9 participants being removed (3 children, 2 young adults and 4 older adults). One older adult was also removed due to difficulty deriving an appropriate auditory threshold.

General Interference

Response time and accuracy data were submitted to two separate 2 (sensory condition) x 3 (age group) ANOVAs to compare unimodal to cross-modal interference. Significant interactions were followed up with simple main effects analyses adjusted for multiple comparisons via Bonferroni correction. Significant main effects of age group were followed with post-hoc quadratic tests to examine whether each interference type followed a U-shape trajectory across age groups. Across analyses, a series of Bonferroni corrected t-tests were also performed to compare ratio scores to 1. This provides a comparison of each interference score relative to baseline. The results of these comparisons are indicated within each Figure. Results for response times and accuracy are reported below and shown in Figure 9.

Response Times

There was no effect of age group on response time ratios (*F(*2, 102) = 3.17, *p* = .05, η_p^2 =.06). A significant effect of sensory condition (*F(*1, 102) = 9.51, *p* = .003, η_p^2 = .09) occurred, which interacted with age group (*F(*2, 102) = 5.26, *p* = .007, η_p^2 = .09). Although general interference was higher (i.e. caused more slowing) under unimodal compared with cross-modal conditions, this arose due to a significant difference in older adults (*p* < .001), which was not present in young adults (*p* = .77) or children (*p* = .16). A quadratic test indicated a significant U-shape function under unimodal conditions (*F(*1, 102) = 8.58, *p* = .004, η_p^2 = .08) which did not occur under cross-modal conditions (*F(*1, 102) = .01, *p* = .93, η_p^2 < .001).

Accuracy

There was no main effect of sensory condition on accuracy (*F(*1,102) = 4.02, *p* = .5, η_p^2 = .04), no interaction between sensory condition and age (*F(*2, 102) = 3.72, *p* = .03, η_p^2 = .07) and no main effect of age group (*F(*2, 102) = 1.14, *p* = .32, η_p^2 = .02). Quadratic trends did not reach significance in either unimodal (*F(*1, 102) = 5.93, *p* = .02, η_p^2 = .06) or cross-modal (*F* (1, 102) = 1.04, *p* = .31, η_p^2 = .009) conditions.

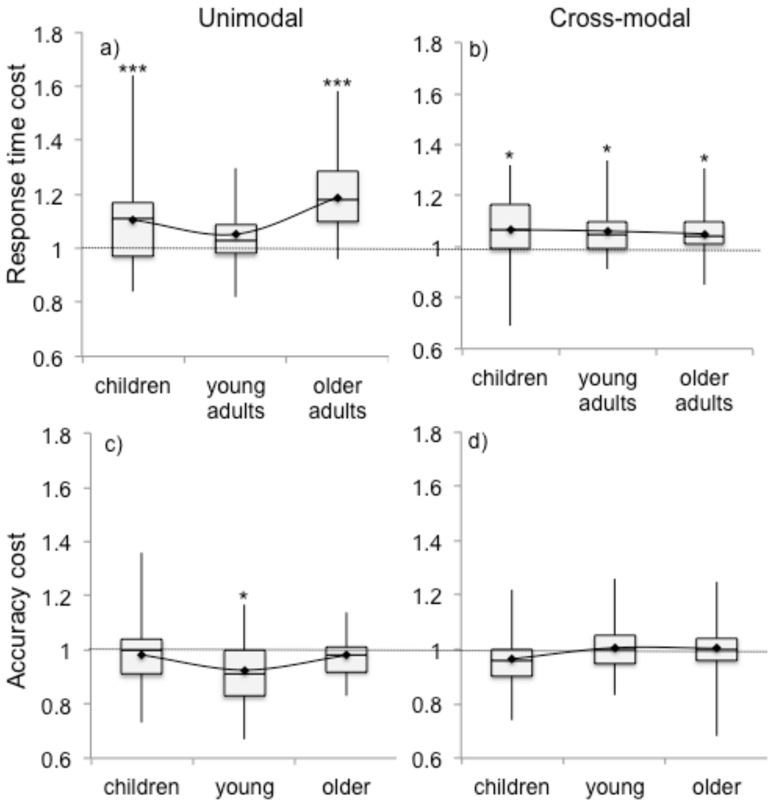


Figure 9. General interference in unimodal (left) and cross-modal (right) conditions in terms of response time (top) and accuracy (bottom). Response time ratios greater than 1 indicate slowing. Accuracy ratios less than 1 indicate accuracy decrements. Black diamonds = means used for analyses. Line connecting diamonds = the extent of the quadratic (U-shape) trajectory. Asterisks = results of t-tests comparing ratio to 1 (* = .05, ** = .01, ***= .001) Bonferroni corrected for 6 comparisons. Box = interquartile range. Whiskers = minimum and maximum points. Central line = median.

Stimulus- and response-interference

To examine differences in the contribution of stimulus- and response-interference under unimodal and cross-modal conditions, stimulus- and response-interference ratios for response time and accuracy were submitted to two separate 2 (sensory condition: unimodal, cross-modal) x 2 (interference type: stimulus-interference, response-interference) x 3 (age group: children, young adults, older adults) ANOVAs. Results for response times and accuracy are reported below and shown in Figure 10.

Response Times

Unimodal interference was greater than cross modal interference ($F(1,102) = 9.34$, $p = .003$, $\eta_p^2 = .08$) and this interacted with age group ($F(2, 102) = 4.86$,

p = .01, η_p^2 = .08), but did not interact with interference type (*F(*1, 102) = .62, p = .54, η_p^2 = .01). The difference between stimulus- and response-interference did not reach significance (*F(*1, 102) = 4.35, p = .04, η_p^2 = .04) and did not interact with age group (*F(*2, 102) = 0.09, p = .92, η_p^2 = .002) or sensory condition (*F(*1, 102) = 1.36, p = .25, η_p^2 = .01). There was no three-way interaction between sensory condition, age group and interference type (*F*(2, 102) = .62, p = . 54, η_p^2 = .01). In line with analyses of general interference, the interaction between sensory condition and age group arose because interference was greater in unimodal compared with cross-modal conditions but only in older adults ($p< .001$) and not young adults (p = .79) or children (p = .14).

Quadratic trends did not reach significance for stimulus- or response-interference under unimodal (stimulus-interference: *F*(1, 102) = 5.32, p = .02, η_p^2 = .05, response-interference: *F*(1, 102) = 1.377, p = .24, η_p^2 = .01) or cross-modal (stimulus-interference: *F*(1, 102) = .02, p = .88, $\eta_p^2 < .001$, response-interference *F*(1, 102) = .001, p = .98, $\eta_p^2 < .001$) conditions.

Accuracy

There was no main effect of sensory condition for accuracy (*F*(1, 102) = 4.29, p = .04, η_p^2 = .04). The effects of interference type (*F*(1,102) = 5.27, p = .02, η_p^2 = .05) and age group (*F*(2, 102) = 5.27, p = .47, η_p^2 = .01) also failed to reach significance. The interaction between sensory condition and age failed to meet significance (*F*(2, 102) = 4.22, p = .02, η_p^2 = .07), as did the interaction between interference type and age group (*F*(2, 102) = .38, p = .68, η_p^2 = .01) and the three-way interaction between sensory condition, interference type and age group (*F*(2, 102) = 1.55, p = .22, η_p^2 = .03). However, a significant interaction occurred between sensory condition and interference type (*F*(1, 102) = 11.86, $p < .001$, η_p^2 = .10). Stimulus- and response-interference only significantly differed from one another under unimodal (*p* < .001) but not cross-modal (*p* = .566) conditions.

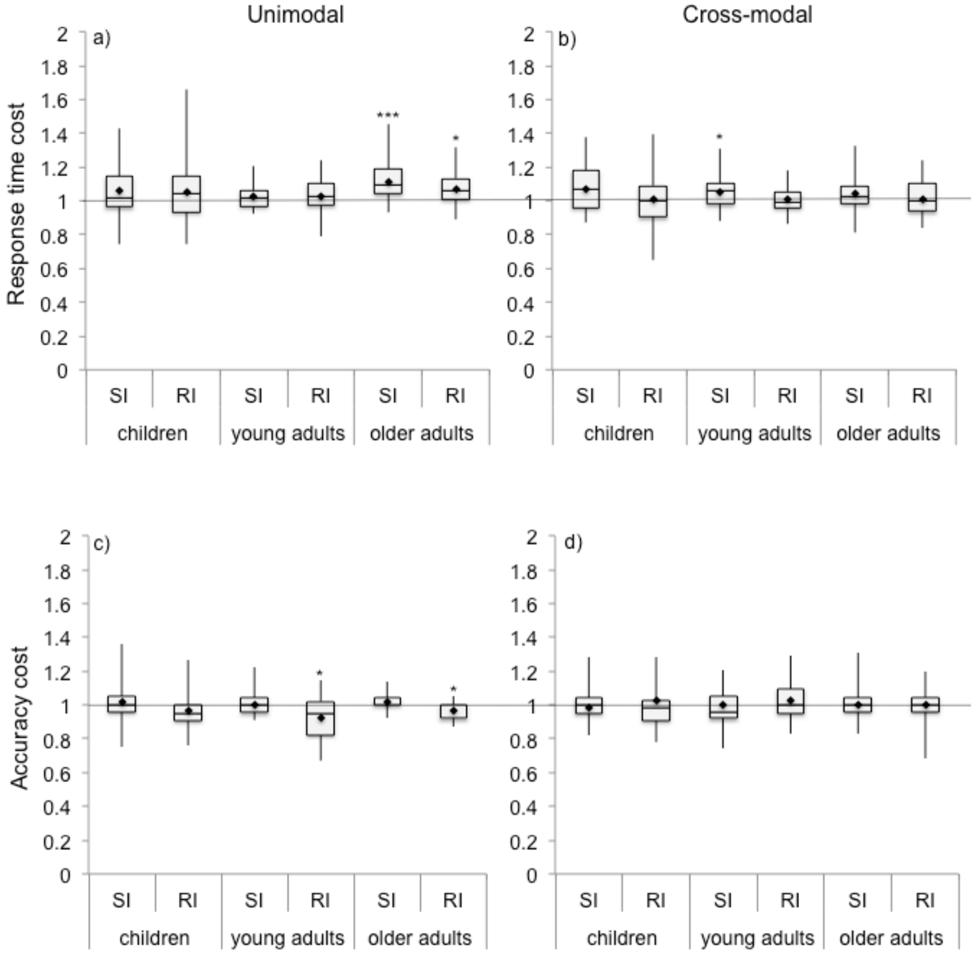


Figure 10. Stimulus (SI) and response (RI) interference ratios in terms of response time (top) and accuracy (bottom) under unimodal (left) and cross-modal (right) conditions. For response times, values higher than 1 indicate slowing. For accuracy, values lower than 1 indicate accuracy decrements. Black diamonds indicate means used for analyses. Asterisks indicate t-tests comparing ratio to 1 (* = .05, ** = .01, ***= .001) Bonferroni corrected for 12 comparisons.

Developmental Trajectory Analyses

Due to the wide age range encompassed within the child (6 - 11 years) and older adult (60 - 84 years) samples, two separate 2 (sensory condition) x 2 (interference type) ANCOVAs were performed in which age in decimals was held as a covariate, thus detecting whether age influenced the pattern of effects observed in these age groups. These analyses were conducted for both response time ratios and accuracy ratios and are shown in Figure 11.

Response Times

Developmental trajectory analyses on response times yielded no main effects or interactions in adults or children (Table 3).

	Children			**Older adults**		
Effect	***F(1,40)***	***p***	η_p^2	***F(1,30)***	***p***	η_p^2
Sensory Condition	2.13	.15	0.05	.214	.65	.01
Sensory condition x Age	1.57	.22	0.04	.915	.35	.03
Interference type	0.03	.87	<.01	.58	.58	.01
Interference Type x Age	0.00	.96	<.01	.5	.49	.02
Sensory condition x interference type	0.61	.44	0.02	.44	.51	.02
Sensory condition x interference type x Age	1.05	.31	0.03	.475	.5	.02

Table 3. Results from 2 (sensory condition: unimodal, cross-modal) x 2 (interference type: stimulus-, response-interference) ANCOVA on response times in children and older adults. No significant effects were found. Residuals for child and older adult comparisons 40 and 30 respectively.

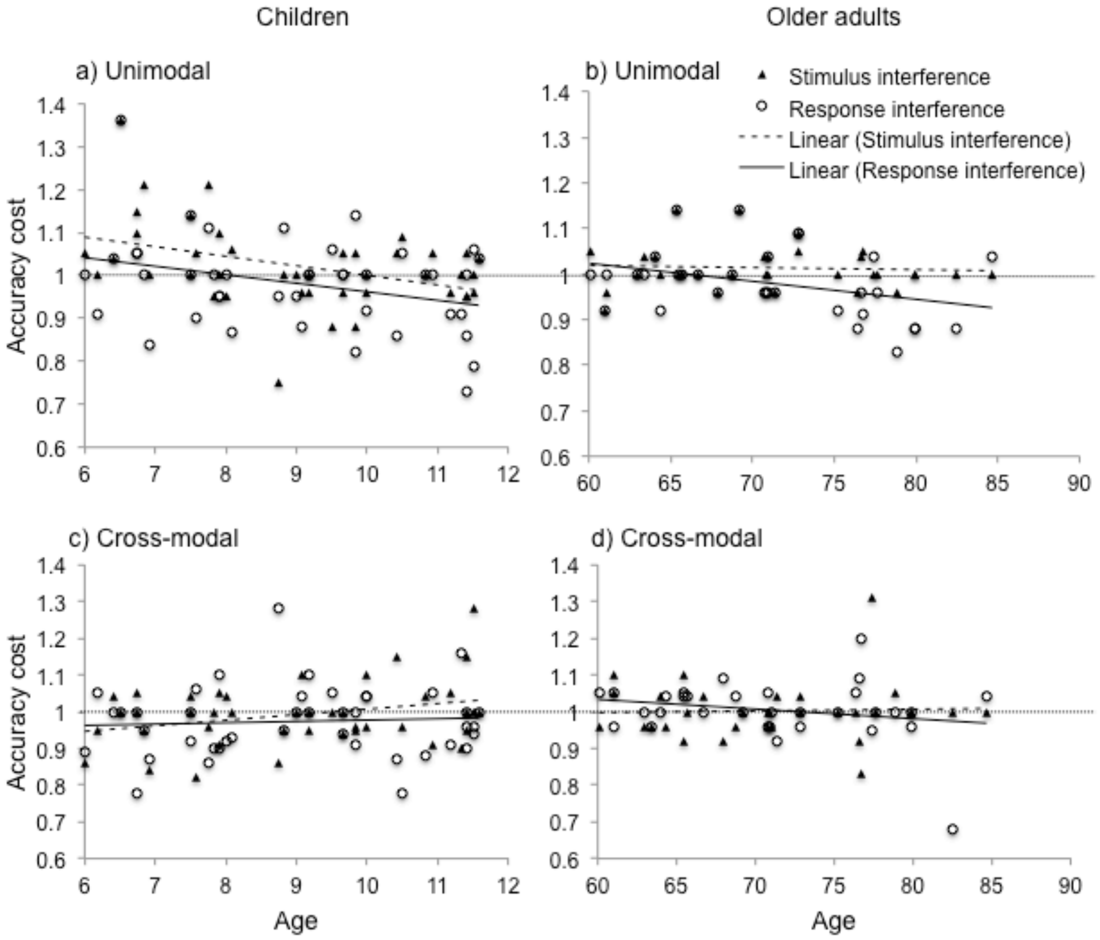


Figure 11. Accuracy ratios for stimulus-interference (triangles and dashed line) and response-interference (circles and continuous line) under unimodal (top) and cross-modal (bottom) conditions. Ratios are shown for children (left) and older adults (right). A significant reduction in accuracy cost occurred across childhood under unimodal conditions.

Accuracy

In children the ANCOVA showed a main effect of sensory condition (*F*(1, 40) = 8.83, *p* = .005, η_p^2 = .16) which covaried with age (*F*(1, 40) = 8.27, *p* = .006, η_p^2 = .15). This occurred because unimodal accuracy costs increased (i.e. accuracy was worse) with age whilst this did not occur under cross-modal conditions. There was no effect of interference (*F*(1, 40) = .77, *p* = . 38, η_p^2 = .02) and this did not covary with age (*F*(1, 40) = .26, *p* = .62, η_p^2 = .01). There was no interaction between sensory condition and interference type (*F*(1, 40) = 6.09, *p* = .02, η_p^2 = .12) and this did not covary with age (*F*(1, 40) = 4.97, *p* = .03, η_p^2 = .1).

It should be noted that the effects of stimulus- and response-interference upon accuracy costs are contradictory. Stimulus-interference would increase accuracy (as the distractor primes the correct response) whilst response-interference would reduce accuracy (as the distractor primes the incorrect response). Thus increased accuracy costs (i.e. reduced accuracy) in unimodal conditions across childhood can to be attributed to decreases in stimulus-interference and increases in response-interference with development (Figure 11; Cragg, 2016).

In older adults the effects of sensory condition and interference type did not reach significance ($F(1, 30) = .43$ $p = .52$, $\eta_p^2 = .02$ and $F(1, 30) = 2.7$ $p = .11$, $\eta_p^2 = .07$ respectively). Neither of these effects were shown to covary with age ($F(1,30) = .63$, $p = .43$, $\eta_p^2 = .02$ and $F(1, 30) = 3.34$, $p = .08$, $\eta_p^2 = .09$ respectively). There was no interaction between sensory condition and interference type ($F(1, 30) = .016$, $p = .9$, $\eta_p^2 = .001$).

Response Time Distributions

Unimodal response-interference but not cross-modal response-interference increased accuracy costs (see above). Developmental trajectory analyses also suggested increased unimodal, but not cross-modal, accuracy costs across childhood. Together these findings suggest that response-interference may occur under unimodal conditions (and thus reduce accuracy) but not cross-modal conditions.

On the other hand, initial analyses suggested stimulus- and response-interference both slowed response times under unimodal and cross-modal conditions. One possible explanation for contradictory results in response time and accuracy data might be that no differences in response times are seen because data were collapsed across the response time distribution. Chen et al. (2011) show that stimulus- and response-interference occur at different time-points across the response time distribution. Response-interference occurs at longer latencies, whilst stimulus-interference appears uniformly distributed across response times. To address this, the 10th-90th percentiles for each participant's response time distributions were identified (Figure 12). If interference occurred at specific time points in the response time distribution a main effect of percentile would be expected. If stimulus- and response-interference occur at different latencies, as predicted by Chen et al. (2011), a percentile by interference type interaction would be expected. If unimodal and cross-modal stimulus- and response-interference are different a three-way interaction between percentile, interference type and sensory condition would be expected. Finally, if these differences are only observed in some and not all age groups a four-way interaction would be expected.

A 2 (sensory condition) x 2 (Interference type) x 9 (percentile) x 3 (age group) ANOVA showed no main effect of sensory condition (*F*(1, 102) = 4.75, *p* = .03, η_p^2 = .04) or interference type (*F*(1, 102) = .550, *p* = .46, η_p^2 = .01) or age (*F*(2,102) = 2.56, *p* = .08, η_p^2 = .05). There was a main effect of percentile (*F*(2.081[8], 229.12) = 14.439, *p* < .001, η_p^2 = .12). There was no two-way interaction between interference type and percentile (*F*(2.24, 228.03) = 1.039, *p* = .362, η_p^2 = .01), and the three-way interaction between percentile, interference type, and sensory condition did not reach significance (*F*(2.25, 229.12) = 2.53, *p* = .08, η_p^2 = .02). There was however a significant four-way interaction between percentile, sensory condition, interference type, and age (*F*(4.49, 229.12) = 4.05, *p* = .002, η_p^2 = .07).

Post-hoc t-tests showed that in children unimodal response-interference was significantly higher than stimulus-interference at the 90th percentile (*p* = .002). Conversely, under cross-modal conditions, the opposite pattern occurred and stimulus-interference was significantly higher than response-interference at the 90th percentile (*p* = .01). In young adults, and older adults, stimulus- and response-interference did not significantly differ

---

[8] Any degrees of freedom throughout this thesis reported to the decimal are Greenhouse Geisser corrected, indicating that Mauchleys test for sphericity showed this assumption had been violated.

from one another at any point in the response time distribution, under unimodal or cross-modal conditions.

The analysis of response time distribution indicated that in children unimodal and cross-modal interference occurring at longer latencies arose from different types of interference. Under unimodal conditions response-interference was highest at longer latencies whilst stimulus-interference remained stable across response times. Under cross-modal interference, stimulus-interference peaked at longer latencies, whilst response-interference remained stable.

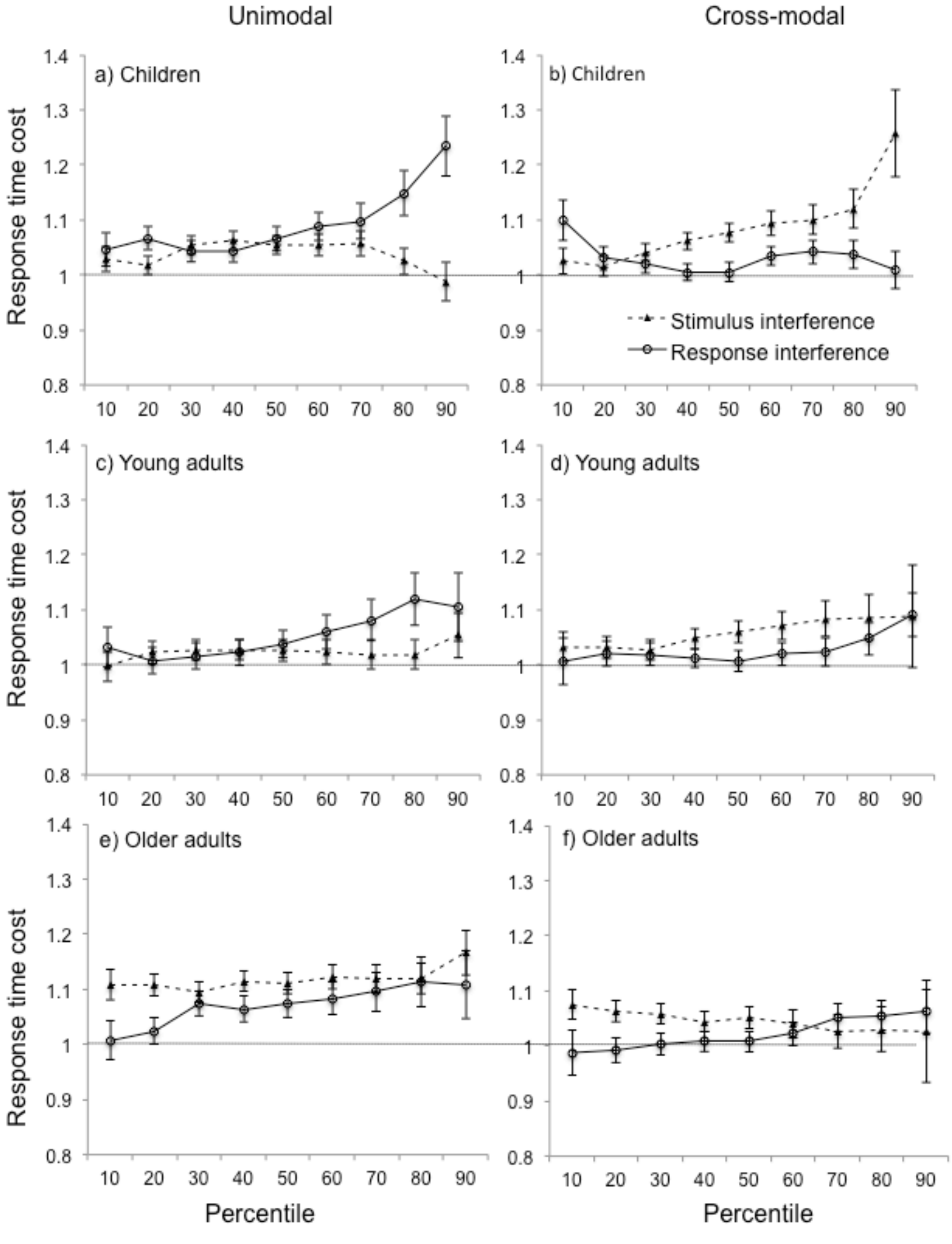


Figure 12. Response time distributions for stimulus- and response-interference under unimodal (left) and cross-modal (right) conditions in children (a/b) young adults (c/d) and older adults (e/f). Each data point shows the mean response times for 10th-90th percentiles. Error bars indicate standard error.

Discussion

Experiment 1 showed unimodal and cross-modal interference differed in multiple ways. Firstly, unimodal interference manifested a U-shape trajectory from childhood, to young adults to old age. Cross-modal interference did not manifest a U-shape trajectory and older adults showed substantially more interference under unimodal compared with cross-modal conditions. Secondly, unimodal but not cross-modal response-interference reduced accuracy. Thirdly, between the ages of 6 and 11 years, accuracy decreased for unimodal, but not cross-modal tasks, although note that this may have resulted from the facilitatory influence of stimulus-interference conditions in early childhood. Finally, in childhood, unimodal and cross-modal interference showed opposing patterns of stimulus- and response-interference across the response time distribution. Under unimodal conditions response-interference peaked at longer response times whilst stimulus-interference remained stable. Under cross-modal conditions stimulus-interference peaked at longer latencies whilst response-interference remained stable.

However, “cross-modal” distractions in Experiment 1 were always auditory distractors whilst participants focused on vision. Thus, results may have occurred due to the distractor being auditory in nature as opposed to cross-modal per se. Evidence suggests a shift in sensory weighting across the lifespan, such that children prefer auditory information whilst adults give precedence to visual information (Barnhart, Rivera, & Robinson, 2018; Colavita, 1974; Diaconescu, Alain, & McIntosh, 2011; Nava & Pavani, 2013). Furthermore, older adults with mild levels of age-related hearing loss show cortical reorganisation, such that auditory cortices are recruited for visual tasks (Campbell & Sharma, 2014). Given shifts in sensory weighting and cortical allocation, it might be expected that children would be more susceptible to distraction from auditory sources whilst older adults are more susceptible to distraction from visual sources. This is in line with the existing debate as to whether older adults can ignore vision whilst focusing on audition (Guerreiro, Adam, et al., 2014; Guerreiro, Anguera, et al., 2014; Guerreiro et al., 2015, 2010, 2013; Van Gerven & Guerreiro, 2016).

Given the existing literature, it was predicted that older adults should be able to focus on vision whilst ignoring audition (as seen in Experiment 1) but should find it more difficult to ignore vision whilst focusing on audition. Furthermore, based on shifts in sensory dominance, it was predicted that young children should find auditory distractors more difficult to ignore whilst focusing on vision than vice versa.

Experiment 2

Experiment 2 aimed to explore whether the pattern of results seen in Experiment 1 held for two types of cross-modal distraction: focusing on vision whilst ignoring auditory distractors (as in Experiment 1) and focusing on audition whilst ignoring visual distractors.

Method

Participants

Sample size calculation was conducted in accordance with Experiment 1, indicating a need for a minimum of 29 participants per age group (87 in total), a criterion that was met by all three samples. Thirty young adults (mean age 25.79 years, range 22 - 33, 21 female), 53 children (mean age 9.38 years, range 6 - 11 years, 32 female) and 36 older adults (mean age 71.38 years, range 60 - 84 years, 21 female) took part. All participants were recruited using the same methods reported in Experiment 1.

One child was excluded due to developmental disorders reported by parents. Five older adults were later excluded, two due to inability to derive an appropriate hearing threshold and three due to the use of hearing aids.

Equipment

The equipment used was identical to those detailed in Experiment 1. However the experimenter used a headphone splitter to check that responses being made in the practice trials were correct. This was necessary as in the new cross-modal condition participants had to respond to a spoken word.

Stimuli

Stimuli were identical to those used in Experiment 1 apart from that neutral trials contained no information other than the relevant dimension (i.e. a coloured rectangle to be identified or a spoken word to be identified). As in Experiment 1, thresholds were assessed to present all stimuli at 10x (20dB) above threshold with a maximum visual contrast of 100% and auditory presentation of 65dB. Auditory stimuli had to be presented at maximum for 27 (of 52) children, 2 (of 30) younger adults, and 16 (of 30) older adults included in the final analysis. However, stimuli were still judged to be clearly audible in these participants (*M* =16.67dB above threshold range = 6-20dB above threshold).

Procedure

The task was identical to that used in Experiment 1, however, there was no unimodal condition (i.e. ignoring written words whilst focusing on the colour of the rectangle) and this was replaced with a second type of cross-modal condition. The "ignore auditory" condition was identical to the cross-modal condition used in Experiment 1. In this condition, participants were instructed to sort coloured "tickets" (rectangles) into two boxes based on the colour of the rectangle whilst ignoring a spoken word. In the new "ignore visual" condition, participants were instructed to sort the tickets based on the colour-word they heard, whilst ignoring the actual colour of the ticket. An emphasis was made that participants were not allowed to close their eyes and must focus on the ticket at all times.

Analysis and Results

Replication of Ignore Auditory Result

The first goal of Experiment 2 was to replicate the findings from the cross-modal condition of Experiment 1 (i.e. compare the cross-modal condition of Experiment 1 to the "ignore auditory" condition in Experiment 2). Two 2 (experiment) x 3 (age group) ANOVAs were used to compare general interference ratios in terms of response time and accuracy between the three age groups. Two further 2 (experiment) x 2 (interference type) x 3 (age group)

ANOVAs were used to compare stimulus- and response-interference between age groups. To assess support for the null (that there was no difference between Experiments) Bayesian factors were calculated alongside frequentist statistics.

	Accuracy					**RT**				
Models	**BF $_M$**	**BF $_{01}$**	***F***	***p***	η_p^2	**BF $_M$**	**BF $_{01}$**	***F***	***p***	η_p^2
General interference										
Null	2.89	1.00				1.25	1.00			
Exp.	0.76	2.63	1.35	.25	<.01	0.73	1.54	1.51	.22	<.01
Stimulus- and response-interference										
Null	.001	1.00				11.07	1.00			
Exp.	2.9e-4	5.24	1.19	.28	<.01	2.35	3.7	1.5	.22	<.01

Table 4. ANOVA and Bayesian statistics for comparison of effects between Experiment 1 and 2 for accuracy and response times (RT). BF $_M$ = change from prior to posterior model odds. BF $_{01}$ = Bayes factor for each model against the alternative (favour for the null). Exp. = Experiment

Table 4 shows the resulting statistics comparing general interference and stimulus- and response-interference (in terms of response time and accuracy) between experiments. Critically both sets of analyses indicated no significant difference between Experiment 1 and 2. Bayesian analyses in all cases were in favour of the null compared with the alternative and this was most convincing when considering stimulus- and response-interference separately. Differences in accuracy between experiments were 5.24 times more likely under the null hypothesis and differences in response times were 3.7 times more likely under the null hypothesis. Thus, the “ignore auditory” condition of Experiment 2 was considered as a replication of the cross-modal condition in Experiment 1.

Auditory Vs. Visual Distractors

The main goal of Experiment 2 was to explore whether the findings from Experiment 1 generalised to cross-modal distraction occurring in the opposing modality. If this were the case we would expect that the U-shape trajectory seen across age groups in unimodal conditions in Experiment 1 would not be seen with either type of cross-modal distractor (older adults would be good at ignoring both auditory and visual distractors). Alternatively, if the findings from Experiment 1 could be explained due to the nature of the cross-modal distractor, older adults should maintain the ability to ignore auditory distractors whilst focusing on vision, but not vice versa.

As in Experiment 1, I first outline general interference with both auditory and visual cross-modal distractors in each of the age groups (Figure 13) and follow this by separating stimulus- and response-interference (Figure 14).

General Interference

Response times

A 2 (distractor type: auditory, visual) x 3 (age group: children, young adults, older adults) ANOVA showed a main effect of age group (F(2, 109) =4.5, p = .01, η_p^2 = .08). Children showed more interference than older adults, but this did not reach the conservative cut-off for significance (*p* = .02). There was no significant difference between general interference in children and young adults (*p* = .12) or younger and older adults (*p* = 1). There was no difference between the two distractor types (*F*(1, 109) = .18, *p* = .67, η_p^2 = .002). There was no significant interaction between distractor type and age group (*F*(2, 109) = .85, *p* = .43, η_p^2 = .02).

Accuracy

A 2 x 3 ANOVA showed no main effect of distractor type (*F*(1, 109) = .67, *p* = .41, η_p^2 = .01) and no main effect of age group (*F*(2, 109) = 1.48, *p* = .23, η_p^2 = .03). There was no interaction between distractor type and age group (*F*(2, 109) = 1.18, *p* = . 31, η_p^2 = .02).

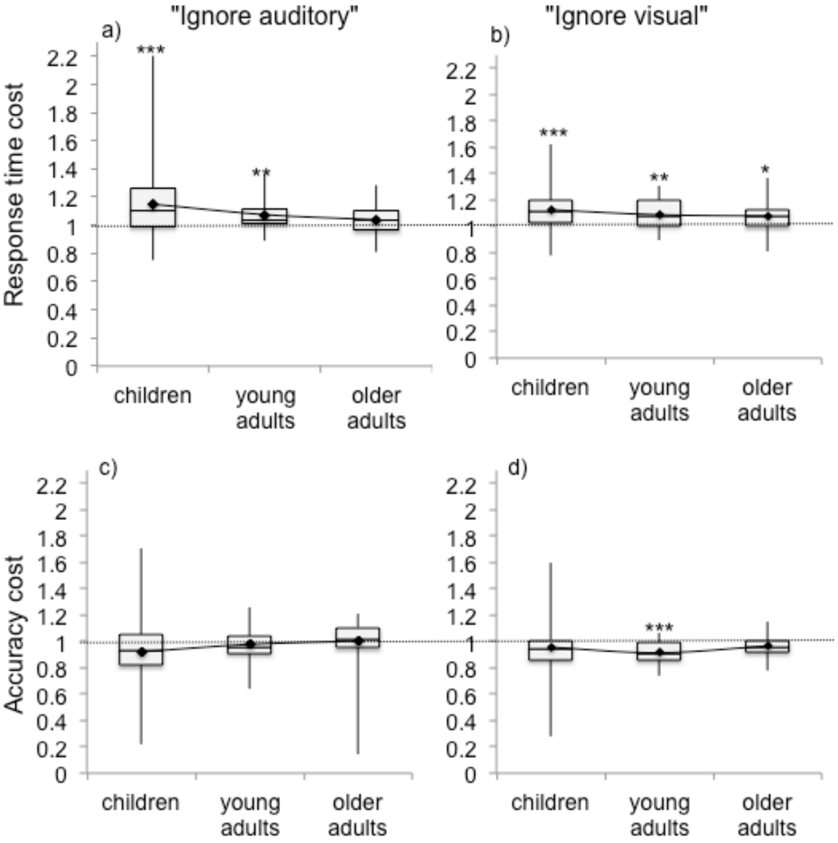


Figure 13. Response time (top) and accuracy (bottom) ratios for general interference in cross-modal conditions with auditory distractors (left) and visual distractors (right) in children, young adults and older adults. For response times, ratios higher than 1 indicate slowing. For accuracy, ratios lower than 1 indicate accuracy decrements. Black diamonds indicate means used for analyses. The line demonstrates the extent of the quadratic (U-shape) trajectory. Asterisks indicate t-test comparing ratio to 1 (* = .05, ** = .01, *** = .001) Bonferroni corrected for 6 comparisons.

Stimulus and Response-interference

Response times

A 2 (distractor type) x 2 (interference type) x 3 (age group) ANOVA showed a main effect of age group (*F*(2, 109) = 5.056, *p* = .01, η_p^2 = .09). Children showed significantly more interference (slowing) than older adults (*p* = .01) but not young adults (*p* = .11). Young adults did not significantly differ from older adults ($p = 1$). There was no effect of interference type ($F(1, 109) = .074$, $p = .79$, $\eta_p^2 = .001$) but age interacted with the effect of interference type ($F(2, 109) = 5.024$, $p = .01$, $\eta_p^2 = .08$). Children showed more stimulus-interference compared with younger and older adults, however, only the comparison between children and older adults reached significance ($p = .02$ and $p = .001$, respectively). Children did not differ from younger or older adults with regards to response-interference ($p = 1$ for both comparisons). Stimulus-interference caused significantly more slowing compared with response-interference in children ($p = .003$) but this difference was not significant in young adults ($p = .62$) or older adults ($p = .18$). There was no effect of distractor type ($F(1, 109) = .325$, $p = .57$, $\eta_p^2 = .003$) and no three-way interaction ($F(2, 109) = 0.24$, $p = .79$, $\eta_p^2 = .004$).

Accuracy

A 2 x 2 x 3 ANOVA showed a significant effect of interference type ($F(1, 109) = 14.12$, $p < .001$, $\eta_p^2 = .11$) which interacted significantly with age group ($F(2, 109) = 5.95$, $p = .004$, $\eta_p^2 = .09$). Accuracy costs were larger for response-interference vs. stimulus-interference (this finding did not occur under Experiment 1), and this difference only reached significance in children ($p < .001$) and not young adults ($p = .12$) or older adults ($p = .86$). Notably, this may also have occurred because stimulus-interference facilitated correct responses in children (producing mean ratios higher than 1), resulting in differences in accuracy between stimulus- and response-interference.

The main effects of age group and distracter type did not reach significance (*F*(2, 109) = .64, *p* = .53, η_p^2 = .01; *F*(1, 109) = .19, *p* = .67, η_p^2 = .002). There was no interaction between distractor type and age group (*F*(1, 109) = 1.35, *p* = .26, η_p^2 = .02), no interaction between distractor type and interference type (*F*(1, 109) = .18, *p* = .67, η_p^2 = .002) and no three-way interaction between distractor type, interference type, and age group (*F*(2, 109) = 2.12, *p* = .13, η_p^2 = .04).

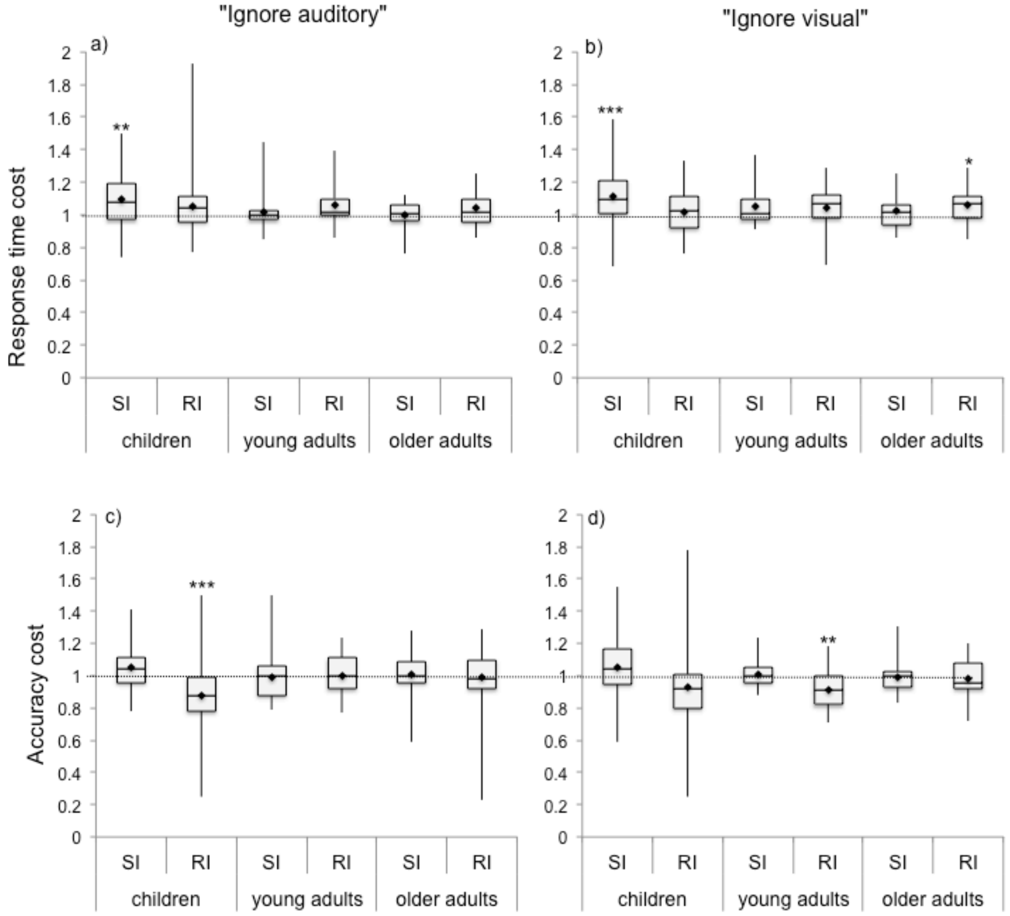


Figure 14. Stimulus- (SI) and response-interference (RI) with auditory (left) and visual (right) cross-modal distractors in terms of response time (top) and accuracy (bottom). For response time ratios, values greater than 1 indicate slowing. For accuracy ratios, values less than 1 indicate accuracy decrements. Black diamonds show means used for analyses. Asterisks indicate t-test comparing ratio to 1 (* = .05, ** = .01, ***= .001) Bonferroni corrected for 12 comparisons.

Discussion

The results from Experiment 2 suggested that neither type of cross-modal distraction produced the U-shape trajectory seen in the unimodal condition in Experiment 1. Rather, older adults appeared able to ignore both types of cross-modal distraction.

Notably, in Experiment 1 only unimodal and not cross-modal conditions produced accuracy costs (i.e. accuracy decrements) due to response-interference (there was a main effect of interference type that interacted with sensory condition). In Experiment 2, however, there was a main effect of interference type for cross-modal distraction. Response-interference produced accuracy costs with both visual and auditory distractors. This difference, however, arose only from children and may, in part, be due to facilitation on stimulus-interference trials in children. Increased stimulus-interference in children was evidenced in response times in Experiment 2, in which children were significantly slower than young adults and older adults but this was a result of stimulus-interference.

General Discussion

This was the first study to compare unimodal and cross-modal Stroop interference across childhood and old age whilst also considering sensory differences based on individual thresholds, and the first study to separate stimulus- and response-interference within a cross-modal paradigm. The questions addressed were a) whether unimodal and cross-modal interference follow similar patterns of development and deterioration across the lifespan b) if stimulus- and response-interference contribute to cross-modal, as well as unimodal, interference and c) whether the relative contribution of stimulus- and response-interference under unimodal and cross-modal conditions changes across the lifespan. Overall the findings indicated that different mechanisms underpin unimodal and cross-modal interference-control and that these mechanisms are differentially susceptible to developmental maturation and ageing. I begin this section by addressing each research question in turn before focusing in detail on interference-control in older adults and children.

Unimodal and cross-modal interference did not follow similar patterns of development and deterioration. Experiment 1 showed that unimodal interference was highest in children and older adults as compared with younger adults, producing a U-shape trajectory. Children also struggled to ignore auditory distractors whilst focusing on vision. However, older adults maintained the ability to ignore audition whilst focusing on vision. This finding was replicated in Experiment 2 where two different cross-modal conditions were compared. As such, unimodal and cross-modal interference do not appear to follow the same patterns of development and deterioration across the lifespan.

Stimulus- and response-interference were found to contribute differentially to cross-modal and unimodal interference. Experiment 1 showed that unimodal interference arose from both stimulus- and response-interference whilst cross-modal interference arose mainly from stimulus-interference. Under unimodal conditions, response times were slowed by the presence of conflicting information mapped to the same response. Participants were further slowed, and made errors, if the conflicting information was mapped to a different response. However, cross-modal response-conflict was not sufficient to produce accuracy decrements. This was also the case for adults, but not children, in Experiment 2. This suggests cross-modal interference arises mainly from stimulus-interference whilst unimodal interference takes effect at both the stimulus and response processing levels.

The relative contribution of stimulus- and response-interference under unimodal and cross-modal conditions changed across age groups. It has been proposed that children show more stimulus- than response-interference compared with adults (Cragg, 2016). The current findings support this, but also suggest the contribution of stimulus- and response-interference in childhood may differ between unimodal and cross-modal conditions. In Experiment 1 younger children were more accurate on stimulus-interference trials (showing facilitatory effects). This suggests more stimulus-interference in childhood. However, when comparing the pattern of stimulus- and response-interference across the response time distribution under unimodal and cross-modal conditions, children showed different patterns under unimodal and cross-modal conditions. Under unimodal conditions response-interference peaked at the longest response time latencies, whilst stimulus-interference remained constant (Chen et al., 2011). However, the opposite pattern was seen under cross-modal conditions. Experiment 2 also suggested children were more susceptible to stimulus-interference, as children, but not adults, were significantly slowed by stimulus-interference. Furthermore, accuracy in children was significantly lower on response-interference compared with stimulus-interference conditions, which, as in Experiment 1, may have partly resulted from facilitation of accuracy on stimulus-interference trials. In Experiment 2 cross-modal distractors only reduced accuracy in children. Thus, children appear to process cross-modal distraction differently from younger and older adults and this may in part be due to cross-modal interference occurring at different levels (i.e. also at the response selection levels) of processing in childhood.

Maintained Cross-Modal Interference-control in Ageing

The findings from Experiments 1 and 2 provided substantial support for the hypothesis that cross-modal interference is less susceptible to age-related decline (Guerreiro et al., 2010). Surprisingly, this was the case for both visual and auditory cross-modal distractions. This result is in line with fMRI data suggesting equivalent down-regulation of visual and auditory processing in older and younger adults during cross-modal attention (Guerreiro et al., 2015). However, this finding contradicts findings showing older adults may suppress audition whilst focusing on vision but not vice versa (Guerreiro, Adam, et al., 2014; Guerreiro, Anguera, et al., 2014; Van Gerven & Guerreiro, 2016). One explanation of differing findings proposed by Guerreiro et al. (2015) is that asymmetrical effects (i.e. an ability to ignore audition whilst focusing on vision but not vice versa) are seen in tasks in which auditory and visual information are presented concurrently, but symmetrical effects (i.e. an ability to ignore both visual and auditory cross-modal distraction) might occur when information is presented sequentially. In contrast to this, this study shows symmetrical effects in a task where stimuli were presented simultaneously. Two further explanations may account for the maintained suppression of visual and auditory distractors seen in this study. Firstly, the perceptual load of the task in this study may have been higher, permitting fewer cognitive resources for distractibility (Matusz et al., 2015). Secondly, it is arguable that the "ignore visual" condition in Experiment 2 was easier than the "ignore auditory" condition (see Limitations and Future Directions page 79).

It is possible that the demands of the current task and the number of stimuli presented (colour rectangle, written words/visual babble, spoken words/auditory babble and Brown noise) may have required more cognitive resources compared with previous literature (Guerreiro, Adam, et al., 2014; Guerreiro, Anguera, et al., 2014). Indeed, this study presented background noise throughout, and it has been shown that older adults require more cognitive resources to decipher speech in noise (Getzmann, Wascher, & Falkenstein, 2015) and systematic reviews support a link between speech in noise comprehension and cognitive ability (Dryden, Allen, Henshaw, & Heinrich, 2017). Accumulating evidence also suggests increased perceptual effort of general speech processing in ageing (Gagné, Besser, & Lemke, 2017). For example, using a dual-task paradigm, Tun, Mccoy and Wingfield (2009) found that, even when words were presented at supra-threshold intensities, older adults showed poorer performance on a secondary task (tracking a mouse on a screen) when recalling auditory information. Combined, the effects of increased perceptual load and increased perceptual effort of listening in older adults may have left fewer cognitive resources to be allocated to visual information, thus reducing cross-modal visual distraction (Lavie, 1995). Indeed, it has been shown that limited cognitive resources may sometimes shield younger children from cross-modal distraction (Matusz et al., 2015). Future research should aim to investigate whether perceptual load, listening effort and dual-task performance may predict the extent to which older adults are distracted in cross-modal environments.

As discussed in the introduction to this thesis, one mechanism underlying the ability to suppress cross-modal distractions is the suppression of activity in task-irrelevant sensory cortices. For example the suppression of auditory processing regions whilst focusing on vision (Ghatan, Hsieh, Petersson, Stone-Elander, & Ingvar, 1998; J. A. Johnson & Zatorre, 2005; Kawashima et al., 1995; Mozolic et al., 2008; Weissman et al., 2004). Neuroimaging studies investigating this process in ageing appear mixed. Some findings show older adults manifest less down-regulation of auditory cortices when processing visual information (Hugenschmidt, Mozolic, Tan, Kraft, & Laurienti, 2009) and some suggest age-equivalent suppression of visual and auditory cortices in cross-modal attention (Guerreiro et al., 2015) while others suggest intact suppression of auditory but not visual processing in cross-modal attention (Guerreiro, Anguera, et al., 2014). Other neuroimaging studies also suggest different, compensatory strategies may be used in older adults when ignoring irrelevant information (Allen & Payne, 2012). It is possible that this may also be the case for cross-modal control. Peiffer et al. (2009) found that older adults inhibit distinctly different regions of occipital cortex when ignoring visual information compared with young adults. Similarly, Diaconescu et al. (2013) found posterior parietal and medial frontal activity was increased in older adults relative to younger adults when presented with cross-modal stimuli, and found that this activity was related to faster detection of cross-modal stimuli. These findings suggest that the neural mechanisms used to support cross-modal interference control in ageing may undergo reorganisation, which may help support the normal behavioural performance seen in this study.

Increased Stimulus-interference and Differential Processing of Cross-Modal Distraction in Childhood

The current findings support the claim that children experience more stimulus-interference compared with adults (Cragg, 2016). In Experiment 1 these effects are seen in the developmental trajectory analysis, in which younger children counter-intuitively manifested higher accuracy under unimodal conditions, which decreased with age. As shown in Figure 4 this likely arose due to the combined effects of increasing response-interference and decreasing stimulus-interference across childhood. Increased response-interference in later childhood resulted in accuracy decrements, as the incongruent distractor primed an incorrect response. Conversely, increased stimulus-interference in early childhood resulted in increased accuracy, as the incongruent distractor primed/facilitated the correct response. Curiously, this effect was not seen under cross-modal conditions. In Experiment 2, both types of cross-modal distractor slowed response times in children, but this appeared attributable to stimulus-interference. Together, these findings suggest that younger children experience more stimulus-interference. Furthermore, the balance of stimulus- vs. response-interference in unimodal distraction changes from early to late childhood, whilst this does not appear to be the case for cross-modal distraction.

Across experiments, children appeared to process cross-modal distraction differently from adults and, as a result, were more susceptible to cross-modal distraction. In Experiment 1 unimodal stimulus- and response-interference followed similar response time distributions to those reported in adults, with response-interference peaking at longer response latencies (Chen et al., 2011). However, children showed the opposite pattern under cross-modal conditions, with stimulus-interference peaking at longer response latencies. This might explain why children, but not adults, appeared susceptible to cross-modal distractions in Experiment 2. It has been proposed that peripheral mechanisms may filter out cross-modal distraction at earlier processing stages (Guerreiro et al., 2010; Haxby et al., 1994), and this appears consistent with the current findings that cross-modal interference arises at stimulus-encoding stages. However, if cross-modal stimulus-interference peaks later in time in children, this suggests cross-modal distractors are more difficult to suppress at peripheral stages in childhood. Furthermore, given that cross-modal distractors reduced children's accuracy in Experiment 2, it is possible that cross-modal interference also occurs at later response selection stages in childhood. However, given that this was not clearly evidenced in Experiment 1 these conclusions remain speculative, and warrant further investigation. Nevertheless, these findings suggest differential processing of unimodal and cross-modal distractors in childhood, and suggest children are more susceptible to cross-modal distraction. Support for this is seen in the differences in response time distribution between unimodal and cross-modal distractors in children (Experiment 1), different developmental trajectories for unimodal and cross-modal distraction (Experiment 1), and heightened cross-modal interference (increased slowing and reduced accuracy) in childhood (Experiment 2).

Limitations and Future Directions

The current study provides important findings in an under-represented area of literature. However, considering some limitations may guide future research. Firstly, within this study I focused on age groups in which multisensory and interference control processes are known to be immature, below 11 years (Noel et al., 2016), and susceptible to age-related decline, above 64 years (Comalli et al., 1962; Noel et al., 2016). However, it has been shown that multisensory integration processes continue to mature across adolescents until around 17 years of age (Innes-Brown et al., 2011; Noel et al., 2016) and temporal binding windows progressively increase between the ages of 50 and 64 (Noel et al., 2016). Furthermore, asymmetries in unimodal stimulus- and response-interference have been reported between adolescents (who manifest more response-interference) and middle-aged adults (who show more stimulus-interference; Killikelly & Szűcs, 2013). Investigating these age groups may thus provide insight into the complete lifespan trajectories of unimodal and cross-modal interference.

Secondly, future research should aim to optimise comparisons between sensory conditions. This study implemented a variant of the established De Houwer (2003) paradigm that has been utilised in developmental (Killikelly & Szűcs, 2010) and ageing (Killikelly & Szűcs, 2013) contexts to separate stimulus and response-interference. Using this variant enabled comparison of unimodal and cross-modal conditions in which visual information was relevant (Experiment 1) and compare cross-modal conditions in which visual and auditory information was relevant (Experiment 2). Furthermore, through using a colour patch version of the Stroop task prevented confounding unimodal and cross-modal conditions with integrated vs. separate Stroop tasks (Macleod, 1998). Nevertheless, this design had some limitations. Firstly, in Experiment 1 there was no unimodal auditory condition. It is recognised that “fully crossed” paradigms are under-represented in the literature (Guerreiro et al., 2010; Van Gerven & Guerreiro, 2016) and future research should implement such designs to enable full comparison between unimodal and cross-modal interference control. Secondly, in Experiment 2, the relevant visual dimension was a colour, whilst the relevant auditory dimension was a word. Stroop interference has been attributed to imbalanced ease of access to colour vs. word information, whereby colour naming is more difficult than word reading (Melara & Algom, 2003; Roelofs, 2005). It is therefore possible that the “ignore visual” (colour) condition was easier than the “ignore auditory” (word) condition. Given this, the possibility cannot be excluded that older adults were able to perform the task due to the ease of access to spoken word information. This might also be one explanation of why the current findings contradict past research, suggesting older adults struggle to ignore vision (Van Gerven & Guerreiro, 2016). However, it would still be expected that focusing on audition whilst ignoring vision would be difficult in older adults, given age-related increases in listening effort (Tun et al., 2009), reduced dual task performance (Gagné et al., 2017; Verhaeghen, Steitz, Sliwinski, & Cerella, 2003), reduced speech in noise comprehension (Getzmann et al., 2015), and shifts in sensory dominance in favour of vision (Diaconescu et al., 2013). Nevertheless, future research should be able to decipher whether ease of word access may account for maintained ability in ageing by equating the type of visual and auditory stimuli in cross-modal tasks.

Conclusions

The findings from this study suggest that the ability to ignore distraction within and across senses undergo different lifespan trajectories. These findings form empirical support for the theory that cross-modal interference is less susceptible to age-related decline (Guerreiro et al., 2010; Van Gerven & Guerreiro, 2016) but extend this to show that older adults may be able to ignore audition whilst focusing on vision and vice versa in cross-modal Stroop tasks. Conversely children appear more susceptible to both unimodal and cross-modal Stroop interference compared with adults. This might be because children process cross-modal distractions differently, however, this warrants further investigation.

Chapter 4: Do Colour Vision Deficits Increase Stroop Interference in Ageing?

In Chapter 3 older adults were impaired only for the unimodal Stroop task. This Chapter addresses whether some age-related effects may be attributed to a loss of colour vision. Using the unimodal Stroop condition from Chapter 3, alongside a low saturation unimodal task, I attempted to "simulate" ageing in young adults. Furthermore, using event-related potentials (ERPs) alongside the Stroop paradigm used in Chapter 3, I assessed whether sensory deficits specifically increase stimulus-interference. ERPs associated with stimulus-interference (N450 and stimulus-locked lateralised readiness potentials; LRPs) and response-interference (response-locked LRPs) were compared. It was predicted that colour saturation would specifically increase stimulus-interference and this would be evidenced in larger amplitude of the N450 and stimulus-locked LRPs. Conversely it was predicted that colour saturation would not affect response-locked LRPs, which have been associated with response-interference. In contrast to the hypotheses, behavioural data showed similar interference effects between high and low saturation conditions, and N450 and stimulus-locked LRP amplitudes did not differ between saturation conditions. However, response-locked LRPs to stimulus-incongruent trials were significantly lower in amplitude (indicating less preparation of the correct response) under low saturation vs. high saturation conditions. These findings are discussed in relation to task differences between the current study and previous research showing effects of colour saturation on Stroop performance.

Introduction

Research has consistently shown that older adults manifest larger Stroop effects than younger adults (Cohn, Dustman, & Bradford, 1984; Comalli et al., 1962; Davidson, Zacks, & Williams, 2003). Some theories attribute this ageing effect to reduced frontal lobe function (Milham et al., 2002; West, 1996b; West & Bell, 1997) and subsequent inhibitory control deficits (Hasher & Zacks, 1988; West & Alain, 2000). However, other theories attribute these effects to general slowing (Verhaeghen & De Meersman, 1998) and sensory differences (Anstey et al., 2002; Ben-David & Schneider, 2009, 2010). In contrast to inhibitory deficit theories, these perspectives suggest cognitive ability (assessed via the Stroop task) in older adults has been underestimated, and peripheral mechanisms, such as sensory change, may influence Stroop performance with age.

An imbalanced ease of access to word and colour information, *dimensional imbalance*, has been proposed to drive Stroop interference – as easier access to the word vs. the colour results in more interference (Melara & Algom, 2003). Age-related reductions in colour vision have therefore been proposed to increase Stroop interference. In a meta-analysis of 13 published studies Ben-David and Schneider (2009) found an increase in the speed of word reading relative to colour naming with age, indicating increased dimensional imbalance. In line with this, Anstey et al (2002) report colour vision as the single largest predictor of Stroop performance with age. Ben-David and Schneider (2010) also found that reducing the saturation of colour in Stroop stimuli resulted in performance decrements in younger adults comparable to those seen in older adults. Together, these findings suggest a sensory origin of colour-word Stroop effects in ageing.

The findings presented in Chapter 3 suggested increases in both stimulus- and response-interference with old age. If colour quality is considered a perceptual property, it follows that reduced colour vision may contribute specifically towards increased stimulus-interference in age. However, whether this also biases response competition is less clear. The study presented in this Chapter uses the same paradigm implemented in

Chapter 3 to assess whether limiting colour quality could specifically increase stimulus-interference and not response-interference in a group of young adults. This behavioural design was combined with an event-related-potential (ERP) approach, enabling detection of neural processes that may not be detectable at the behavioural level. For example, it is possible that compensatory strategies might be employed to support similar behavioural performance in young adults, and this could be inferred using EEG although might not be as clear in behavioural findings. A second reason for using ERPs in the current Chapter was to establish protocol to be used in later experiments (Chapter 5 and 6) in which the neural correlates of unimodal and cross-modal interference were to be compared.

This study, and ERP studies throughout this thesis, focused on ERPs that have previously been associated with stimulus-interference (the N450 effect and stimulus-locked lateralised readiness potentials; LRPs) and response-interference (response-locked LRPs). These components are outlined below.

The N450 effect

The N450 effect, also known as medial frontal negativity MFN (Chen, Bailey, Tiernan, & West, 2011), N^{Inc} (Donohue, Appelbaum, Mckay, & Woldorff, 2016) and Ni (Kray, Eppinger, & Mecklinger, 2005), is a lower amplitude potential occurring under incongruent vs. congruent conditions around 450ms post-stimulus presentation, resulting in a negative (incongruent – congruent) difference wave (Appelbaum, Meyerhoff, & Woldorff, 2009; Augustinova, Silvert, Ferrand, Llorca, & Flaudias, 2014; Coderre, Conklin, & Van Heuven, 2011; Donohue, Todisco, et al., 2013; Killikelly & Szűcs, 2013; Larson, Kaufman, & Perlstein, 2009; Lorist & Jolij, 2012; Markela-Lerenc et al., 2004; Naylor, Stanley, & Wicha, 2012; Szűcs & Soltész, 2007, 2010, 2012; K. Wang, 2011; West, Jakubek, Wymbs, Perry, & Moore, 2005). This ERP has been localised to the anterior cingulate cortex (ACC) and attributed to general conflict detection (Milham et al., 2001; Swick & Jovanovic, 2002).

Studies have found the N450 effect is sensitive to dimensional imbalance. For example, Kray et al (2005) found prolonged N450 effects for a colour naming version of the Stroop task versus a word naming version. Further to this, many studies have shown the N450 effect is not increased by response conflict, suggesting the N450 is associated with stimulus-interference (Killikelly & Szűcs, 2013; Szűcs & Soltész, 2012; West, Bowry, & McConville, 2004; however see Chen et al., 2011). For example, using a numerical version of the Stroop task, West et al (2004) found that the N450 did not significantly differ between conditions in which incongruent stimuli were in the response set, compared to when they were not in the response set. Furthermore, they suggest that the ACC, the proposed source of the N450 response, may detect semantic conflict but not response conflict. In a more recent study, Szűcs and Soltész (2012) sought to explicitly identify whether the N450 was linked with stimulus- or response-interference. They manipulated the amount of response conflict that was present in each trial through either pre-cuing the correct response hand (removing response interference) or not pre-cuing the correct hand. It was found that the N450 effect did not differ between pre-cued and not pre-cued trials. As such, Szűcs and Soltész (2012) define the N450 as reflecting stimulus-interference and not response interference. Given that the N450 effect has been linked to stimulus-interference, and that it appears sensitive to dimensional imbalance (Kray et al., 2005), it was expected that *a larger N450 effect should be seen under low saturation vs. high saturation Stroop conditions*, reflecting increased dimensional imbalance and subsequent stimulus-interference.

Stimulus- and Response-locked Lateralised Readiness Potential (LRP)

The second group of ERPs focused on in this thesis are lateralised readiness potentials (LRPs). LRPs are an index of response preparation calculated through comparing evoked responses over the left and right motor cortices. LRPs are calculated as (Coles, 1989):

Equation 6

$$LRP = [(ER\text{-}EL)\text{left hand response} + (EL\text{-}ER)\text{right hand response}]/2$$

In which ER is the amplitude of the Evoked potential over the Right motor cortex and EL is the amplitude over the Left motor cortex. A more negative deflection therefore indicates preparation of the correct response.

Two types of LRPs have been measured in the literature. *Stimulus-locked LRPs* are time-locked to stimulus presentation whilst *response-locked LRPs* are time-locked to the response. These components have been associated with stimulus processing and response preparation, respectively. For example, Smulders, Kok, Kenemans, and Bashore (1995) used a choice response time task in which participants made either a left or right hand response. In this study, stimulus quality (i.e. clear vs. degraded) and response complexity (i.e. one vs. three buttons) were manipulated. The two factors had additive effects on response times. Furthermore, stimulus-degradation affected the latency of the stimulus-locked LRP but not the response-locked LRP, whilst response complexity increased only the response-locked LRP latency. This is in line with the findings of Masaki, Wild-Wall, Sangals, and Sommer (2004) who found that response-locked LRPs were modulated by control over response velocity whilst stimulus-locked LRPs were not. Furthermore, Killikelly and Szűcs (2013) found stimulus-locked LRP amplitude did not significantly differ between stimulus- and response-interference conditions. Together, these findings suggest that the stimulus-locked LRP is associated with stimulus encoding whilst the response-locked LRP is linked to response processing. Given the similarity between the manipulations in Smulders et al (1995) and the current study (i.e. limiting stimulus quality), It was predicted that *lowering saturation would affect (lower) the amplitude of stimulus-locked LRPs and not response-locked LRPs.*

Increased Stray-light and Colour Saturation in Ageing

A second way in which this study advanced previous research was selecting saturation level based on an existing model of visual ageing. Ben-David and Schneider (2010) “simulated” ageing in young adults through limiting colour saturation. However, in this study, an arbitrary level of saturation was selected. Ben-David and Schneider (2010) highlighted the variable nature of colour vision changes with age (Nguyen-Tri, Overbury, & Faubert, 2003; Werner & Steele, 1988) and recognised that reducing colour saturation by an arbitrary amount cannot precisely emulate the colour vision deficits seen in older adults. As such, I attempted to select a more biologically plausible level of saturation to simulate ageing Stroop effects.

Encoding of hue, brightness and saturation all deteriorate between 25 and 75 years of age (Clay Smith, 1943; Cooper et al., 1991; Okajima, Tsuchiya, & Yamashita, 2002). The detection of short wavelength light is particularly affected by ageing (Kraft & Werner, 1999), mapping onto a loss in short wavelength cones in the retina (Haegerstrom-Portnoy, Hewlett, & Barr, 1989). Nevertheless, the most commonly reported visual deficit seen in ageing is increased glare perception, resulting from thickening of the lens and increased opacity of intraocular fluid (Hennelly, Barbur, Edgar, & Woodward, 1998; Van Den Berg, 1995). These changes reduce the efficiency by which light is focused onto the retina, resulting in increased *stray-light* and, consequently, glare. This glare "washes out" colour saturation. Thus, although colour vision changes in many ways with age, lowering saturation is a plausible approach to simulating ageing. However, the level of saturation selected could be based on models of change in stray-light (glare) with age.

Hennelly et al (1998) defined the increase in stray-light with age using the stay-light parameter, k':

Equation 7

$$k' = 6.6 + 0.09exp\ (a^{0.627677} - 8)$$

Where k' represents the integral of the scatter function between $\theta = 2.2^\circ$ and infinity and a represents age (illustrated in Figure 15a). Using this model, I aimed to more closely emulate the effect of stray-light on saturation perception with age. To implement the model, the opacity of a white desaturation mask overlaying onscreen stimuli was manipulated. To approximately simulate the stray-light seen in a 70-year-old participant (the average age of older adults in Chapter 3) the opacity of the mask was set to 60% (indicated by the cross-hair in Figure 15b). Critically, I must emphasise that changes in lens opacity and intraocular glare are a few of many dimensions of colour vision that alter with age. Nevertheless, this model provides a more objective approach for selecting the level of saturation at which colours were presented.

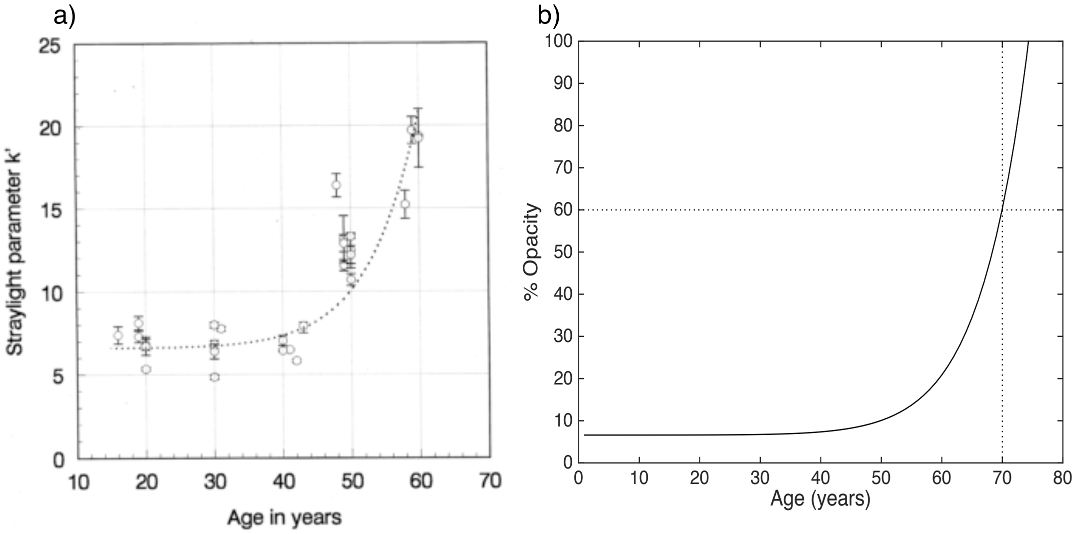


Figure 15. a) Figure taken from Hennelly et al (1998) and b) implementation of the model to the opacity of a desaturation mask. The cross-hair in b marks the masker opacity selected based on the mean age of participants in the experiments of Chapter 3.

The Current Study

In sum, this study aimed to assess whether reducing colour saturation increased Stroop interference, and whether this takes effect at stimulus-encoding stages. Young adults were presented with the unimodal Stroop task outlined in Chapter 3, alongside a low saturation version of the same task (in which the level of saturation was selected based on the model of Hennelly et al (1998)). It was predicted that, if stimulus-interference was greater in low saturation conditions but response-interference was not affected by saturation:

1) Response time and accuracy ratios should show more stimulus-interference effects but no differences in response interference effects under low saturation compared with high saturation conditions.
2) Low saturation conditions would result in larger N450 amplitude and lower stimulus-locked LRP amplitude (because more negative LRPs indicate stronger preparation of correct response).
3) The amplitude of the response-locked LRP should not differ between high and low saturation conditions.

Method

Participants

Thirty-two[9] young adults (Mean age = 24.2 years; range = 19-30 years; 17 female) were recruited. Participants were staff and students of the University of Nottingham. All participants self-reported being right handed, having normal or corrected to normal vision and hearing, and being fluent in English (29 had English first language, 1 Spanish, 1 Icelandic, 1 Dutch, 1 German). Due to a software crash, one participant's EEG data was not usable (final n = 31).

---

[9] An a priori power analysis was not run for this study. However this sample size is consistent with that used in Chapter 3 and therefore should be powered for a similar number of comparisons. To aid interpretation of null findings, and identify where statistical comparisons may have been underpowered Bayes factors indicating support for the null (i.e. BF_{01}) are provided alongside frequentist statistics.

Equipment

Visual stimuli were presented via a Mac Pro on a Samsung 23” S23A700D monitor (resolution 1820x1080 @ 120 Hz) at a viewing distance of ~57cm maintained via a chinrest. Auditory stimuli were presented via Senheiser HMD280 PRO headphones (for calibration and stimulus presentation methods see Chapter 2). EEG data were collected using a Macbook Pro using Actiview acquisition software. Triggers were set using a Cedrus StimTracker.

EEG data were recorded in an electrically shielded Faraday cage using a 64-channel BioSemi Active Two electrode array at a sampling rate of 2024Hz (later down-sampled to 511Hz). Electrodes placed below the right eye and on the right temple were used to measure vertical and horizontal eye movements, respectively. Recordings were referenced online using a right mastoid reference (later re-referenced to an average).

Stimuli

The stimuli used throughout all sections of the procedure in this study were identical to those outlined in Chapter 3. However, a white “desaturation mask” was presented over the entire screen to create a “low saturation” Stroop condition. During the Stroop task the opacity of this mask was set to 60% (the opacity level selected to simulate stray-light in a 70 year old observer – Figure 15). Although individual thresholds for how visible colours were behind the mask were obtained for each participant (these are presented in Analysis of Threshold data section of the appendix, page 315) the opacity of the mask was the same across participants. Due to luminance differences between the monitor used in this study and that used in Chapter 3, the luminance of the colour rectangles were; red = 54.2cd/m^2, green = 177cd/m^2, blue = 20.83cd/m^2, yellow = 212.75cd/ m^2.

Procedure

The procedure of the current study consisted of three parts, a threshold task, a response time task and a Stroop task.

Threshold task

The protocol for deriving thresholds was identical to the staircase procedure outlined in Chapter 1 Experiment 1 (adult participants). However, two additional sets of staircases were added to derive thresholds for reading words behind the white desaturation mask and for naming colours behind the white desaturation mask (whilst the mask varied in opacity – see Threshold task section of the appendix, page 309). The former of these was so that I could attempt to present written words 10x above threshold in both high saturation and low saturation conditions. The second of these was so that I could assess how visible the colour was relative to the word for each individual. As this second analysis is not the focus of this Chapter the visibility of colour information will not be discussed further. However, in brief, the results of this analysis showed that the detectability of colour information relative to word information was significantly higher under high saturation vs. low saturation conditions.

Response time task

The response time task was used to assess participants' speed for naming colours relative to reading words. Thus, I assessed whether lowering colour saturation increased the dimensional imbalance between colour vs. word accessibility. This was also secondary to the main goals of this study, therefore details of this task procedure and analysis are detailed in the appendix (pages 310 and 313 respectively). In short, this analysis showed response times did not differ between colour naming and written word naming.

Stroop task

The Stroop task implemented is illustrated in Figure 16. This task was identical to the unimodal task used in Chapter 3 with the following adjustments:

1. The cross-modal condition was replaced with a low saturation unimodal condition in which a desaturation mask occluded all onscreen stimuli.
2. There were more trials (288 per saturation condition; 72 per congruency condition) to allow more accurate measurement of ERPs.
3. Visual babble was used instead of the word “Brown” on neutral trials[10].
4. The trial structure was altered to reflect that used in previous ERP studies separating stimulus and response-interference (Killikelly & Szűcs, 2013). This included increasing the duration of stimulus presentation and adding “blank” periods following stimulus offset to allow time for movement artifacts to subside[11].

---

[10] Babble was used on neutral trials instead of the word Brown because in the first set of Experiments (Chapter 3) the word “Brown” resulted in stimulus-interference.

[11] During piloting the brief presentation of a fixation cross was intended to tell participants when to blink, thus minimising artifacts in the ERPs overlapping stimulus presentation (as was the case in Killikelly & Szűcs, 2013). However this instruction was then not implemented to keep the task simple.

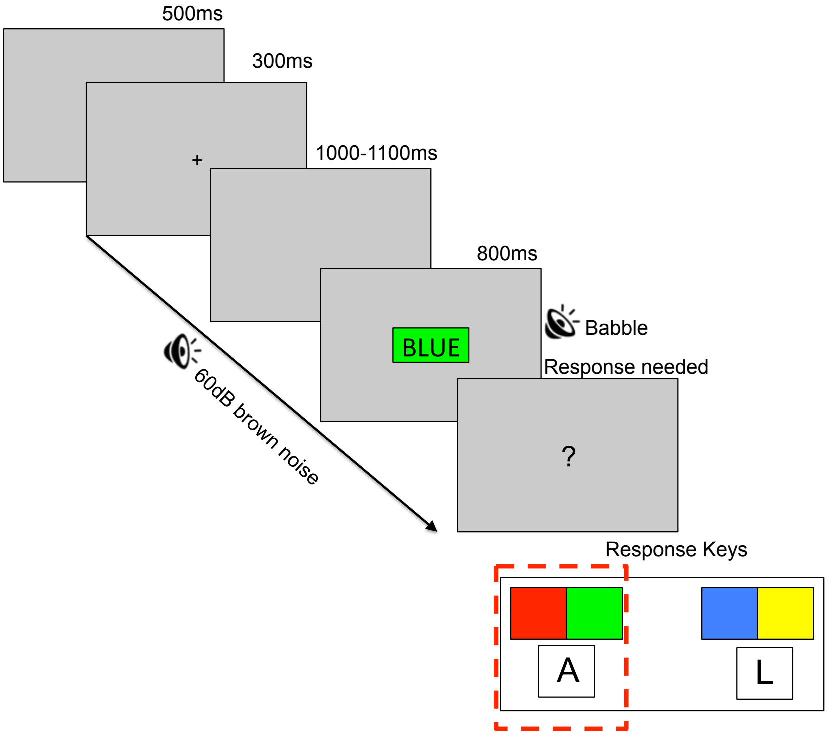


Figure 16. Design of Stroop protocol. Participants identified the colour of a rectangle using two response keys. Participants could respond as soon as they saw the rectangle however if a response had not been made following stimulus offset a question mark was presented signalling the need for a response.

Analysis

The analysis of response times and accuracy was identical to the approach used in Chapter 3. No participants were identified as outliers, however, one participant was removed due to a technical difficulty with EEG measurement (final n = 31). In Chapter 3, I discussed findings with *p* values <=.01 as significant. This was to remain conservative given the exploratory nature of the study and the large number of comparisons to discuss. However, in the current chapter, and the remaining chapters in this thesis, I discuss *p* values <=.05 as significant. This was firstly due to the increasingly directional nature of hypotheses, secondly due to the fewer number of comparisons to discuss and thirdly, for EEG chapters, to maintain consistency in reporting of significance between behavioural and EEG data (for which a more liberal alpha was used to avoid Type 2 errors – further detail provided in the discussion of ERP analysis, page 96).

EEG pre-processing

The pre-processing pipeline used is illustrated in Figure 17. EEG data were re-referenced to an average and pre-processed offline using Fully Automated Statistical Thresholding for EEG artefact Rejection (FASTER) in MATLAB (Nolan, Whelan, & Reilly, 2010). A 95Hz low-pass filter, 1Hz high-pass filter and 50Hz notch filter was applied. Bad channels were identified, rejected and interpolated if the channel's variability, mean correlation with other channels and Hurst exponent fell more than 3 standard deviations (SDs) away from the average across channels (excluding EOG channels) across all time-points[12]. These parameters therefore, respectively, isolated channels that were highly variable, non-correlated and presented extreme trends, or no trend, relative to other channels.

Epochs were time-locked to stimulus onset (-200-800ms) and baseline corrected -200-0ms pre stimulus onset. Epochs were rejected if the trial was identified as an outlier or incorrect response in the behavioural data. Epochs containing artifacts were identified and automatically rejected if the amplitude range and variability of the epoch was more than 3 SDs from the average across epochs, and if the average amplitude of one channel within that epoch deviated more than 3 SDs from the mean amplitude of that channel across epochs. Outlier removal resulted in 2.95-20.76% of trials being removed (*M* = 11.51%, *SD* = 4.31%). Epoch rejection during pre-processing resulted in 1.29-4.77% of trials being removed (*M* = 2.95%, *SD* = 0.86%). Thus, 76.47-93.86% of all epochs were *accepted* (*M* = 85.54%, *SD* = 4.4%).

Independent component analysis (ICA) was used to isolate and reject components associated with artifacts. Artefacts associated with eye movements were identified if a component's correlation with the EOG channels was more than 3 SDs away from other component's correlations with the EOG channels. Artifacts occurring at single channels (suggesting electrode "pop-off") were identified if a component was spatially focused on a single electrode relative to other components (i.e. the spatial kurtosis was

---

[12] The Hurst exponent is a value ranging from 0-1 indicating long-range dependence (i.e. tendency to promote trends) within a signal (Nolan et al., 2010).

more than 3 SDs away from the mean spatial kurtosis of components). Finally, components associated with residual white noise (i.e. electrical interference) were identified based on the slope of the component's power spectra, and the Hurst exponent and median gradient of the component's time-course (more than 3 SDs from average).

It should be noted that response-locked LRPs were not subjected to ICA and component rejection. As shown in the preprocessing pipeline in Figure 17, FASTER performs ICA on epoched data. In the current study data were epoched -200 – 800ms relative to stimulus presentation prior to ICA and component rejection. However, response-locked LRPs required a different time-window, -800 – 200ms around the *response*, rather than around stimulus-presentation. As such, the pre-ICA data was used to "re-epoch" data relative to responses rather than stimulus-presentation for each participant. Response-locked LRPs were then calculated using this epoched data. A limitation of this approach is that response-locked LRPs may have been contaminated by artifactual components rejected during ICA. However, the main purpose of ICA was to identify eye-blink components that have highest influence on frontal channels. As LRPs were calculated using channels C3 and C4 it was judged that these channels would be less influenced by eye blinks, thus justifying the use of pre-ICA data. Notably, this limitation is addressed using a different approach in the following EEG chapters (Chapter 5).

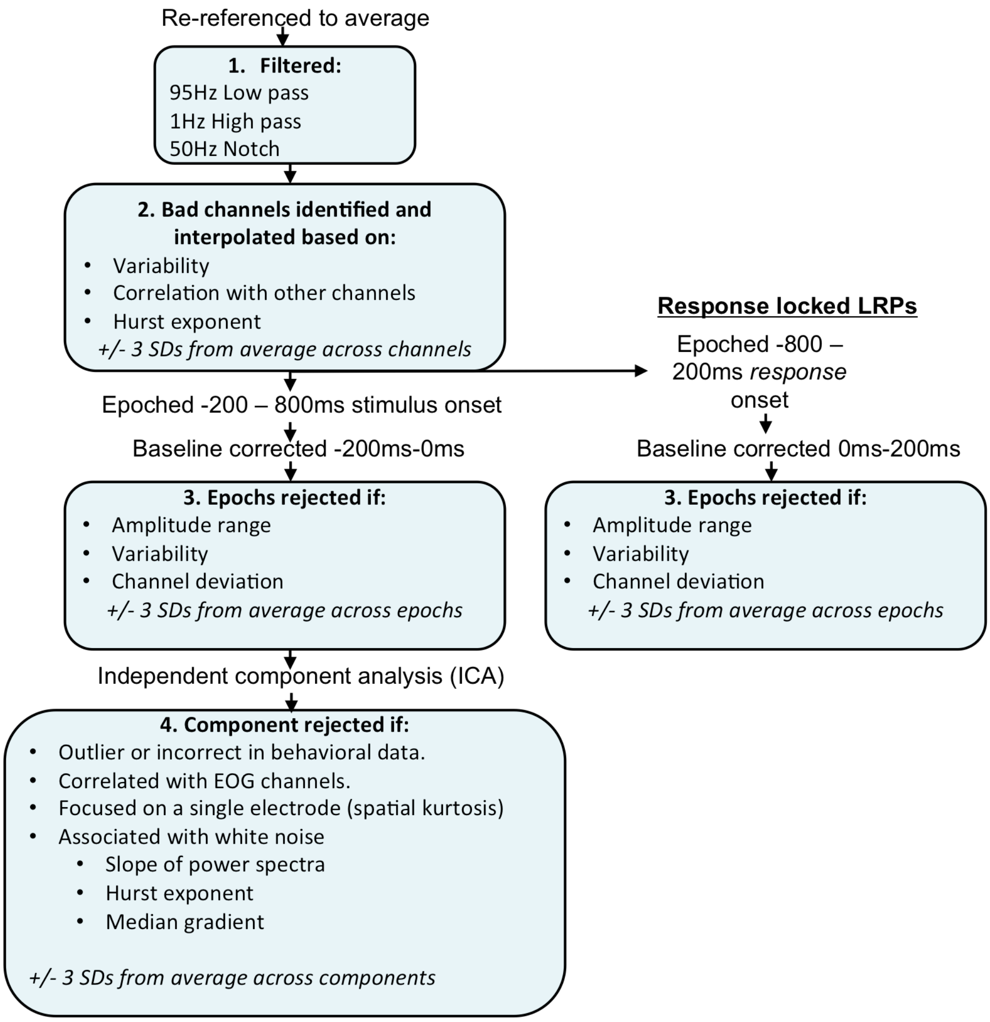


Figure 17. EEG pre-processing pipeline used to reject bad channels and identify artifacts. This pipeline was implemented automatically with FASTER (Nolen et al., 2010) in MATLAB. Because response-locked LRPs were calculated using a different time-window, and using channels C3/C4, which were considered less susceptible to eye-blinks, these LRPs did not undergo ICA.

ERP analysis

Although the N450 effect is a commonly used marker of conflict detection, the time window and location at which the N450 has been studied is heterogeneous. Studies have used time windows as early as 250ms to as late as 550ms, and reported effects over frontal, central and parietal electrodes. As such, I used an exploratory statistical approach to isolate components of interest – Mass Univariate Analysis.

The Mass Univariate approach compares all time-points at all electrodes to identify locations and time-points of interest (Groppe, Urbach, & Kutas, 2011). To correct for multiple comparisons, a t-max permutation test is used. Here a series of t-tests identify the most extreme t-value across electrodes at time-point *x* (t^{Max}). A specified number of permutations[13] are then used to derive new t^{Max} values. The corrected *p*-value then indicates the number of times the new t^{Max} values exceeded the original. For example, if 1 in 8 permutations derived a more extreme t^{Max}, *p* = .125 because there is a 1/8 probability of a more extreme *t* value occurring when the null hypothesis is true. This exploratory technique does not depend upon a priori knowledge regarding the electrodes and time-points of interest. As such, this approach was deemed appropriate for identifying the time window and electrodes of interest in which to compare ERPs.

Mass univariate analyses were implemented using the Mass Univariate ERP toolbox in MATLAB (Groppe et al., 2011). As the extent of multiple corrections implemented by the mass univariate approach may entail a loss of statistical power, and a risk of Type 2 errors, *p* values of .05 were considered as significant (in comparison to the conservative *p* value of .01 used in Chapter 3). To further limit the number of comparisons, the time window of 0-800ms post stimulus onset was used to find effects associated with general interference (i.e. differences between response-incongruent and congruent) collapsed across sensory conditions, and then further comparisons (i.e. saturation x congruency ANOVAs) were conducted using the mean amplitude of responses within the time window of interest at the identified electrodes of interest. Notably, as the N450 effect is a difference wave this made identification of individual peaks in the effect challenging. As such, the current analysis and other ERP analyses in this thesis did not investigate peak latency effects and, rather, focused on amplitude comparisons. Notably, several other ERP components have been studied in relation to the Stroop effect. As such this initial exploratory approach allowed me to investigate whether any other effects of interest occurred within the 0-800ms post stimulus onset time-window.

---

[13] In which the voltage of values in randomly selected participants are switched, equivalent to randomly switching conditions A and B in standard permutation testing.

Lateralised Readiness Potentials (LRPs) were calculated as outlined in Equation 7. Stimulus-locked LRPs were compared 250 to 600ms post stimulus onset and response-locked LRPs were compared -300 to 0ms preceeding response initiation. These time windows were selected based on the time-windows used in previous literature (Killikelly & Szűcs, 2013). As LRPs are calculated using only two electrodes (over motor cortices) the outlined explorotory mass-univariate approach was not necessary. Therefore the mean amplitude of LRPs within the specified time windows were compared using an ANOVA.

Results

Stroop performance

As in Chapter 3 the effects of general interference are reported first followed by stimulus- and response-interference[14]. The following was predicted:

1. That general interference would be higher under low vs. high saturation conditions.
2. That stimulus-interference would be higher under low vs. high saturation conditions.
3. That response-interference would not differ between the two saturation conditions.

---

[14] As in Chapter 3, neutral conditions produced a pattern of effects similar to stimulus-interference. Analysis of raw response times and accuracy are therefore within the appendix (page 322).

General interference

Figure 18 shows general interference ratios for response times and accuracy. Response time ratios were not significantly different between high and low saturation conditions ($F(1, 30) = .492$, $p = .488$, $\eta_p^2 = .02$, $BF^{01} = 3.023$).

Accuracy ratios also were not significantly different between high and low saturation conditions ($F(1, 30) = .448$, $p = .509$, $\eta_p^2 = .02$, $BF^{01} = 3.148$).

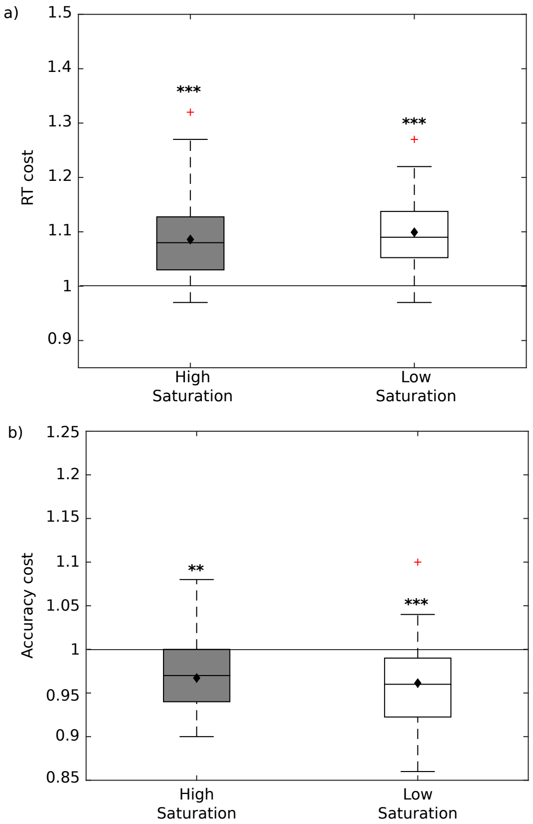


Figure 18. a) Response time (RT) and b) accuracy ratios for high (grey) and low (white) saturation conditions. General interference ratios indicate the ratio of response time/accuracy for response-incongruent vs. congruent conditions. Thus a score higher than 1, shown by continuous horizontal line, reflects slower/more accurate performance and a score lower than 1 reflects faster/less accurate performance. Black diamonds indicate means used for analyses. Asterisks indicate t-tests comparing ratio to 1 (* = .05, ** = .01, ***= .001) Bonferroni corrected for 6 comparisons.

Stimulus- and Response-interference

Figure 19 shows stimulus and response interference ratios under high and low saturation conditions in terms of response times (a) and accuracy (b).

Response times

A 2 (interference type) x 2 (saturation) ANOVA indicated no main effect of saturation (*F*(1, 30) = .553, *p* = .463, η_p^2 = .02, BF_{01} = 4.403) or interference type (*F*(1, 30) = 1.912, *p* = .177, η_p^2 = .06, BF_{01} = 1.255) and no interaction between interference type and saturation (*F*(1, 30) = .195, *p* = .662, η_p^2 = .01, BF_{01} = 3.74). Thus, lowering colour saturation did not specifically increase interference at the stimulus level.

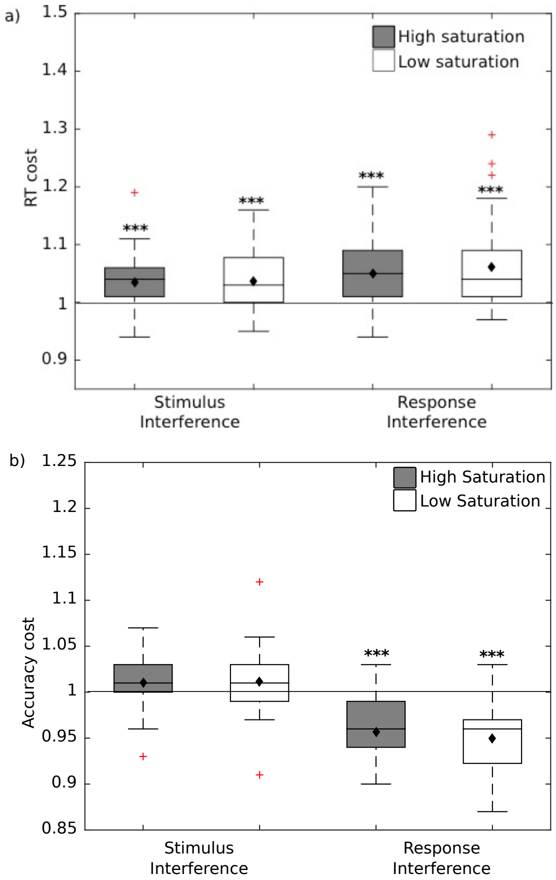


Figure 19. a) Response time (RT) and b) Accuracy ratios showing stimulus *and* response-interference ratios under high (grey) and low (white) saturation conditions. Stimulus-interference indicates response time/accuracy under stimulus-incongruent vs. congruent conditions. Response-interference ratios indicates time/accuracy for response-incongruent vs. stimulus-incongruent conditions. Scores higher than 1, shown by continuous horizontal line, reflects slower/more accurate performance and a score lower than 1 reflects faster/less accurate performance. Black diamonds indicate means used for analyses. Asterisks indicate t-tests comparing ratio to 1 (* = .05, ** = .01, ***= .001) Bonferroni corrected for 6 comparisons.

Accuracy

A 2 (sensory condition) x 2 (interference type) ANOVA showed a main effect of interference type (*F*(1, 30) = 57.387, *p* < .001, η_p^2 = .66, BF_{01}<.001). Response interference but not stimulus-interference reduced accuracy. There was no main effect of saturation (*F*(1,30) = .318, *p* = .577, η_p^2 = .01, BF_{01} = 4.966) and no interaction between saturation and interference type (*F*(1,30) = .427, *p* = .519, η_p^2 = .01, BF_{01} = 3.75).

EEG

The N450 effect

As previously outlined, a mass univariate approach was used to first identify electrodes and time points of interest for the N450 effect (collapsed across saturation conditions), in which the ERP to response-incongruent stimuli should be lower in amplitude than the ERP to congruent stimuli. This analysis showed a significant difference between congruent and response-incongruent ERPs 430-495ms post stimulus presentation over parietal electrodes (P1, P5, and P03; Figure 21; all corrected *p*-values 0.0428 - 0.0108; corrected t^{Max} -scores -5.29 - 5.29) in which lower amplitude ERPs occurred under response-incongruent vs. congruent conditions. This was accompanied by an effect of opposing polarity at frontal electrode F6 at 510ms post stimulus presentation. Together, this was identified as the N450 effect. The topography of this effect under high and low saturation conditions is shown in Figure 21. Individual ERPs for each electrode, used for the subsequent ANOVA, are shown in Figure 20. Notably, no other significant effects were found in the 0-800ms time window. As such, I will focus only on the N450 effect and LRP effects.

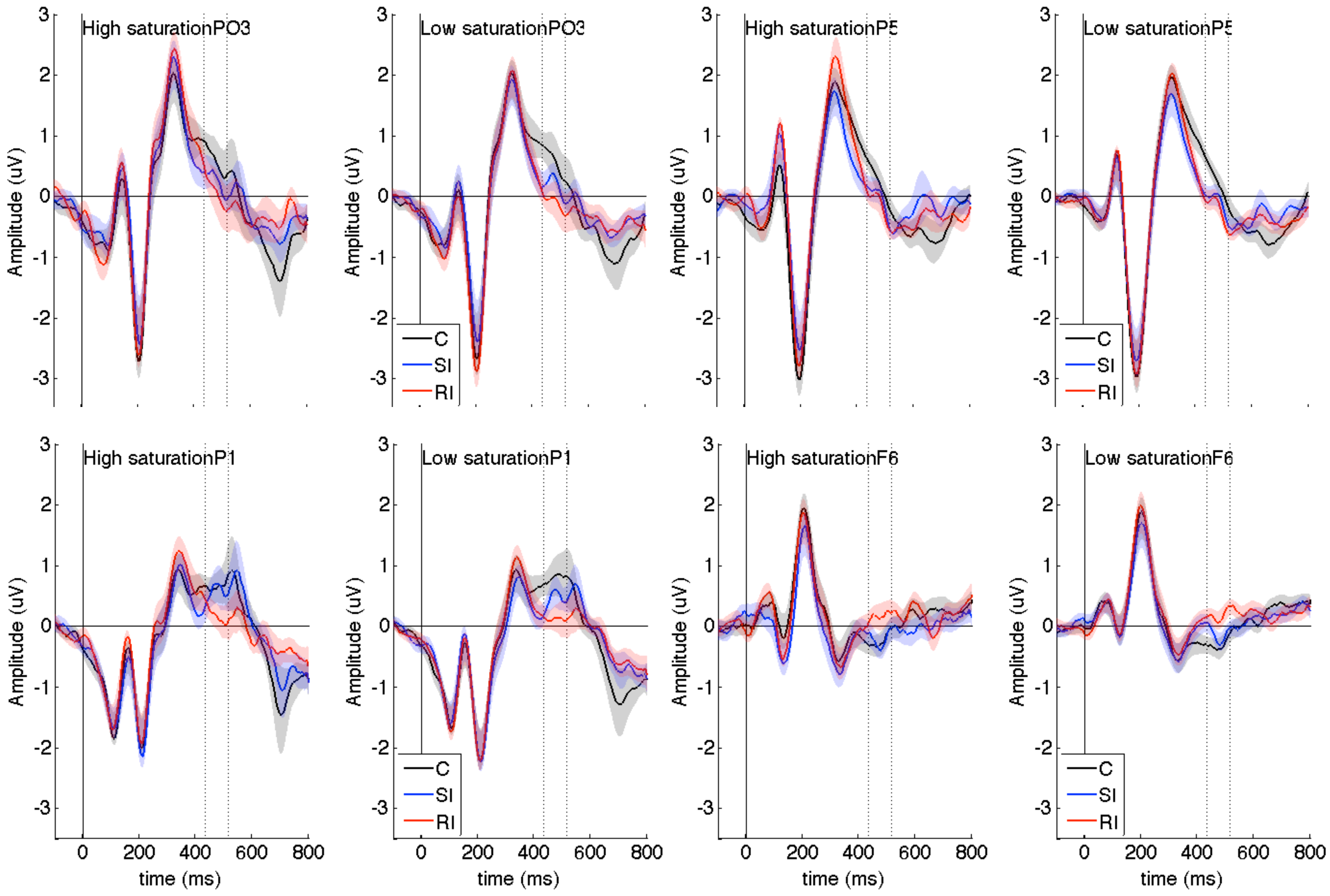


Figure 20. ERPs in response to congruent (black), stimulus-incongruent (blue) and response-incongruent (red) conditions under high and low saturation. ERPs are shown at PO3, P5, P1 and F6. C = congruent, SI = stimulus-incongruent RI = response-incongruent. Shaded regions indicate the standard error of the mean at each time point.

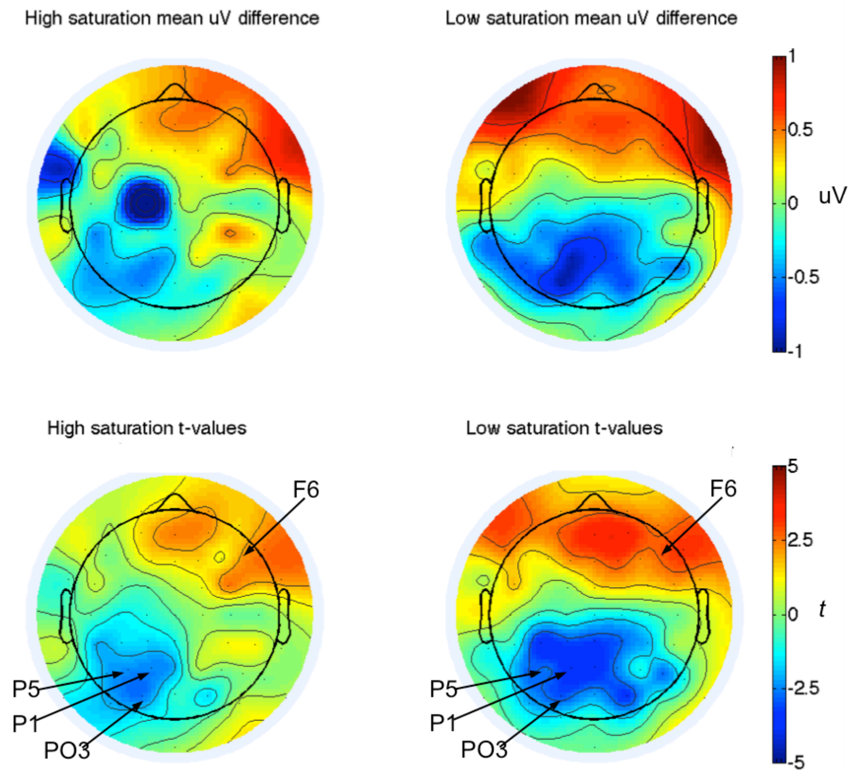


Figure 21. Topographical distribution of mean voltage (upper) and t-score (lower) for the response-incongruent – congruent difference under high saturation (a and c) and low saturation (b and d) conditions between 430 and 495ms. Electrodes showing significant effects are highlighted.

The identified time window and electrodes for the N450 effect was then used to guide comparisons of saturation and interference type using a 2 (saturation) x 3 (congruency) x 4 (electrode) ANOVA to compare mean amplitude within the time window of 430-510ms.

The results of the ANOVA showed a main effect of congruency ($F(1.33, 39.8) = 10.88$, $p = .001$, $\eta_p^2 = .204$, $BF_{01} = .077$). There was a significant difference in mean amplitude between congruent vs. stimulus-incongruent ($p = .008$, $BF_{01} = .157$) and congruent vs. response-incongruent ($p < .001$, $BF_{01} < .001$). Critical to the current assumption that the N450 reflects stimulus interference, there was no significant difference between stimulus-incongruent and response-incongruent ERPs ($p = .157$, $BF_{01} = 2.18$). This supports the assumption that the N450 reflects stimulus interference, as response incongruence did not further modulate the effect. However, it should be noted that the Bayes factor suggests only anecdotal support for the null (i.e. $BF_{01} < 3$).

In contrast to the hypothesis that saturation would increase stimulus interference, and therefore N450 amplitude, there was no interaction between congruency and saturation (*F*(2, 60) = 1.599, *p* = .211, η_p^2 = .05, BF_{01} = 15.58). This suggests that lowering saturation did not alter the amplitude of the N450 effect.

There were several other null and one significant finding that are not relevant to the current hypotheses. In brief; there was no main effect of saturation overall (*F*(1.00, 30.00) = .87, *p* = .357, η_p^2 = .03, BF_{01} = 8.624), or electrode (*F*(18.99, 49.27) = 2.68, *p* = .088, η_p^2 = .08, BF_{01} <.001), no interaction between electrode and saturation (*F*(2.08, 62.52) = .532, *p* = .597, η_p^2 = .02, BF_{01} = 59), and no three-way interaction between electrode, saturation and congruency (*F*(1.21, 0.59) = 2.08, *p* = .109, η_p^2 = .07, BF_{01} = 37.26). There was a significant interaction between electrode and congruency (*F*(2.63, 78.98) = 10.6, *p* <.001, η_p^2 = .26, BF_{01} <.001). However, this was due to inverse polarity over frontal vs. parietal locations. As this was not of primary interest to the current hypotheses this will not be discussed further.

In sum, the current findings support the assumption that the N450 effect was associated with stimulus-interference (because there was no significant difference between stimulus-incongruent and response-incongruent ERPs). However, the amplitude of this effect was not modulated by colour saturation. Thus, lowering colour saturation did not result in any effects in N450 amplitude.

Stimulus-locked LRPs

A second marker proposed to represent stimulus-interference is the stimulus-locked LRP. The current hypothesis was that the stimulus-locked LRP should be significantly smaller for incongruent stimuli under low saturation conditions, indicating lower preparation of the correct response due to stimulus-interference.

Stimulus-locked LRPs under low saturation and high saturation conditions are shown in Figure 22. A 2 (saturation) x 3 (congruency) ANOVA showed no main effect of saturation ($F(1, 30) = 1.654$, $p = .208$, $\eta_p^2 = .05$, $BF_{01} = 4.156$), no main effect of congruency ($F(21.66, 49.73) = 1.195$, $p = .310$, $\eta_p^2 = .04$, $BF_{01} = 5.317$) and no interaction between saturation and congruency ($F(2, 60) = 1.208$, $p = .306$, $\eta_p^2 = .04$, $BF_{01} = 3.39$). Firstly, this suggests that saturation had no main effect on stimulus-locked LRP effects. Furthermore, the findings from this comparison suggest congruency effects were not evidenced in the stimulus-locked LRP. This is contrary to the assumption that the stimulus-locked LRP is sensitive to stimulus-interference.

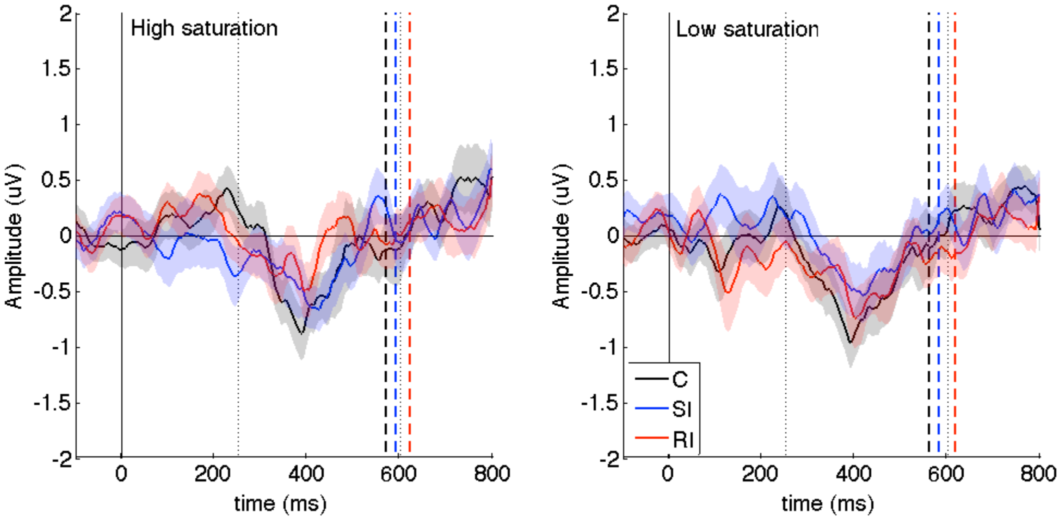


Figure 22. Stimulus locked LRPs calculated using electrodes C3 and C4 under high saturation and low saturation conditions. LRPs are shown for congruent (C), stimulus-incongruent (SI) and response-incongruent (RI) conditions. Vertical dashed line show average response time for congruent (black), stimulus-incongruent (blue) and response-incongruent (red) trials relative to the stimulus-locked LRPs.

Response-locked LRPs

A final analysis performed was comparison of response-locked LRPs – proposed to reflect response-interference. It was hypothesised that this LRP should show lower amplitude under response-incongruent vs. congruent conditions (showing the effect is sensitive to response-interference). Furthermore, it was hypothesised that this would not differ between saturation conditions (if saturation is linked with stimulus-interference and not response-interference). Response-locked LRPs under high and low saturation conditions are shown in Figure 23.

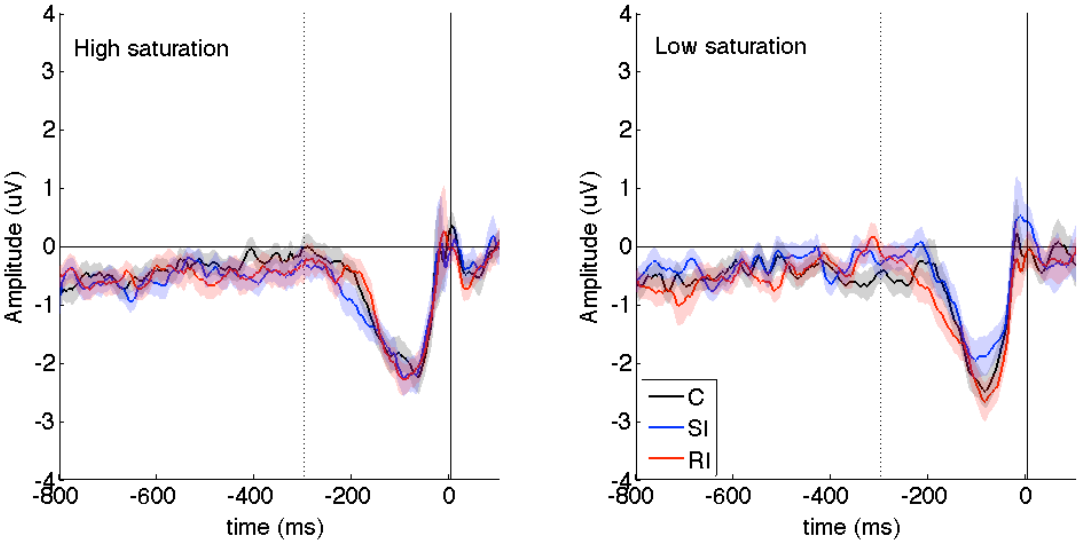


Figure 23. Response-locked LRPs under high saturation (left) and low saturation (right) conditions. C = congruent, SI = stimulus-incongruent, RI = response-incongruent. Shaded areas indicate the standard error of the mean at each time point. 0ms indicates time point where response was made. Horizontal dashed line indicates time window used for comparison (-300 to 0ms).

A 2 (saturation) x 3 (congruency) ANOVA showed no main effect of saturation ($F(1, 30) = .07$, $p = .79$, $\eta_p^2 = .002$, $BF_{01} = 5.948$) and no main effect of congruency ($F(1, 30) = 1.03$, $p = .363$, $\eta_p^2 = .033$, $BF_{01} = 9.705$) but an interaction between saturation and congruency ($F(1, 30) = 3.73$, $p = .03$, $\eta_p^2 = .11$, $BF_{01} = 1.246$). Simple main effects analyses showed that this interaction occurred because under stimulus-incongruent conditions the amplitude of the response-locked LRP was significantly smaller under low saturation conditions compared with high saturation conditions ($p = .035$). Contrary to the hypothesis that the response-locked LRP would be sensitive to response-interference, there was no difference in amplitude between congruency conditions within high and low saturation conditions respectively.

Discussion

The current study aimed to assess whether reducing colour saturation increased Stroop interference (Ben-David & Schneider, 2010), and whether this takes effect at stimulus-encoding stages. The findings were as follows. General interference did not differ between saturation conditions, and there was also no difference in stimulus-and response-interference between conditions. Following this, there was also no difference in the amplitude of the N450 effect and stimulus-locked LRP between saturation conditions. However, there was an unexpected finding in the response-locked LRPs. Specifically, response-locked LRPs on stimulus-incongruent trials were smaller under low saturation vs. high saturation conditions. In this section, I will first discuss potential explanations for the lack of differences between high and low saturation conditions. Following this, I will discuss the implications of the ERP and LRP effects seen in this study.

In contrast to Ben-David and Schneider (2010), the current findings show no behavioural differences in Stroop performance between high and low saturation conditions. Furthermore, there was no difference in either stimulus- or response-interference between saturation conditions (in terms of behaviour or ERPs). This suggests colour saturation does not affect Stroop interference and does not specifically increase stimulus-interference. One explanation of this finding is that colours were not “washed out” enough to directly replicate the effects seen in Ben-David and Schneider (2010). Supplementary analyses of thresholds (see appendix page 315) showed that colours were more detectable than words in both saturation conditions, although this difference was significantly smaller for low saturation conditions (suggesting a shift in dimensional imbalance). However, analyses also showed no difference in speed of word reading relative to colour naming for high or low saturation colours. As such, lowering colour saturation did not appear to affect accessibility of colour and word information in terms of response times. However, as these measures (thresholds and response times) were not obtained in Ben-David and Schneider's (2010) study it is difficult to conclude that this explains the null findings in this study.

Another explanation of this finding might be that the current task was easier, and more susceptible to practice effects than that used by Ben-David and Schneider (2010). Ben-David and Schneider used a standard, colour-word, Stroop paradigm, in which the participants identified the font colour of a written word whilst ignoring the word meaning. Conversely, the study in this Chapter used a colour-patch Stroop. Thus, the amount of colour information was greater and arguably more salient (perhaps also explaining why colour information was more accessible as previously described). Furthermore, colour and word information could be more easily segregated as opposed to being considered features of the same, integrated, object. It has been shown that practice results in substantial improvement in “segregated” Stroop tasks compared with integrated colour-word Stroop tasks (Flowers & Stoup, 1977; MacLeod, 1991). As this study used 576 (2 x 288) trials compared with the 36 trials used by Ben-David and Schneider, it is likely that practice effects may have lessened interference effects in this study, subsequently masking differences between saturation conditions.

However, other design features in this study should have increased interference in this study relative to previous research. First, Ben-David and Schneider used a Stroop task in which trials were presented within blocks of the same congruency, thus congruency effects should have been decreased through *the Gratton effect,* in which congruency effects are lower when a trial is preceded by the same congruency condition (Gratton, Coles, & Donchin, 1992). In contrast to this, the current study presented congruency in a randomised order and, although more incongruent trials were presented, many of these were actually informative of the correct response (i.e. stimulus-incongruent), thus, participants should have been more likely to attend to distracting information. In support of the claim that practice effects should not have affected this study, a supplementary analysis assessing behavioural performance for the first (96 trial) block of each saturation condition only still showed no differences in behavioural performance between saturation conditions. These findings support the conclusion that null findings in this study may not have resulted from practice effects, and therefore saturation does not affect interference in the Stroop task. Furthermore, considering the Bayes factors reported in this study, it does not appear that null findings resulted from underpowered comparisons.

Nevertheless, several findings in the EEG data do suggest differences between saturation conditions. First, the topographical distribution of the N450 effect (Figure 21) showed more widespread effects in the low saturation compared with the high saturation condition. This suggests a wider spread distribution of neural activity under low saturation vs. high saturation conditions. This could reflect either increased interference effects or compensatory neural activity to produce similar behaviour despite increased task difficulty. However, differences in topography were not statistically tested therefore at present this interpretation remains qualitative and speculative.

A second, unanticipated, effect was that the response-locked LRP to stimulus-incongruent trials showed smaller amplitude under low saturation conditions. This suggests more stimulus-interference under low saturation vs. high saturation conditions. Furthermore, this suggests that the response-locked LRP may also be sensitive to stimulus interference. However, as the difference between congruent and stimulus-incongruent response-locked LRPs was not significant in either condition this conclusion remains tentative. The response-locked LRP has been associated with response-interference rather than stimulus-interference (Killikelly & Szűcs, 2013; Masaki et al., 2004; Smulders et al., 1995). Therefore, it would be expected that response-incongruent conditions should manifest lower amplitude, even positive, responses. However in the current findings response-locked LRPs showed similar, negative, morphology across congruency conditions (apart from the lower amplitude seen on low saturation conditions). A likely cause of this is that only correct responses were included in the ERP/LRP analyses (and too few incorrect responses were made for an adequate analysis of these trials), thus all potentials would show negative deflection.

Although secondary to the primary aims of this study, the current findings support the suggestion that the N450 effect reflects stimulus-interference (Szűcs & Soltész, 2012). This is because there was a significant difference between stimulus-incongruent and congruent conditions in the N450 time window, however, there was no significant difference between stimulus-incongruent and response-incongruent ERPs within this time window. Given this finding, the N450 is considered to reflect stimulus-interference in the following chapters. In contrast to this, no congruency effects were evident in either the stimulus-locked or response-locked LRPs. This is surprising, given that LRPs have been used as markers of incongruency and response conflict in previous literature (Killikelly & Szűcs, 2013; Li et al., 2017; Masaki et al., 2004; Szűcs, Soltész, Bryce, & Whitebread, 2009) . As such although the experiment reported in the following Chapter did also focus on LRPs, the limitations of using LRPs as markers of stimulus- and response-conflict are to be discussed in Chapter 5.

Conclusions

In contrast to previous findings, reducing colour saturation did not increase interference in the current Stroop paradigm. Furthermore, there were no increases in either stimulus- or response-interference and no differences in N450 or stimulus-locked LRP amplitude (proposed to reflect stimulus interference). However, the topography of N450 effects did appear more widespread under low saturation conditions. Furthermore, the response-locked LRP initiated by stimulus-incongruent trials was significantly smaller under low saturation vs. high saturation Stroop conditions. Although these might be indicative of discrete differences in neural processes between high and low saturation, this conclusion remains tentative. Finally, the N450 was associated with stimulus-interference, however, congruency effects were not clear in either the stimulus-or response-locked LRPs in this study. These components will be explored and discussed further in Chapter 5.

Chapter 5: ERP Markers of Unimodal and Cross-modal Stroop Effects

Chapter 3 suggested different neural mechanisms underlie unimodal and cross-modal Stroop interference. Unimodal interference occurred at stimulus encoding and response selection stages. Cross-modal interference occurred mainly from stimulus-interference and appeared resistant to ageing effects. In this chapter, this hypothesis is tested further using the event-related-potential (ERP) protocol established in Chapter 4.

ERPs associated with stimulus-interference (N450 and stimulus-locked lateralised readiness potentials, LRPs) and response-interference (response-locked LRPs) were compared under unimodal and cross-modal conditions. It was predicted that although unimodal and cross-modal effects would show similar amplitude of stimulus-interference components, response-locked LRPs would only show congruency effects under unimodal conditions. Behavioural findings supported the theory that unimodal but not cross-modal interference arises from response conflict. In contrast to Chapter 4, the N450 effect was associated with response interference, and therefore appeared larger under unimodal conditions (although this failed to reach significance). In line with the hypothesis that cross-modal interference arises from stimulus-interference, the stimulus-locked LRP showed more interference under cross-modal conditions. Contrary to hypotheses, but in line with Chapter 4, the response-locked LRP did not show congruency effects in either sensory condition.

Introduction

Although multiple studies have investigated the neural correlates of ignoring audition whilst focusing on vision (Townsend, Adamo, & Haist, 2006) and vice versa (Čeponiene, Westerfield, Torki, & Townsend, 2008), a direct comparison of the time-course of unimodal and cross-modal Stroop interference is yet to be conducted. The findings presented in Chapter 3 of this thesis suggested the time-course of unimodal and cross-modal Stroop interference differ. Cross-modal effects may occur at earlier, stimulus-encoding, stages, whilst unimodal interference also arises from response competition.

Additional evidence suggests unimodal and cross-modal interference occur at different stages. Elliott et al, (2014) mapped the response time distribution of unimodal and cross-modal Stroop interference. Whilst unimodal interference increased with response latency, cross-modal interference did not. Response times under cross-modal conditions were also approximately 100 ms faster than under unimodal conditions. One explanation of these findings is that cross-modal interference control occurs earlier in processing, enabling participants to “block out” distractors prior to response selection. As such, although Stroop interference occurs at longer latencies under unimodal conditions, this would not be the case for cross-modal conditions. However, in Chapter 3 of this thesis unimodal and cross-modal interference manifested similar response time distributions in young adults. Furthermore, response times overall under unimodal and cross-modal conditions appeared comparable. Given this, we cannot rule out a similar time-course of effects between unimodal and cross-modal interference control.

The finding that cross-modal interference control is less susceptible to ageing supports the idea that unimodal and cross-modal interference have different neural loci. Guerreiro et al. (2010) propose that unimodal interference control relies on cortical mechanisms known to be susceptible to age-related decline such as the prefrontal cortex (Milham et al., 2002; Prakash et al., 2009; West, 1996a). Conversely, cross-modal auditory distraction may be suppressed by peripheral mechanisms less susceptible to ageing. As discussed in the introduction to this thesis, evidence shows that irrelevant auditory information can be gated at the level of the cochlea, prior to actual cortical processing (Giard, Collet, Bouchet, & Pernier, 1994). EEG studies using the paired-click-paradigm, in which the auditory evoked response to a second of two clicks is smaller due to gating, also suggest irrelevant auditory information is gated as early as 50 ms post stimulus onset (Jones, Hills, Dick, Jones, & Bright, 2016) and that this gating mechanism is maintained in ageing (Gmehlin, Kreisel, Bachmann, Weisbrod, & Thomas, 2011). Together, these findings support the hypothesis that cross-modal interference-control may take effect earlier in processing and this may be why cross-modal interference control is maintained in ageing.

The aim of the current study was to identify whether unimodal and cross-modal Stroop effects manifest similar time-courses using event-related potentials (ERPs). As in Chapter 4, this study focused on ERPs associated with stimulus-interference (the N450 effect and stimulus locked lateralised readiness potentials; LRPs) and response-interference (response-locked LRPs). Based on the prediction that cross-modal interference arises at stimulus-encoding stages and not response-selection stages. The following predictions were made:

1) The N450 would show similar amplitude between stimulus- and response-interference and between unimodal and cross-modal conditions (thus showing stimulus-interference in both conditions).
2) The stimulus-locked LRP would show similar amplitude between stimulus- and response-interference and between unimodal and cross-modal conditions (again showing stimulus-interference in both conditions).
3) The response-locked LRP would show lower amplitude on response-incongruent vs. congruent and stimulus-incongruent trials, and this effect would occur under unimodal but not cross-modal conditions (thus showing less correct hand preparation under response-incongruent trials, i.e. response-interference, but only under unimodal conditions).

Method

Participants

Thirty-two participants, 19 female, 29 right-handed (aged 18-32 years, *M* =

23.6 years) were recruited from the Macquarie University[15] volunteer register. Twenty-five[16] participants reported English as a first language (remaining participants reported English as a second language; mean age of English language acquisition was 8.57 years, range 6-11 years). Participants were paid $30 (Australian dollars) for their participation. The data for this study were gathered as part of a larger study validating a portable EEG system (EMOTIV Epoc+) and fast periodic visual stimulation for measuring cognitive control[17]. Participants therefore took part in two tasks: a Stroop task and passive target detection task whilst wearing a “research-grade” Neuroscan EEG overlaid by a “gaming” Emotiv EEG system. The focus of this Chapter is the comparison of unimodal vs. cross-modal Stroop ERPs only using the research-grade EEG (Neuroscan) recordings. Comparisons with EMOTIV are presented in Chapter 6.

Equipment

The experimental task was programmed and presented in PsychoPy version 1.85.2 (Peirce, 2007, 2009). Visual stimuli were presented via an Optoplex 910 desktop running Windows 7 on an A0C G2770 monitor (1920x1080, refresh rate 144Hz, viewing distance ~57cm). Auditory stimuli were presented via Phillips SHS4700/37 ear-clip headphones (for calibration details see general methods, Chapter 2 page 27).

EEG data were recorded via a 32-channel Neuroscan system (SynAmps2) using Curry software. Data were sampled at 1000Hz with an online band-pass filter of 1 to 100Hz, referenced online to the left ear lobe and later re-referenced to an average. Vertical eye movements were recorded with electrodes above and below the left eye. Horizontal eye movements were measured using electrodes placed at the outer canthi of each eye. The ground electrode was positioned between FPz and Fz.

---

[15] This study was conducted as part of an overseas institutional visit to Macquarie University, Sydney, Australia, in conjunction with data reported in Chapter 6.

[16] As in Chapter 4, sample size was selected in accordance with Chapter 3 and Bayes factors are reported in the results section to indicate comparisons that may have been underpowered.

[17] Further details of this project are available via open science framework https://osf.io/xbk9c/ and https://osf.io/8cdqa/

Stimuli

The stimuli were identical to those used in Chapter 3, Experiment 1, with the following exceptions:

1) Babble stimuli were presented instead of the word “Brown” on neutral trials.
2) Due to luminance differences in the monitor the luminance of the colour rectangles were; red=61.25cd/m^2, green=208.77cd/m^2, blue=18.61cd/m^2 and yellow=243.8cd/m^2
3) Continuous Brown noise was presented at 55dB[18].

Procedure

The procedure consisted of a threshold task, to identify thresholds for reading and hearing colour-words, followed by a Stroop task. The threshold task was identical to the method-of-adjustment used in Chapter 3. The Stroop task was identical to the Stroop task used in Chapter 3, Experiment 1. However, the timing structure of each trial and the number of trials per condition followed that used in Chapter 4 (allowing time for movement artifacts to subside following each response), with the exception that the duration of the stimulus was 482ms rather than 800ms (allowing the timing of auditory and visual stimuli to be matched) and the conditions were unimodal and cross-modal (rather than two unimodal conditions).

In the Stroop task, written and spoken information were presented 10x (20dB above) each individual’s threshold.

Analysis

Behavioural data were analysed using the same methods outlined in Chapter 3. Outlier removal resulted in the removal of 3 (of 32) participants. Analyses of threshold data and raw response time and accuracy scores are available within the appendix (pages 320 and 323). As outlined in Chapter 4 (page 93)

---

[18] Due to a calibration error, all auditory stimuli were presented 5dB below the expected value. Thus, Brown noise was 55dB rather than the 60dB used in other chapters. This also meant auditory stimuli were capped at maximum for all participants. Supplementary analysis exploring whether this influenced unimodal vs. cross-modal Stroop interference are therefore provided in the appendix (page 321).

I discuss *p* values <=.05 as significant. Furthermore, I report Bayes factors to evaluate support for the null hypothesis alongside frequentist statistics.

EEG pre-processing

The pre-processing pipeline used for EEG data is illustrated in Figure 24. Data were down-sampled to 500Hz, band-pass filtered from 1-95Hz and re-referenced to an average of all electrodes (excluding eye channels). Bad channels and noise-burst artifacts were first identified using the “clean_rawdata” plugin for EEGLAB which uses the Artifact Subspace Reconstruction (ASR) method (summarised in Figure 24 boxes 2, 3, and 4; Miyakoshi & Kothe, 2014). Following this, independent component analysis (ICA) was used to separate components associated with physiological artifacts (i.e. eye blinks). Artifactual components were then identified and removed using an automated classifier MARA (Winkler, Haufe, & Tangermann, 2011). MARA is trained based on the frequency, spatial and temporal characteristics of components manually labelled by experts as being either artifactual or brain sources. This approach has been reported to achieve low (8.9%) classification errors (Winkler et al., 2011).

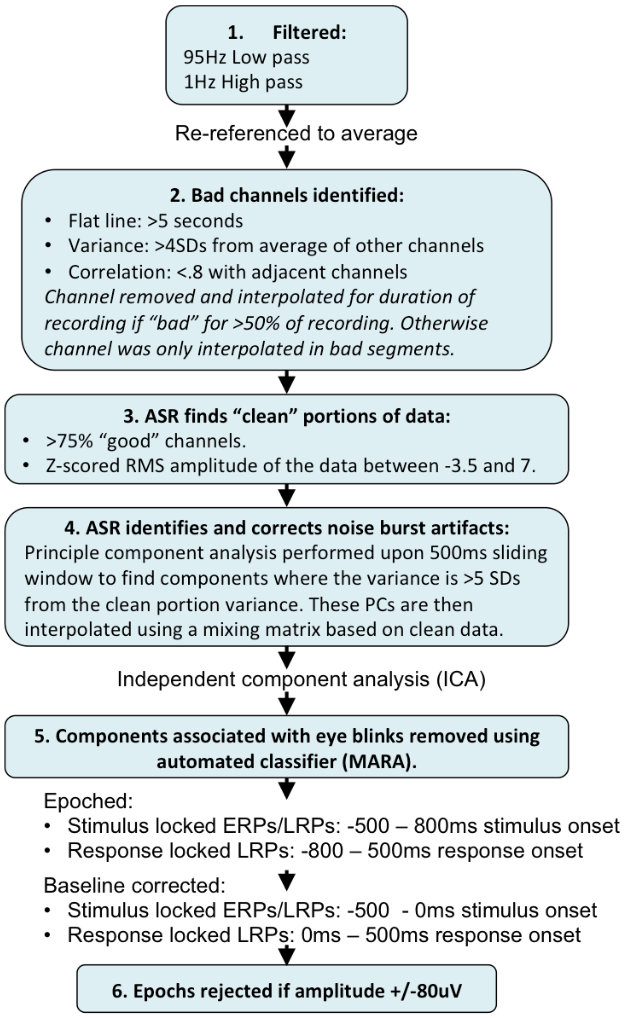
Following pre-processing, data were epoched -500 to 800ms relative to stimulus onset and baseline corrected from -500 to 0ms. Epochs were removed if they contained a voltage -/+80uV. Participants with fewer than 50% trials remaining per congruency condition (in line with Killikelly & Szűcs, 2013) following artefact rejection were not included in analyses. This resulted in the removal of 1 participant. Following removal of this participant, the remaining participants had minimum of 55.55% of trials remaining per condition (unimodal congruent *M* = 86.95% range 66.66-98.61%; unimodal stimulus-incongruent *M* = 86.55% range 70.83-100%; unimodal response-incongruent *M* = 81.4% range 55.55-95.83%; cross-modal congruent *M* = 89.78% range 73.61-100%; cross-modal stimulus-incongruent *M* = 90.33% range 73.61-100%; cross-modal response-incongruent *M* = 88.79% range 70.83-98.61%). This left 28 participants for ERP analysis.

Notably, in contrast to Chapter 4, response-locked LRPs followed the same pre-processing pipeline with the exception that data were epoched -

800 to 500ms pre-post response. Thus, response-locked LRPs in this Chapter were subject to the same ICA and component rejection procedure.

ERP analysis

As in Chapter 4, a mass univariate analysis was used to identify electrodes showing the N450 effect. To minimise the number of comparisons, this analysis identified electrodes showing a significant mean difference between response-incongruent and congruent ERPs (collapsed across sensory conditions) 400-500ms[19] post stimulus presentation (rather than including every time-point in the analysis). The identified electrodes were then used for



1. Filtered:
95Hz Low pass
1Hz High pass

Re-referenced to average

2. Bad channels identified:
- Flat line: >5 seconds
- Variance: >4SDs from average of other channels
- Correlation: <.8 with adjacent channels

Channel removed and interpolated for duration of recording if "bad" for >50% of recording. Otherwise channel was only interpolated in bad segments.

3. ASR finds "clean" portions of data:
- >75% "good" channels.
- Z-scored RMS amplitude of the data between -3.5 and 7.

4. ASR identifies and corrects noise burst artifacts:
Principle component analysis performed upon 500ms sliding window to find components where the variance is >5 SDs from the clean portion variance. These PCs are then interpolated using a mixing matrix based on clean data.

Independent component analysis (ICA)

5. Components associated with eye blinks removed using automated classifier (MARA).

Epoched:
- Stimulus locked ERPs/LRPs: -500 – 800ms stimulus onset
- Response locked LRPs: -800 – 500ms response onset

Baseline corrected:
- Stimulus locked ERPs/LRPs: -500 - 0ms stimulus onset
- Response locked LRPs: 0ms – 500ms response onset

6. Epochs rejected if amplitude +/-80uV

Figure 24. EEG pre-processing pipeline used for Experiments presented in Chapter 5. At the channel interpolation stage (Box 2) only 15 participants had channels removed completely. The number of channels removed ranged from 1 – 3, *M* = 1.53.

---

[19] The selected 400-500ms time-window was selected based on the observed effects in Chapter 4.

further comparisons of mean amlitude using an ANOVA. LRPs were calculated and compared using the same approach outlined in Chapter 4.

Results

Stroop performance

General interference

As shown in Table 5 and Figure 25 no significant differences were found between unimodal and cross-modal interference ratios for either response-time or accuracy. However, as shown in Table 5 general interference was higher under unimodal conditions (i.e. causing more slowing and reduced accuracy) and Bayes factors did not provide strong support for the null.

		Sensory condition					
		Unimodal ***M (SD)***	**Cross-modal** ***M (SD)***	***F***	***p***	η_p^2	BF_{01}
Response times							
Interference type	General	1.06 (.01)	1.04 (.08)	2.597	.118	.09	1.35
	SI	1.01 (.04)	1.02 (.05)	1.694	.204	.06	1.78
	RI	1.05 (.05)	1.01 (.04)	11.99	.002**	.30	0.04
Accuracy							
Interference type	General	0.97 (.05)	0.98 (.05)	0.908	.349	.03	2.47
	SI	1.01 (.03)	1.00 (.04)	1.186	.286	.04	2.12
	RI	0.96 (.04)	0.98 (.05)	4.746	.038 *	.15	0.52

* *p* <= .05, ** *p* <= .01, *** *p* <= .001

Table 5. Statistics comparing unimodal and cross-modal interference types. BF_{01} = Bayes factor showing support for the null. Shaded rows highlight significant differences.

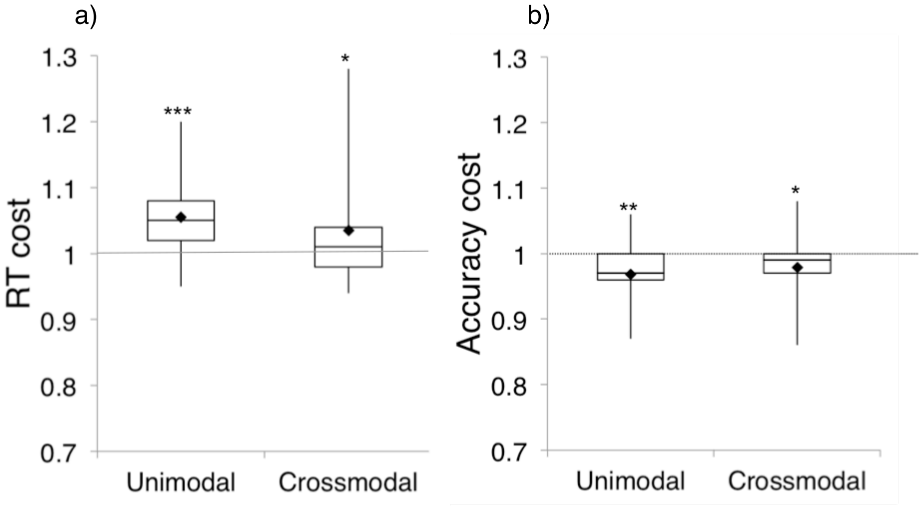


Figure 25. a) Response time (RT) and b) accuracy costs for unimodal and cross-modal conditions. RT ratios higher than 1 indicate slowing and accuracy ratios lower than 1 indicate reduced accuracy for response-incongruent relative to congruent trials. Whiskers = 1^{st} and 4th quartiles; Boxes = 2^{nd} and 3^{rd} quartiles; Central line = median; Black diamonds = means. Asterisks show significance of Bonferroni corrected one-sample t tests comparing ratios to 1; * *p* <= .05, ** *p* <= .01, *** *p* <= .001.

Stimulus and Response Interference

Response times

As shown in Figure 26a and Table 5, a 2(sensory condition; unimodal vs. cross-modal) x 2 (interference type; stimulus-interference vs. response-interference) ANOVA showed no main effect of sensory condition (*F(*1, 28) = 2.975, *p* = .096, η_p^2 = .096, BF_{01} = 2.685) and no main effect of interference type (*F(*1, 28) = 1.694, *p* = .204, η_p^2 = .057, BF_{01} = 1.855), but a significant interaction between sensory condition and interference type (*F(*1, 28) = 8.0, *p* = .009, η_p^2 =.222, BF_{01} = .072). As shown in Table 5 RI caused more slowing under unimodal compared with cross-modal conditions. Furthermore, under

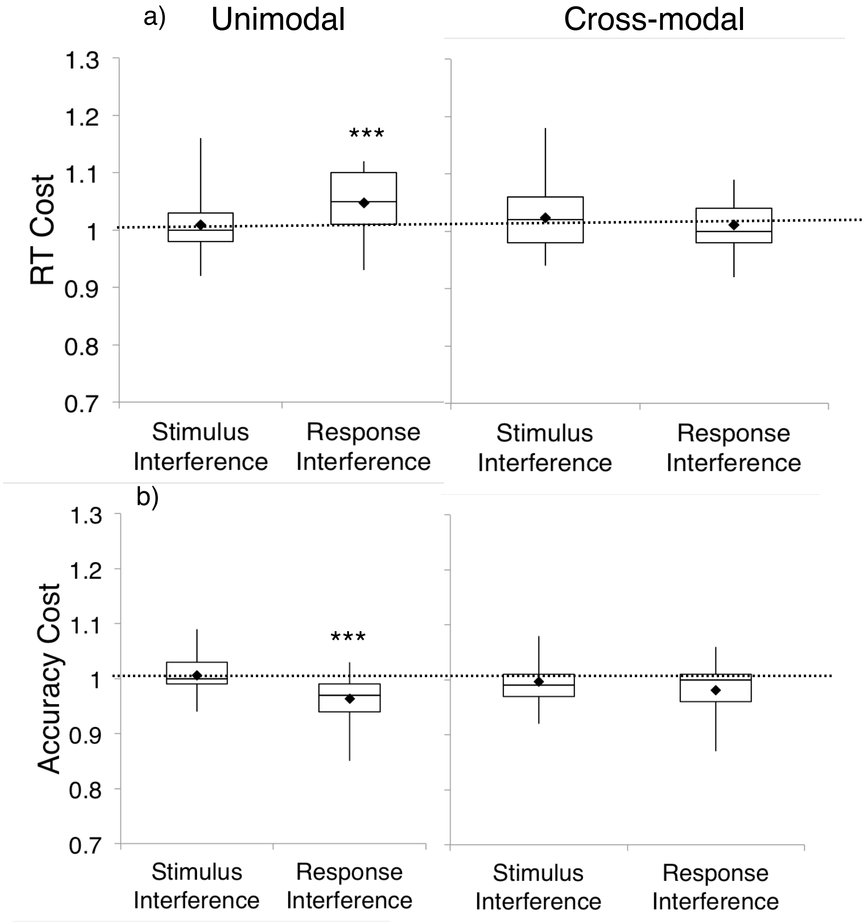


Figure 26. Response time (RT) (a) and accuracy (b) costs for unimodal (left) and cross-modal (right) conditions caused by stimulus-interference and response-interference. Scores higher 1 indicate slowing. Whiskers = 1st and 4th quartiles; Boxes = 2nd and 3rd quartiles; Central line = median; Black diamonds = means; open circles = individual participants data points. Asterisks show significance of Bonferroni corrected one-sample t tests comparing ratios to 1; *** <.001; ** <=.01, *=<.05.

unimodal conditions response-interference was higher than stimulus-interference (*p* = .013, BF_{01} = .066); however, there was no significant difference between stimulus- and response-interference under cross-modal conditions (*p* = .328, BF_{01} = 2.362).

Accuracy

As shown in Figure 26b and Table 5, a 2(sensory condition; unimodal vs. cross-modal) x 2 (interference type; stimulus-interference vs. response-interference) ANOVA showed no main effect of sensory condition (*F(*1, 28) = .865, *p* = .36, η_p^2 = .03, BF_{01} = 4.094) but a main effect of interference type (*F(*1, 28) = 9.71, *p* = .004, η_p^2 = .26, BF_{01} =.003); response-interference (*M* = .972 *SE* = .007) reduced accuracy more than stimulus-interference (*M* = 1.001, *SE* = .005). The effect of interference type interacted with sensory condition (*F(*1, 28) = 4.849, *p* = .036, η_p^2 =.15, BF_{01} = .583). Response-interference reduced accuracy more than stimulus-interference under unimodal conditions (*p* < .001, BF_{01} <.001) but not cross-modal conditions (*p* = .249, BF_{01} =1.612) conditions. Response-interference but not stimulus-interference was also higher (i.e. reduced accuracy more) under unimodal compared with cross-modal conditions.

Stroop performance: Summary

In line with the hypotheses, the current results showed that response-interference only caused additional slowing (above stimulus-interference) and accuracy decrements under unimodal conditions. These behavioural results supported the hypothesis that response interference occurs under unimodal but not cross-modal conditions.

EEG data

The N450 effect

A mass-univariate analysis showed a significant effect at Pz, where the amplitude of the ERP was significantly higher in response to congruent (*M* =

2.14uV, *SE* = .37) compared with response-incongruent (*M* = 1.87uV, *SE* = .36) conditions (t^{Max} (27)=-4.44, *p* = .0016, Cohen's *d* =-.85[20], BF_{01} = .007). This component was identified as the N450 effect and Pz was therefore used for all further comparisons[21].

The time-course and topography of the N450 effect at Pz for each condition is shown in Figure 27. A 2 (sensory condition; unimodal vs. cross-modal) x 3 (congruency; congruent, response-incongruent, stimulus-incongruent) ANOVA was performed to compare the mean amplitude of the ERPs at Pz between 400 and 500ms. This analysis showed a main effect of congruency (*F*(27) = 8.23, *p* <.001, η_p^2 = .23, BF_{01} = 1.13) but no main effect of sensory condition (*F*(27) = 2.81, *p* = .105, η_p^2 = .09, BF_{01} = .157) and no interaction between sensory condition and congruency (*F*(27) = .783, *p* = .462, η_p^2 = .03, BF_{01} = 65.77). The main effect of congruency occurred because the average amplitude of the ERP was significantly lower under response-incongruent vs. congruent (*p* <.001, BF_{01} = .007) and stimulus-incongruent (*p* = .035, BF_{01} = .22) conditions. There was no significant difference in amplitude between congruent and stimulus-incongruent conditions (*p* = 1, BF_{01} = 2.68).

In contrast to the current hypothesis and the results reported in Chapter 4 these results suggest the N450 is associated with response-interference (i.e. there was only a significant difference between response-incongruent and congruent/stimulus-incongruent conditions). Furthermore, the effect of congruency did not interact with sensory condition, suggesting similar effects between unimodal and cross-modal conditions.

---

[20] Cohens d was calculated using Cohens dz for correlated samples based on sample size and t statistics yielded from the tmax permutation test (i.e. $t\sqrt{n}$) (Lakens, 2013).

[21] Exploratory analysis confirmed that no differences between unimodal and cross-modal ERPs occurred between 400 and 500ms at any other electrodes (all t^{Max} values between -1.82 and 0.98; all *p* values >0.6472 ; all *Cohens d* values between -.34 and .19).

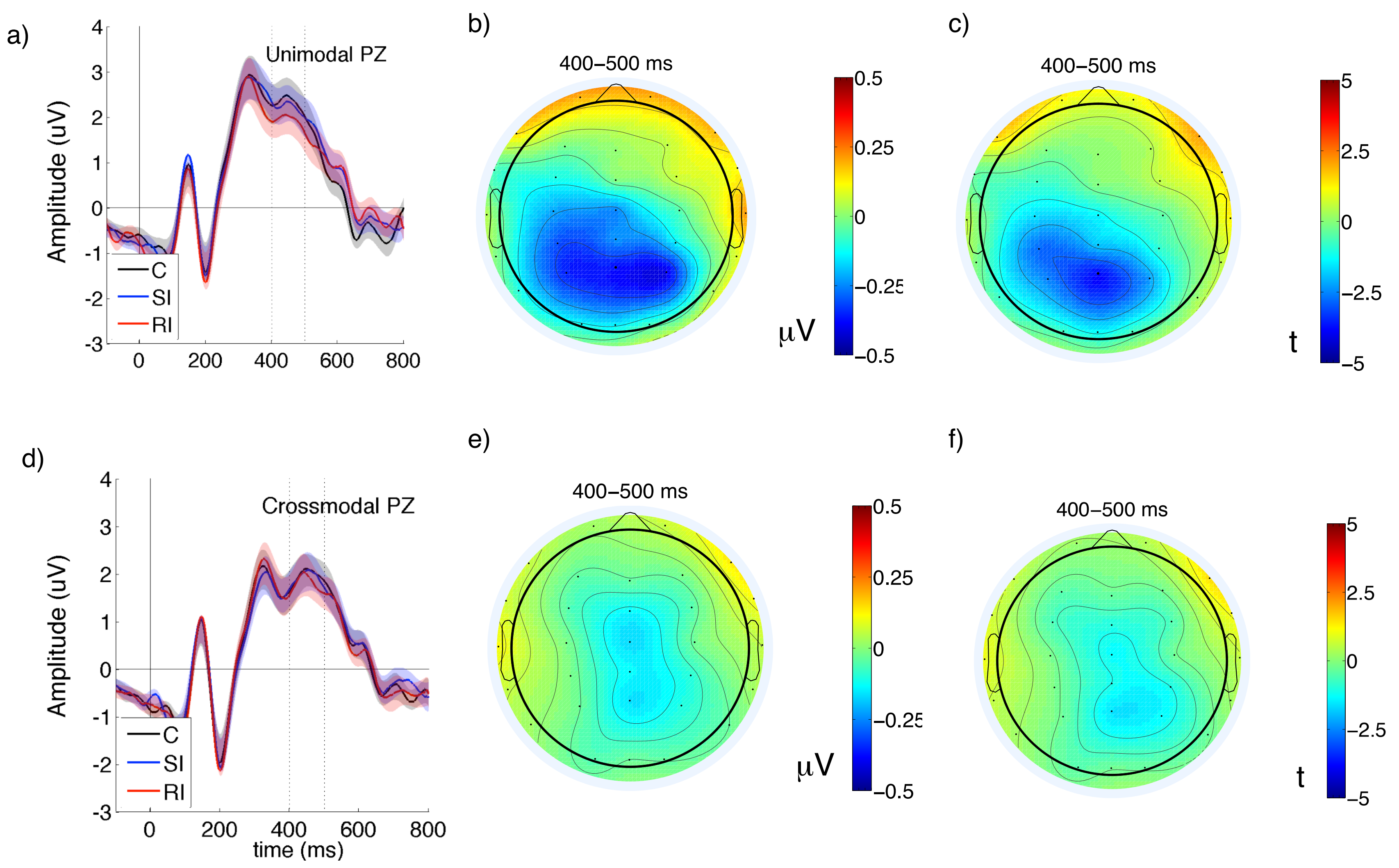


Figure 27. ERPs under unimodal (top row) and cross-modal (bottom row) conditions, shaded areas show standard error of the mean at each time point. Panels show a) ERPs in response to congruent (C - black) stimulus-incongruent (SI - blue) and response-incongruent (RI - red) conditions. b) Voltage topography of mean amplitude of RI-C difference from 400-500ms. c) topography of t-values comparing mean amplitude between 400 and 500ms to 0. Panels d-f then show these for cross-modal conditions.

Stimulus-locked LRPs

Figure 28 shows stimulus-locked LRPs under unimodal and cross-modal conditions. A 2 (sensory condition; unimodal vs. cross-modal) x 3 (congruency; congruent, stimulus-incongruent, response-incongruent) ANOVA showed no main effect of sensory condition (*F*(1, 27) = 2.12, *p* = .149, η_p^2 = .08, BF_{01} = 2.296) and no main effect of congruency (*F*(2, 54) = .09, *p* = .911, η_p^2 = .003, BF_{01} = 15.509) but a significant interaction between sensory condition and congruency (*F*(1.54, 41.46) = 3.78, *p* = .042, η_p^2 = .12, BF_{01} = 0.684). This interaction occurred because the amplitude of the response-incongruent stimulus-locked LRP was larger, and more negative, under unimodal (M = -.146, SE = .08) vs. cross-modal (M = .044, SE = .09) conditions (*p* = .009, BF_{01} =.205). As negative LRPs represent preparation of the correct response, the lower amplitude potential under cross-modal conditions is indicative of more interference. It is notable that this interaction is in the direction predicted (suggesting more stimulus-interference under cross-modal conditions). However, the difference in amplitude between congruency conditions did not reach significance in either unimodal or cross-modal conditions.

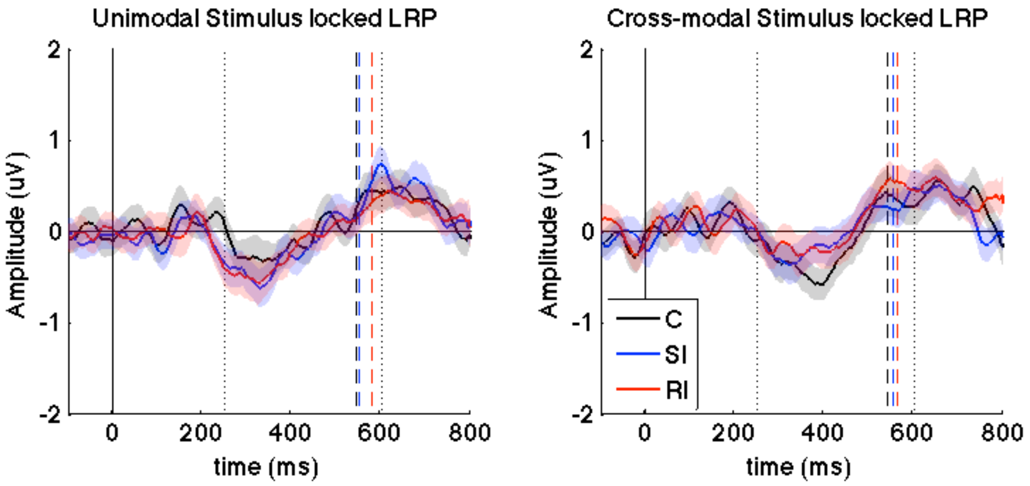


Figure 28. Stimulus-locked lateralised readiness potentials (LRPs) in response to each congruency condition under unimodal (a) and cross-modal (b) conditions. More negative deflections represent preparation of the correct response. C = congruent, SI = stimulus-incongruent, RI = response-incongruent. Shaded areas show standard error of the mean at each time point. Vertical black, blue and red broad dashed lines show average response times under each condition relative to the LRP. Thin dashed black lines indicate the 250-600ms time window in which average amplitude was compared.

Response-locked LRPs

Figure 29 shows response-locked LRPs for each congruency condition under unimodal and cross-modal conditions. A 2 (sensory condition; unimodal vs. cross-modal) x 3 (congruency; congruent, stimulus-incongruent, response-incongruent) ANOVA showed no main effect of sensory condition (*F*(1, 27) = .60, *p* = .445, η_p^2 = .02, BF_{01} = 3.338), no main effect of congruency (*F*(2, 54) = 2.74, *p* = .074, η_p^2 = .09, BF_{01} = 4.636) and no interaction between sensory condition and congruency (*F*(2, 54) = 1.74, *p* = .186, η_p^2 = .06, BF_{01} = 3.13). These findings suggest there was an absence of interference effects in either sensory condition at the level of the response-locked LRP.

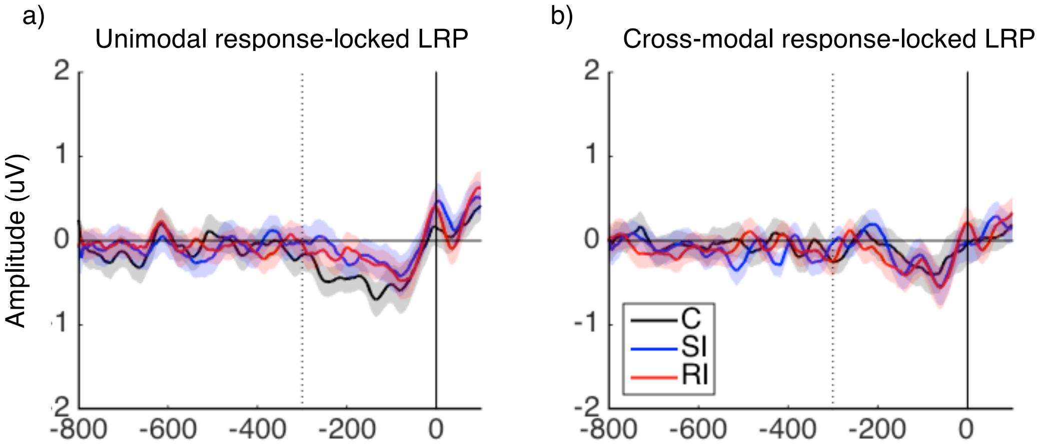


Figure 29. Response-locked, lateralised readiness potentials in response to each congruency condition under unimodal (a) and cross-modal (b) conditions. More negative deflections represent preparation of the correct response. Panels c and d show difference waves for LRPs under unimodal (c) and cross-modal (d) conditions. C = congruent, SI = stimulus-incongruent, RI = response-incongruent. Shaded areas show standard error of the mean at each time point.

Discussion

This study aimed to further identify whether unimodal and cross-modal Stroop effects occur at different processing stages using behavioural and ERP measures. It was predicted that unimodal interference would occur due to stimulus- and response-interference, whilst cross-modal interference would arise from stimulus-interference. These predictions were then extended to ERPs associated with stimulus-interference (N450 and stimulus-locked LRPs) and response-interference (response-locked LRPs). The hypotheses regarding these ERPs and associated results, in brief, were:

1) The N450 would show similar amplitude between stimulus- and response-interference and between unimodal and cross-modal conditions (thus showing stimulus-interference in both conditions). In contrast to this, the N450 did differ between stimulus- and response-incongruent conditions and also did not significantly differ between unimodal and cross-modal conditions.
2) The stimulus-locked LRP would show similar amplitude between stimulus- and response-interference and between unimodal and cross-modal conditions (again, showing stimulus-interference in both conditions). It was found that the amplitude of the stimulus-locked LRP to response-incongruent trials was significantly smaller in cross-modal versus unimodal conditions, suggesting more interference.
3) The response-locked LRP would show lower amplitude on response-incongruent vs. congruent and stimulus-incongruent trials, and this effect would occur under unimodal but not cross-modal conditions (thus showing less correct hand preparation under response-incongruent trials, i.e. response-interference, but only under unimodal conditions). Results showed no effects of congruency or sensory condition on the response-locked LRP.

In this section I will first discuss the findings from the behavioural data (response times and accuracy) and then focus on the ERP/LRP effects.

Behavioural data

Analysis of response-time ratios showed that unimodal response-interference was significantly higher than cross-modal response-interference. Furthermore, unimodal response-interference significantly differed from baseline (suggesting slowing caused by response-interference) whilst cross-modal response-interference did not. In addition to this, unimodal response incongruence reduced accuracy relative to congruent conditions, whilst cross-modal interference did not. These results support the hypothesis that unimodal interference results in part from response-interference, whilst cross-modal interference does not.

Interestingly, stimulus-interference ratios did not significantly differ from baseline in either sensory condition in this study. This is surprising as in Chapter 3 cross-modal stimulus-interference significantly differed from baseline in young adults. Similarly, in Chapter 4, unimodal stimulus-interference significantly differed from baseline in young adults. Given these results, it would be expected that stimulus-interference should also have occurred in the current study.

One possible explanation of null findings could be that distracting stimuli were presented at lower intensities in this study compared with the experiments in Chapters 3 and 4. In this study, a method-of-adjustment was used to derive thresholds, whereas a staircase task was used with adults in previous experiments. As the method-of-adjustment takes considerably less time than the staircase approach, this method was selected in this study for practical reasons. However, it has been shown that threshold values derived from the method-of-adjustment are considerably more variable than those derived from the staircase approach (Podlesek & Komidar, 2006). Furthermore, Cornsweet (1962) highlighted that in comparison with the staircase method, it is less clear what decision criterion subjects use to lower stimulus intensity in the method-of-adjustment. The main constraint of the current method-of-adjustment is that visual stimuli remained on screen, and auditory stimuli could be replayed, until the subject was able to identify the word. This is in contrast to the staircase procedure used in Chapters 3 and 4, in which stimuli were presented once, for limited time, before a decision was required. Thus, in the current study, subjects may have lowered the intensity of distracting stimuli to much lower opacity/volume therefore lowering overall interference and perhaps eliminating stimulus-interference.

To test whether thresholds were significantly lower in this study compared with Chapter 3, an additional 4 (colour background) x 2 (Experiment; this Experiment vs. young adults in Chapter 3 Experiment 1) ANOVA was performed on visual thresholds. No main effect of Experiment was found (*F*(1, 61) = 1.68, *p* = .2, η_p^2=.03). However the main effect of colour background (*F*(3, 183) = 5.01, *p* = .0026, η_p^2=.08) interacted with Experiment (*F*(3, 183) = 7.76, *p* <.001, η_p^2=.11), showing thresholds were significantly lower for reading words on red (*p* = .035), green (*p* = .015) and blue (*p* = .003) but not yellow (*p* = .926) backgrounds in the current experiment compared with young adults in Chapter 3 Experiment 1. This suggests that visual distractors were presented at lower intensity in this study. In terms of auditory stimuli, auditory thresholds appeared similar in this study compared with the young adults in Chapter 3, Experiment 1 (45dB and 43dB respectively). However, due to a calibration error, all auditory stimuli were presented 5dB below the expected value. Due to this, the presentation of all auditory distractors was capped at 60dB for all participants (i.e. not always 20dB above threshold). To identify whether the detectability of visual and auditory stimuli in this study could have influenced Stroop performance, exploratory analyses correlating distractor detectability with Stroop performance were conducted (appendix page 320). The results of these analyses did not suggest that quieter/fainter stimulus presentation influenced unimodal or cross-modal Stroop effects. As such, although stimuli were presented at lower intensities in this study, at present it remains speculative whether this influenced the amount of stimulus-interference seen in the Stroop tasks.

ERP data

The N450 effect

Chapter 4 and previous literature have linked the N450 with stimulus-interference (Killikelly & Szűcs, 2013; Szűcs & Soltész, 2012; West et al., 2004). These studies found that the N450 ERP did not significantly differ between stimulus-incongruent and response-incongruent conditions, suggesting response conflict does not further modulate the N450 effect. However, in the current study there was a significant difference between stimulus-incongruent and response-incongruent N450 effects, suggesting response conflict resulted in a larger N450. Although this finding has not been reported as frequently, this effect was found by Chen et al (2011), who reported N450 effects only for response-incongruent conditions.

One possible explanation of mixed findings is that the N450 effect is susceptible to response conflict when response demands are higher. In Chapter 4 participants completed an extensive threshold task (a staircase procedure) prior to the Stroop task. This may have accustomed participants to the response mappings therefore making response selection easier. In contrast, the current study was preceded by a method-of-adjustment in which participants gave verbal responses to very few stimuli. It is possible that this made the response demands of the current task higher, therefore resulting in response-conflict effects in the N450 response. In support of this Chen et al (2011) used 6 possible colour-word options in comparison to the 4 options used by Killikelly and Szűcs (2013), and two options in other numeric Stroop tasks (Szűcs & Soltész, 2012; West, Bowry, & McConville, 2004). However, whether the size of colour set increases Stroop interference has been questioned (Golden, 1974; McClain, 1983). Nevertheless, future research should investigate the role of response demands in the functional definition of the N450.

Given that the N450 in this study appeared modulated by response-interference, it is surprising that no significant difference was found between unimodal and cross-modal N450 amplitude. If the current N450 effect reflects response-interference, it would be expected that this effect occur for unimodal but not cross-modal conditions. Although this difference did not reach significance in this study, it can be observed that, descriptively, the size of the N450 effect did appear larger under unimodal conditions (see Figure 27). This might imply that the N450 effect in this study reflected response-interference, and was therefore larger under unimodal conditions. However, as this is based on descriptive observation, this remains speculative.

Stimulus-locked LRPs

The stimulus-locked LRP has also been tied with stimulus-interference (Smulders et al., 1995). As such, it was hypothesised that congruency differences in this effect would be greater under cross-modal conditions (i.e. the LRP should be more negative under congruent relative to incongruent conditions). The current findings showed a significant interaction between sensory condition and congruency in the stimulus-locked LRP. However, this occurred because the amplitude of the response-incongruent LRP was smaller under cross-modal versus unimodal conditions. In contrast to the current hypotheses, there was no significant effect of congruency on ERP amplitude in unimodal or cross-modal conditions.

Nevertheless, it is notable that results closely resemble those predicted. Firstly, the significant interaction does suggest a difference in interference effects occurring early in stimulus processing between unimodal and cross-modal conditions. Secondly, as shown in Figure 28 under cross-modal conditions the stimulus-locked LRP was larger for congruent versus stimulus- and response-incongruent conditions (which appeared similar in amplitude) and this effect is less clear under unimodal conditions. Although this remains speculative, this finding forms preliminary evidence to suggest that stimulus-locked LRPs detect subtle stimulus-interference effects under cross-modal conditions that could not be detected at the level of behaviour.

Response-locked LRPs

It was hypothesised that response-locked LRPs would only show congruency effects under unimodal conditions. This hypothesis was driven by the theory that unimodal interference arises from both stimulus- and response-interference levels, whilst cross-modal interference arises mainly from stimulus-interference. In contrast to this, the current findings do not indicate any evidence of differences in response-locked LRPs between sensory conditions or between congruency conditions.

As discussed in Chapter 4, LRPs with negative polarity indicate preparation of the correct response. The response-locked LRPs in the current study were all negative in polarity, indicating preparation of the correct response. This is not surprising given that, as in Chapter 4, only correct responses were included (and accuracy was high, leaving too few trials for accurate analysis of incorrect responses). However, it is also possible that response-locked LRPs are more sensitive to response facilitation. The pattern of effects shown in Figure 29 suggests that these facilitation effects were larger in unimodal compared with cross-modal trials, as a more negative deflection was seen for congruent relative to incongruent effects. Although this is in line with behavioural results (showing larger unimodal Stroop effects) this interaction did not reach significance. Furthermore, if the response-locked LRP is sensitive to facilitation, it is interesting that no facilitation was evidenced under stimulus-incongruent conditions. Finally, facilitation effects were not seen in Chapter 4. Therefore this conclusion is tentative.

As with N450 responses, both types of LRPs also did not significantly differ between stimulus- and response-interference. This is interesting given that the response-locked LRP has been more strongly associated with response interference (Killikelly & Szűcs, 2013; Masaki et al., 2004; Smulders et al., 1995). However, it has been highlighted that, when accuracy is high, the response-locked LRP may not provide the most precise measure of response conflict. Szűcs et al (2009) highlight that, due to volume conduction, if motor cortex activity for preparation of the correct and incorrect hand occurs simultaneously, activity in the motor cortices will cancel one another out. Thus, it may not be possible to detect activity associated with incorrect hand activation when accuracy is high. Indeed, Szűcs et al. (2009) were only able to derive an LRP marker of incorrect hand activity on correct trials in one of four experiments. It is possible that this limitation also held true for the current study. One finding supporting this is that the response-locked LRPs in this study were much lower in amplitude (<1uV) whereas the response-locked LRPs in Chapter 4 were larger (>2uV). This small amplitude response would suggest some cancelling out of effects. As such future research may want to include additional measures of response preparation, such as electromyography (EMG) in correct and incorrect response hands (Killikelly & Szűcs, 2013).

Conclusions

In summary, the current study aimed to further investigate whether unimodal and cross-modal interference occurred at different processing levels. Behavioural findings showed that unimodal interference arose from response competition whilst this did not occur under cross-modal conditions. In contrast to the hypothesis that N450 effects would reflect stimulus-interference, findings suggested N450 effects in this study may have been linked with response interference, and, as such, effects appeared larger under unimodal conditions (although this did not reach significance). In line with literature linking stimulus-locked LRPs with stimulus-interference (Smulders et al., 1995), the response-incongruent stimulus-locked LRP was significantly smaller under cross-modal versus unimodal conditions and congruency effects on the stimulus-locked LRP did appear to be larger under cross-modal conditions. This finding supports the theory that cross-modal interference takes place at earlier processing stages. However, although in line with apriori predictions, some of these observations remain descriptive and therefore speculative. Finally, response-locked LRPs did not show any significant effects of congruency and did not differ between sensory conditions. This finding is discussed in relation to the limitations of LRPs (i.e. volume conduction) and the response-locked LRP reflecting facilitation effects.

Chapter 6: Validating the Emotiv EPOC© portable EEG system for studying the N450 effect

Low-cost portable EEG may provide a means of collecting large-scale electrophysiological data in real-world, multisensory, contexts. In this chapter, I investigated whether portable EEG (the Emotiv EPOC© system) could measure neural markers of unimodal and cross-modal Stroop interference (the N450 effect). EEG was simultaneously recorded from research grade (Neuroscan) and Emotiv EEG systems whilst young adults (n = 27) performed unimodal and cross-modal Stroop tasks. Fewer trials were retained in the Emotiv EEG following pre-processing (although >80% of trials were retained in both systems). Intra-class correlations (ICCs) showed high similarity in ERP waveforms between EEG systems. However ICCs were lower for frontal electrodes, potentially due to higher susceptibility to eye-movement and blink artifacts. Both EEG systems detected an N450 effect under unimodal conditions, but not cross-modal, conditions. These findings provide promising support for the use of Emotiv EPOC© in studying late ERP components such as the N450. However, I also discuss several challenges that exist within the current system, and how these may be addressed.

Introduction

A central goal of cognitive neuroscience is to understand how the brain works in complex real-world environments (Matusz, Dikker, Huth, & Perrodin, 2018). However, most traditional research relies upon the assumption that laboratory-based findings generalise to the real world. Some potential solutions to this problem are a) collecting data from larger sample sizes, b) creating sharable datasets (Poldrack et al., 2017), c) creating paradigms that emulate real-world contexts (i.e. comparing unimodal and cross-modal paradigms), and d) studying the brain and behaviour in situ, outside of the laboratory (see Matusz et al., 2018 for discussion). Portable brain imaging methods represent a significant advancement towards achieving these goals. However, the scope of questions that can be answered using this technology must first be evaluated. This study contributes towards the literature validating portable EEG for studying ERPs by evaluating whether Emotiv EPOC© can measure Stroop N450 effects. I also extend existing cognitive control literature by using this technology to compare unimodal and cross-modal paradigms.

The limitations of laboratory based EEG

EEG has significantly advanced our understanding of human cognition, enabling measurement of the time-course of information processing. However, most research grade EEG systems suffer three core limitations:

1. Expensive: In terms of purchasing the equipment and creating electrically shielded laboratory environments.
2. Time-consuming: Up to 256 electrodes must be individually connected to the scalp with conductive gel and the skin underneath each electrode prepared with abrasive cream. This can also increase expenses relating to participant payment.
3. Immobile: To reduce artifacts, experiments typically take place in an electrically shielded environment and participants are asked to restrict movement.

These limitations make the use of most research grade EEG systems impractical for research requiring large sample sizes, or specific populations that are not able to travel to the laboratory and may become irritable by lengthy set-up procedures (i.e. patient groups and children). Furthermore, these limitations hamper the application of EEG for studying behaviour in real, multisensory environments. These limitations are critical for the current thesis. Firstly, I aimed to take a lifespan approach in this thesis, entailing work with children and large sample sizes. Secondly, the conclusions I draw in this thesis require testing in the real world. Exploring whether portable

EEG can be used to assess cognitive control in future research is therefore vital.

Literature using portable EEG

Several portable EEG set-ups are available on the market. These include set-ups intended for research (i.e. B-Alert X24, Enobio 20, eego™, waveguard™, Emotiv EPOCflex), commercially available low-cost alternatives intended for non-research purposes such as gaming and meditation (i.e. Muse, Mindwave, Emotiv Insight, Emotiv EPOC©) and, more recently, low-cost set ups intended for research (i.e. Open BCI). However, the capability of low-cost headsets for detecting experimental effects needs thoroughly evaluating before identifying the appropriateness of their use in research.

In a comparison of research (B-Alert X24 and Enobio 20) vs. non-research (Muse and Mindwave) wireless EEG systems, Ratti, Waninger, Berka, Ruffini, and Verma (2017) found all systems capable of measuring comparable resting state oscillatory activity. However, research grade systems had higher test-retest reliability and were less susceptible to artifacts such as eye movements (potentially due to the location of electrodes). Muse and Mindwave also have a limited number of electrodes (2 and 1 respectively). Thus, the research questions that can be answered with these headsets are limited.

The Emotiv EPOC© EEG system (herein referred to as “Emotiv) is a low-cost (~799 USD), 14-channel commercially available headset (www.emotiv.com). The Emotiv system is designed for measuring continuous data to track cognitive states over time in the order of seconds (Hairston, 2012). As such, this system has proven useful in many paradigms not requiring time-locked analyses (Grummett et al., 2015; Pham & Tran, 2012). The affordability of the system has made it possible to conduct novel, group EEG studies. For example, Aspinall, Mavros, Coyne, and Roe (2015) used the system to study oscillatory activity in a group of walkers walking through green and urban environments. Recently, Dikker et al (2017) used the system to study oscillatory synchronisation between 12 high school students whilst learning in the classroom. Through directly comparing simultaneously recorded, wired, research-grade (Neuroscan) and Emotiv EEG, Sufani, De Blasio, McDonald, and Rushby (2015) found comparable alpha-band EEG coherence between the two systems. Together, these findings suggest the Emotiv system is capable and appropriate for studying non time-locked EEG activity.

Several studies have attempted to use the Emotiv system for studying time-locked ERPs. Many studies report the system produces detectable evoked responses, such as the P300, even in adverse conditions such as walking outdoors (De Vos, Gandras, & Debener, 2014; Debener, Minow, Emkes, Gandras, & De Vos, 2012), albeit at a lower signal-to-noise ratio (Duvinage et al., 2013). However, many of these studies removed the Emotiv electrodes from its headband and attached them to a wearable cap to gain better scalp connectivity.

Concerns have been expressed regarding the timing and drift variability in Emotiv relative to wired research grade EEG (Hairston, 2012; Ries, Touryan, Vettel, McDowell, & Hairston, 2014). Hairston (2012) found that the Emotiv system was susceptible to temporal jitter[22] and drift[23], making the Emotiv system in its original form unsuitable for ERP research. As such, several studies have reported difficulties with time-locking Emotiv EEG data to event triggers (Grummett et al., 2015; Kotowski, Stapor, Leski, & Kotas, 2018; Ries et al., 2014). One solution to this problem is to conduct an offline timing correction procedure, following which the ERPs in Emotiv appeared comparable to other EEG systems (Hairston, 2012). A second approach is to directly inject the trigger information into the EEG signal. In a series of validation studies, Badcock et al, (Badcock et al., 2013, 2015; de Lissa, Sorensen, Badcock, Thie, & McArthur, 2015; Petit et al., in prep) used a wireless trigger system to directly inject the event trigger into the Emotiv EEG

---

[22] Hairston (2012) notes two forms of jitter: Stimulus-related jitter, in which the trigger/stimulus is presented at a time different to expected, and recording-related jitter, in which the recording equipment records the trigger to have occurred at a different time to the actual presentation time.

[23] Temporal drift refers to a slow cumulative difference in the timing coefficients of two systems. This occurs when the internal clocks of two systems (i.e. the EEG recording and the stimulus presentation computer) are slightly out of sync.

signal (Thie, 2013). These studies were able to derive valid and reliable ERPs for the face selective N170 (de Lissa et al., 2015), auditory evoked mismatch negativity effects (Badcock et al., 2013, 2015) and, most recently, N400 effects associated with semantic incongruence (Petit et al., in prep). Using an alternative system to inject event markers directly into the Emotiv EEG Kotowski et al, (2018) also recently found early posterior negativity (EPN) effects for emotion processing using the Emotiv system. Together, these findings indicate strong support for the use of Emotiv in studying time-locked ERPs.

To date the Emotiv system has been validated (i.e. directly compared against research grade systems) for several early sensory evoked ERPs, including the N170 face selective response (de Lissa et al., 2015), auditory mismatch negativity (MMN) (Badcock et al., 2013, 2015) and the P300 response (Barham et al., 2017; Duvinage et al., 2013). Fewer studies have investigated whether Emotiv can be used for later components. It has been reported that the system can detect N400 semantic incongruence effects (Ousterhout, 2016; Petit et al., in prep). However, to my knowledge, only one study (Petit et al, in prep) has compared later non-sensory related components between Emotiv and research grade EEG. This study found the Emotiv system capable of detecting N400 effects, although at a lower amplitude. This lower amplitude has been attributed to the lower signal-to-noise ratio in Emotiv (Barham et al., 2017), which may be due to higher impedances and susceptibility to artifacts.

The aim of the current study was to validate whether the Emotiv system could be used to study the Stroop N450 effect (for details of the effect see Chapter 4 page 84). The N450 effect is a commonly used marker of conflict detection. As such, if Emotiv is capable of studying this effect this opens up the possibility to study conflict detection in the real world. Furthermore, comparing the system’s ability to detect unimodal and cross-modal conflict indicates the types of real-world conflict that portable EEG may be sensitive to[24].

Method

Participants

The participants in this study were the same participants that took part in the Experiment presented in Chapter 5. However, only participants yielding usable data from both EEG systems were included in the analysis (final n = 27).

Equipment

The experimental stimuli and equipment used to present stimuli were identical to those outlined in Chapter 5. The Neuroscan EEG was also identical to that outlined in Chapter 5. However, only the 12 electrodes corresponding to the locations of the Emotiv electrodes were included in the analysis (see Figure 30).

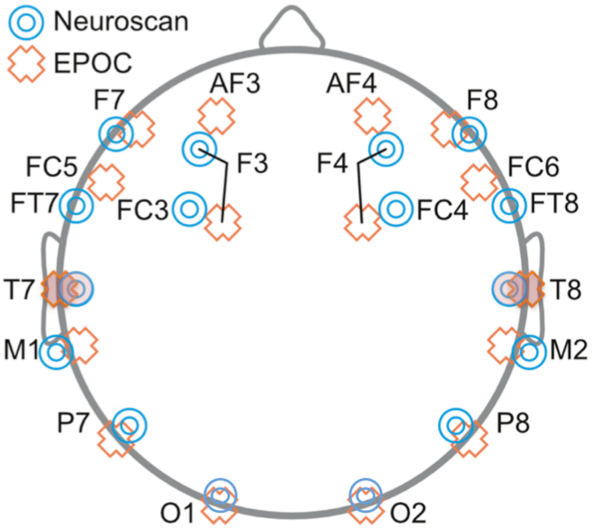


Figure 30. Placement of Emotiv (blue circles) and Neuroscan (Orange crosses) electrodes. Electrodes shaded in red (T7/T8) were used for event marking and not included in analyses. M2/M1 were attached to the ear lobes. Adapted from Badcock et al. (2015).

---

[24] Unlike previous chapters in this thesis, this study does not focus on lateralised readiness potential (LRPs). This is because these potentials are calculated using central electrodes positioned over the motor cortex (C3/C4) and these electrodes are not available in the Emotiv headset.

Emotiv EEG

The Emotiv EPOC© system (herein referred to as “Emotiv”) was set up over the top of the research grade Neuroscan system (see Figure 31a). As with the Neuroscan, each electrode location was prepared with abrasive gel (Nu Prep). Electrodes were then connected using cotton pads soaked in sodium chloride saline solution. Holes were cut in the Easycap to allow the Emotiv sensors to reach the scalp. Electrode connectivity was tested using the TestBench software (version 3.1.21) and sensors were adjusted until they reached the “green” level, reported to be equivalent to 220 kΩ (Badcock et al., 2015).

EEG data were recorded from 12 scalp electrode sites, aligned with the 10-20 system: AF3, F7, F3, FC5, P7, O1, O2, P8, FC6, F4, F8, FC4. Two reference electrodes were also wired to M1 and M2 and placed on the right and left ear lobes, which respectively served as a ground reference, to compare the voltage of other electrodes, and a feed-forward reference, to reduce electrical interference. EEG was collected using the Emotiv TestBench software with an online sampling rate of 256 Hz and processed online with a high-pass filter of 0.16Hz and a low-pass filter of 83Hz, digitized at 2048Hz and filtered with a 5th order sinc notch filter 50-60Hz.

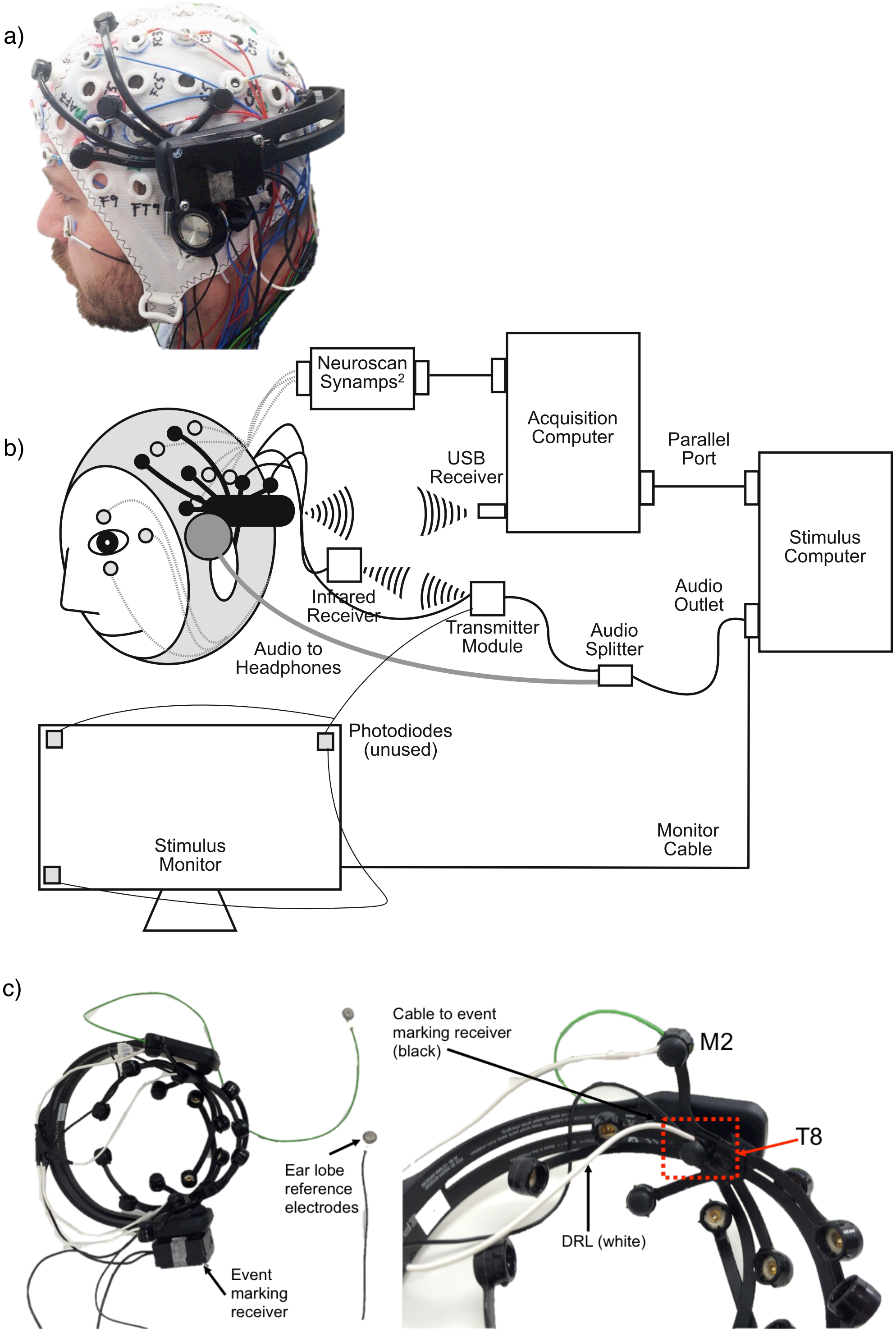


Figure 31. a) Neuroscan (white) and Emotiv EPOC (black) dual set up. b) Schematic of Neuroscan (grey) and Emotiv (black) EEG set up with wireless event marking system. Note that although photodiodes were connected to the transmitter module these were unused (all triggers were sent via the audio outlet). Adapted from Badcock et al (2013) and de Lissa et al (2015). c) wiring of modified Emotiv headset.

Event marking

The event marking system used to insert event markers into the ongoing Emotiv EEG is illustrated in Figure 31b. Electrodes T8/T7 were “sacrificed” to receive event markers from a wireless transmission system (Thie, 2013). The system consisted of a transmitter and a receiver unit linked using an infrared light. The transmitter unit was set up such that it could receive a signal either from a visual stimulus presented to a photodiode covering the corner of the monitor (de Lissa et al., 2015), or from a tone presented at stimulus onset via a headphone splitter from the computer’s audio-output (Badcock et al., 2015). However, signals were only presented via the audio-port (i.e. a constant grey background was always present underneath the photodiode). The tones used to generate the signal were 100ms in duration and of four different frequencies depending on the congruency condition (congruent = 500Hz, stimulus-incongruent = 750Hz, response-incongruent = 1000Hz and neutral = 1250Hz). Once the transmitter unit received a signal this was converted to an infrared light, which was then picked up by the receiver units on the headset. The receiver unit then converted this into a pulse that was injected into the T7/T8 channels of the Emotiv headset. The duration of the pulse varied depending on the frequency[25].

Electrodes T7 and T8 were attached to a Driven Right Leg (DRL) through wires and 4700-ohm resistors that mimicked connection with the scalp. Event markers were then identified semi-automatically using the matEPOC toolbox in MATLAB (Badcock & Thei, 2015). Events in the T7/T8 channel exceeding +/-50uV were detected automatically and checked manually before aligning the event markers with the identified pulses.

---

[25] Following data collection it was realized that the transmitter unit could only produce two pulse durations (100ms for a low frequency, <1000Hz, tone and 200ms for a high frequency, >1000Hz, tone (Thie, 2013)). However conditions could still be identified from the order of conditions noted in the experiment log file.

Analysis

EEG pre-processing

The pre-processing pipeline used is illustrated in Figure 32. This pipeline closely resembled the approach used in previous Emotiv validation studies (Badcock et al., 2013, 2015; de Lissa et al., 2015). To ensure fair comparison between the two EEG systems, data from the research grade system were down-sampled to 256Hz to match the portable system and electrodes not contained in the portable EEG system were excluded from analysis. EEG from both systems were band-pass filtered 1-30Hz and re-referenced to an average of all electrodes in their respective system (excluding eye channels and mastoids). Independent component analysis (ICA) was then used to separate components associated with eye blink artifacts. Initially, removal of eye blinks using the identified components from ICA was attempted using the automated artifact classifier used in Chapter 5 (MARA; Winkler et al., 2011). However, most automated systems identify eye blinks using either correlation with the EOG channels (such as in FASTER; Nolan et al., 2010) or the spatial distribution of the component (i.e. whether the component has frontal distribution or peaks on at a “border” as in MARA). Due to the predominantly frontal distribution of electrodes in Emotiv, these approaches were not judged to be appropriate and components were manually rejected. Components were rejected if they had frontal distribution AND temporally coincided with a voltage fluctuation on the EOG channels. Note that, due to the frontal distribution of electrodes in the Emotiv system, both of these conditions had to be met.

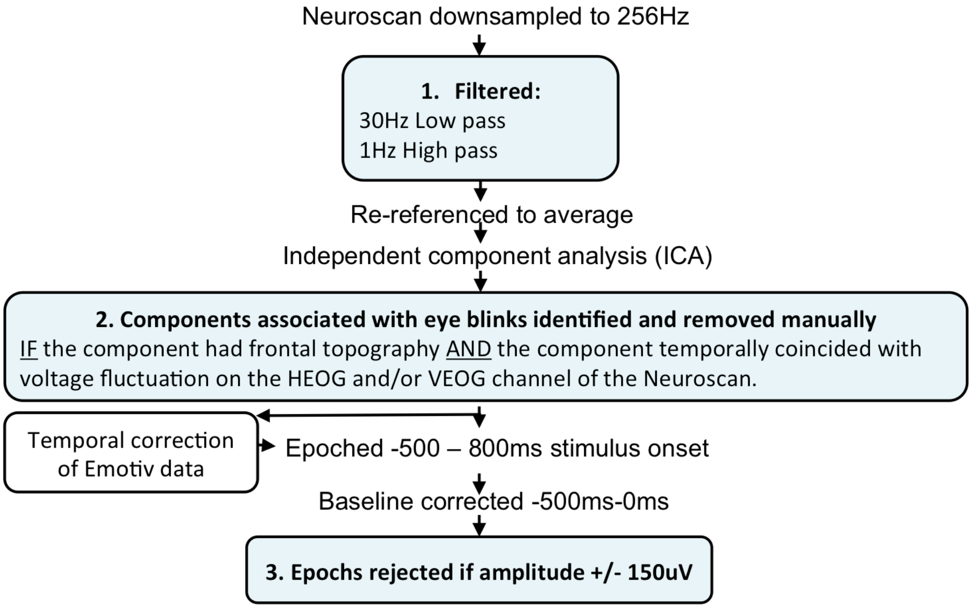


Figure 32. Pre-processing pipeline used with both Neuroscan and Emotiv EEG.

Data were epoched -500 to 800ms pre/post stimulus onset. Epochs were rejected if they contained voltage fluctuations larger than +/-150uV amplitude, if they were responded to incorrectly or excluded as response time outliers in the behavioural data, and if the trial was not available in both EEG systems (i.e. due to missing triggers in the Emotiv system).

Participants were only included in analyses if they were not identified as behavioural outliers and if they retained more than 50% of trials in both the Emotiv and Neuroscan EEG following pre-processing. Following this, two participants were excluded based on behavioural results, a third participant was removed based on both limited trials and behavioural performance and a fourth participant was removed due to a technical difficulty with Emotiv EEG. This left 27 participants available for analysis.

Temporal correction of Emotiv data

During pre-processing it was noticed that eye movement artifacts in the Emotiv and Neuroscan EEG differed in latency relative to the event markers. The artifacts in the Emotiv system occurred closer to the event marker compared with the Neuroscan artifacts – suggesting a delay in the Emotiv trigger. To correct for this, Emotiv and Neuroscan data were realigned offline using the eye movement artifacts identified in the data prior to independent component rejection (example continuous data pre- and post-correction shown in Figure 33).

VEOG and HEOG (Neuroscan) channels were used to identify blinks and horizontal eye movements. Horizontal eye movements were identified as bilateral potentials with fast onset and briefly sustained potential with opposing polarity between left and right electrodes (Figure 33 left column). Blinks were identified as bilateral, high amplitude positive potentials with brief duration (Figure 33 right panels; Iwasaki et al., 2005). The four most frontal electrodes in each EEG system were used to search for eye artifacts (F3/F4 of the Neuroscan system, AF3/AF4 of the Emotiv system and F7/F8 in both Neuroscan and Emotiv systems – see Figure 30 for placement).

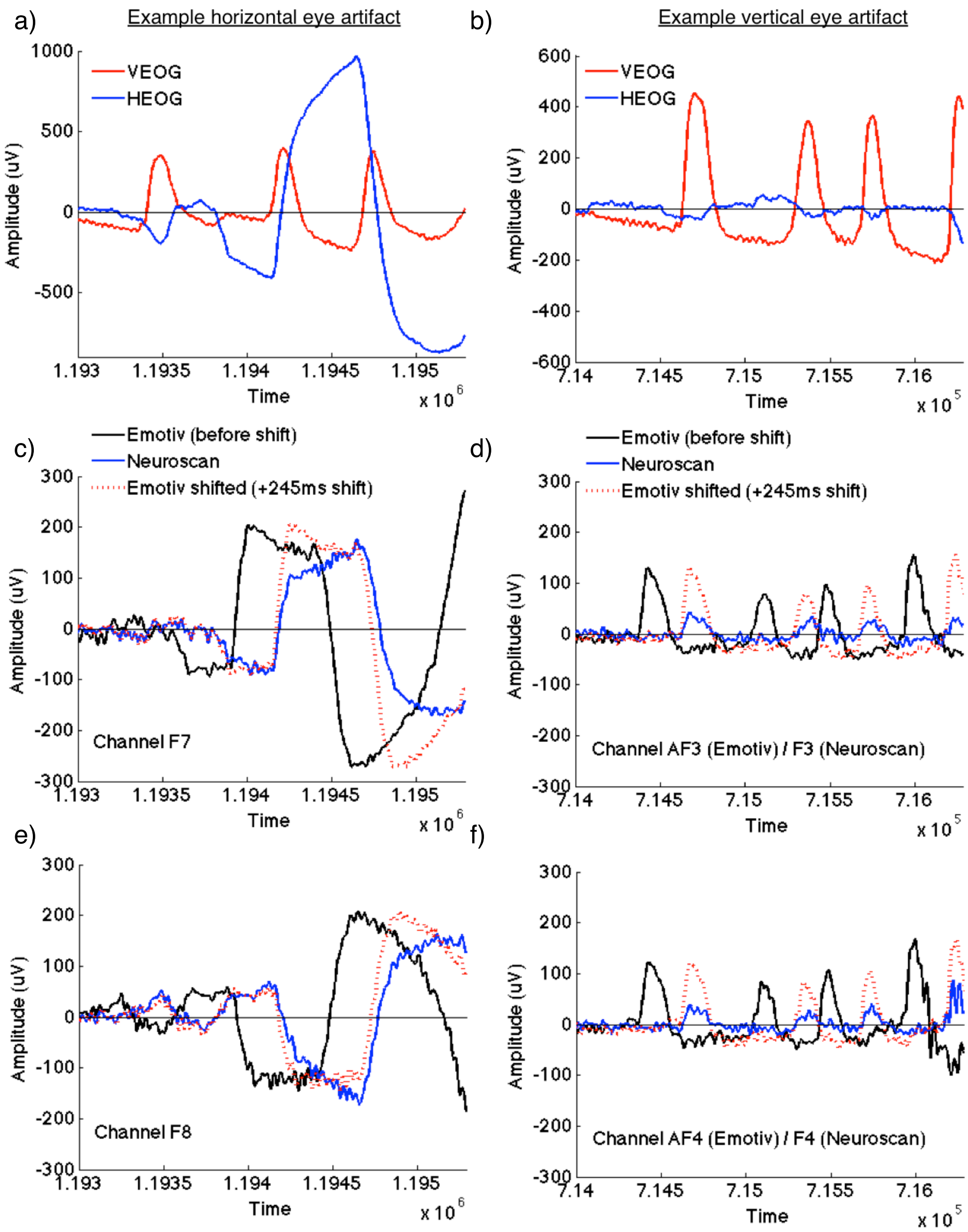


Figure 33. Example eye movement artifacts in a single subject used to realign Emotiv and Neuroscan data. EOG channels were used to identify horizontal (*a*) and vertical (*b*) eye movements and corresponding artifacts. *c – f* show artifacts in the Neuroscan (blue) and Emotiv (black) EEG relative to the event markers in each system. Red dashed lines show Emotiv data following temporal correction. X axis shows time into recording in seconds.

The majority of eye movements occurred during the periods of rest offered every 48th trial. Artifacts were therefore identified using the last event of each block using the time-window -600ms-7000ms pre/post stimulus presentation. The peak of each artifact was manually selected for the Emotiv and Neuroscan data and the time difference calculated. The average time-difference between peaks was then used to realign Emotiv and Neuroscan data. This temporal correction procedure was applied separately to each individual's EEG data (mean lag = 225ms, SE = 9.62ms, range 126 – 309ms). Notably, the measured delay was relatively consistent within participants, although varied between subjects, suggesting some unreliability of the event marking system between testing sessions. This may have resulted from rewiring of the set up between participants, and is discussed as a limitation in the discussion of this Chapter.

Comparison of Emotiv vs. Neuroscan

Emotiv and Neuroscan EEG data were compared on three metrics:

1. Number of accepted epochs following pre-processing.
2. Intra-class correlations (ICCs) to index the similarity of stimulus-locked waveforms between EEG systems. This was conducted for all electrodes.
3. Sensitivity to the N450 effect (i.e. a lower amplitude ERP under incongruent relative to congruent conditions occurring 400-500ms post stimulus presentation).

Each of these comparisons was conducted in parallel for unimodal and cross-modal Stroop effects. Although the current study contained trials that were both stimulus- and response-incongruent (i.e. the design used in Chapter 5), this was not the primary focus of this study. I therefore only focus on response-incongruent (herein referred to as incongruent) vs. congruent conditions. The number of trials retained in other conditions was comparable to those reported here and are presented in appendix, page 326. As with other EEG chapters in this thesis, I focus on the amplitude rather than latency of ERPs. This is firstly because of the challenge of identifying peaks in the difference waves and secondly because a latency correction was applied, making conclusions regarding latency speculative.

Results

Number of accepted epochs

The distribution of number of accepted epochs was negatively skewed. Wilcoxen signed rank tests were therefore used to compare the number of trials retained in each EEG system. As shown in Table 6 a significantly higher number of trials were accepted for Neuroscan compared with Emotiv (of the 72 trials per condition). This is in line with the finding that the Emotiv system was more susceptible to artifacts such as eye blinks, likely due to the more frontal positioning of the electrodes and susceptibility to movement. However, a high enough proportion of each condition was retained for comparisons between systems (i.e. >80%).

	EEG system			
Condition	Neuroscan	Emotiv	*Z*	Shared epochs
Unimodal				
Congruent	63.47 (1.08)	61.63 (1.07)	-3.533***	61.7 (1.05)
Incongruent	58.33 (2.08)	56.73 (2.07)	-3.448***	57.03 (2.05)
Cross-modal				
Congruent	66.27 (0.80)	64.33 (.94)	-3.535***	64.37 (.93)
Incongruent	64.00 (1.07)	62.53 (1.06)	-3.201***	62.53 (1.06)

p* <= .05, ** *p* <= .01, * *p* <= .001

Table 6. Mean (standard error) number of accepted and shared epochs for Neuroscan and EPOC in each condition. Seventy two trials were presented per condition. Wilcoxen Signed Rank Tests (*Z*) were used to test the difference between systems.

ICCs

Waveforms for each condition in each EEG system are illustrated in Figure 34 (unimodal) and Figure 35 (cross-modal). Intra-class correlations (ICC) were computed using the one-way random model reliability analysis method in SPSS. This is the equivalent to approaches previously reported (Badcock et al., 2015; Bishop, Hardiman, Uwer, & Von Suchodoletz, 2007; de Lissa et al., 2015; McArthur, Atkinson, & Ellis, 2009; McArthur & Bishop, 2004). The ICC measures the degree of similarity between two waveforms and produces a score ranging from 0 and 1 with higher scores reflecting a higher degree of similarity between the two EEG systems. As shown in Table 7, ICCs were relatively high across all electrodes. However ICCs were lowest over frontal electrodes (Emotiv AF4/AF3 and Neuroscan F3/F4). One explanation of this is that the most frontal electrodes in the Emotiv system were more susceptible to eye movement artifacts and this may have contaminated ERPs even following independent component rejection.

Unimodal

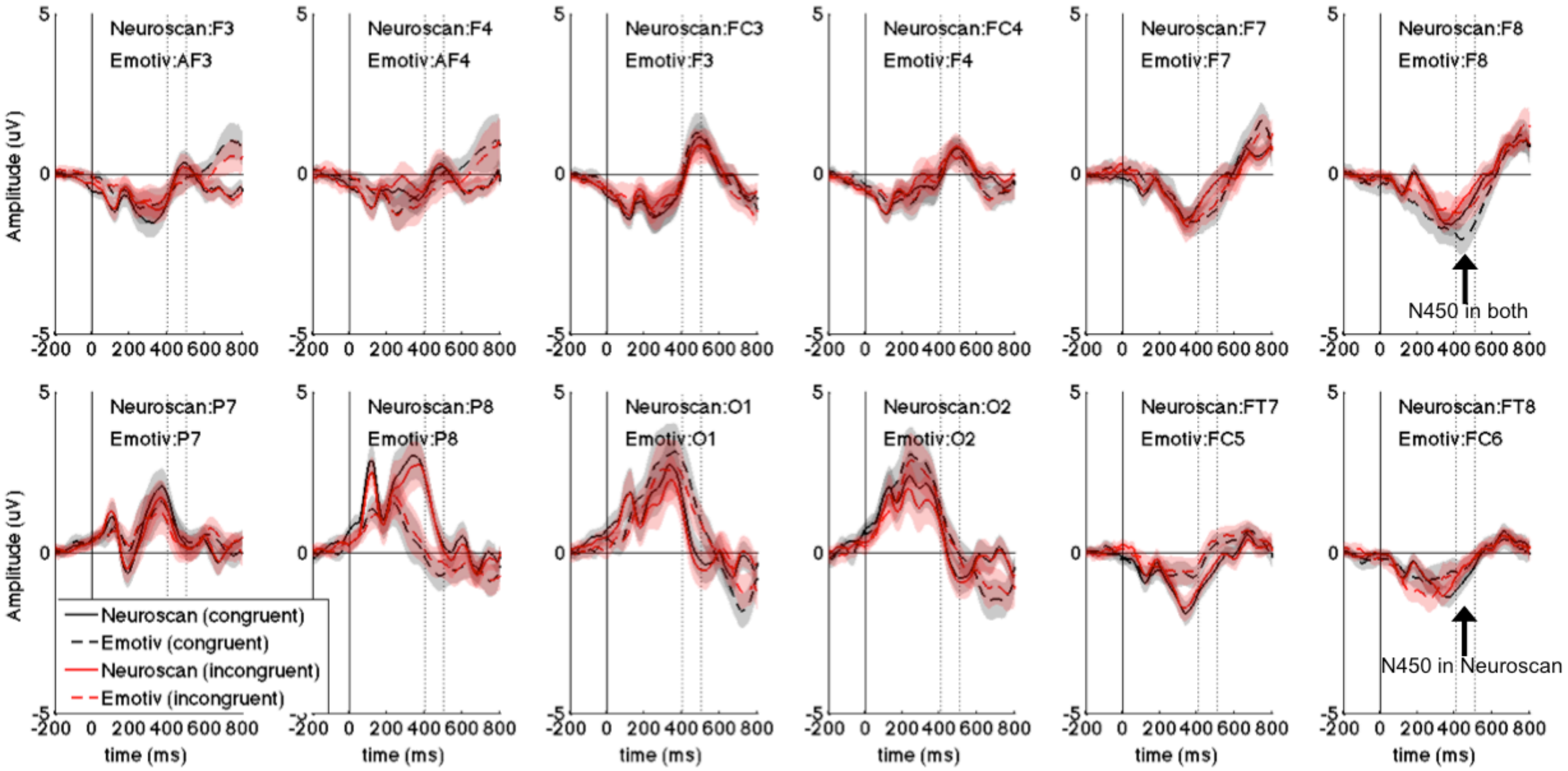


Figure 34. ERPs at each electrode in response to congruent (black) and incongruent (red) stimuli (under unimodal conditions) in the Neuroscan (continuous) and Emotiv (dashed) EEG systems. All ERPs are time-locked to stimulus onset. Vertical dashed lines show the 400-500ms time window used for comparison of the N450 effect. Shaded areas indicate standard error of the mean at each point.

Cross-modal

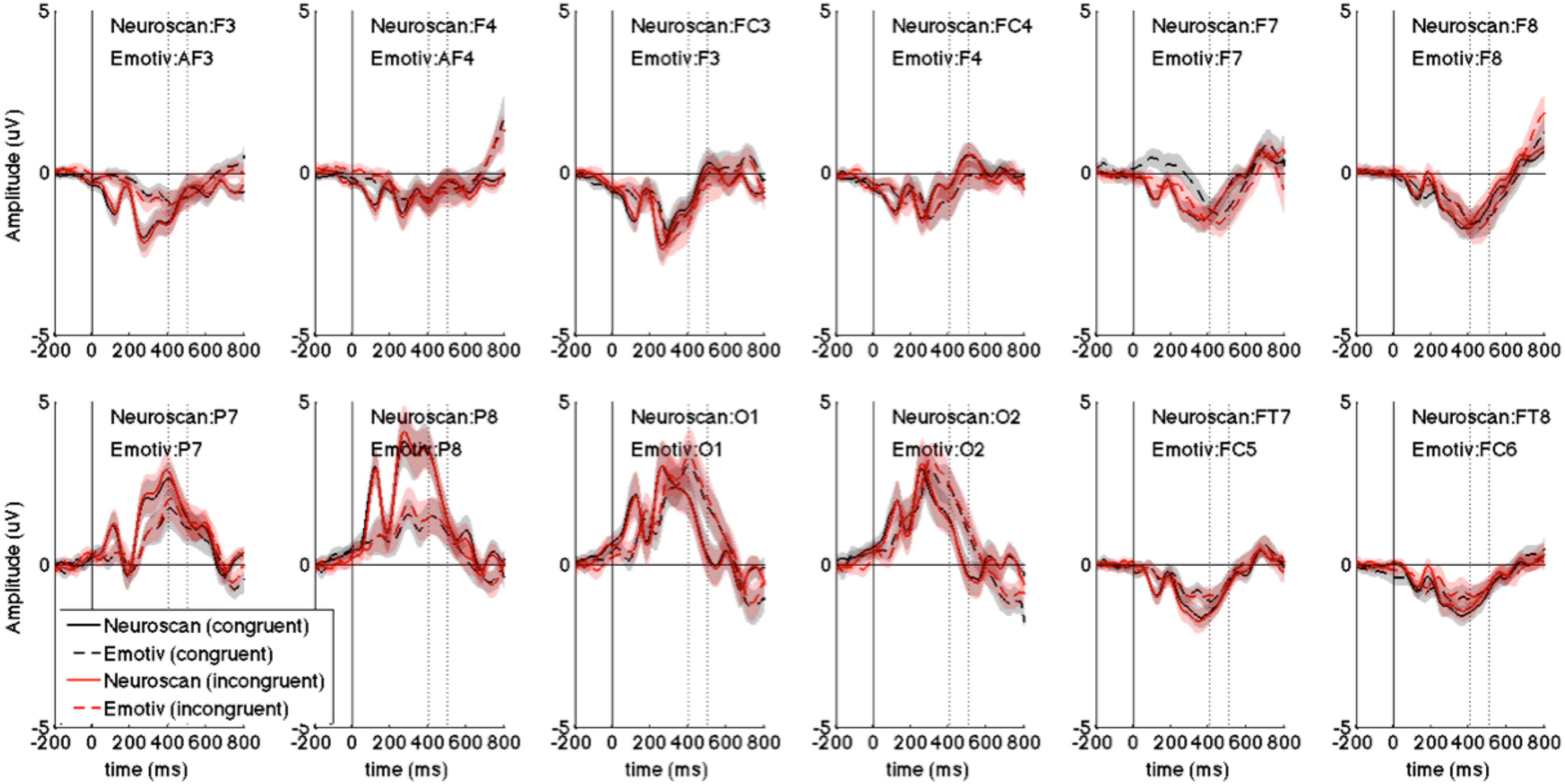


Figure 35. ERPs at each electrode in response to congruent (black) and incongruent (red) stimuli (under cross-modal conditions) in the Neuroscan (continuous) and Emotiv (dashed) EEG systems. All ERPs are time-locked to stimulus onset. Vertical dashed lines show the 400-500ms time window used for comparison of the N450 effect. Shaded areas indicate standard error of the mean at each point.

		Electrode (Emotiv/Neuroscan)											
Condition		AF3/F3	AF4/F4	F3/FC3	F4/FC4	F7/F7	F8/F8	FC5/FT7	FC6/FT8	P7/P7	P8/P8	O1/O1	O2/O2
Unimodal	C	.55	.47	.96	.87	.86	.90	.75	.85	.87	.57	.89	.90
		[.45, .64]	[.36, .58]	[.94, .96]	[.84, .89]	[.82, .88]	[.88, .92]	[.69, .8]	[.82, .88]	[.84, .89]	[.47, .65]	[.82, .89]	[.87, .92]
	IC	.51	.29	.89	.81	.87	.92	.63	.85	.75	.70	.79	.88
		[.40, .61]	[.12, .43]	[.86, .91]	[.76, .85]	[.84, .89]	[.90, .94]	[.55, .71]	[.81, .88]	[.69, .80]	[.63, .77]	[.75, .83]	[.85, .9]
Cross-modal	C	.50	.62	.81	.71	.60	.90	.86	.89	.81	.63	.77	.80
		[.38, .60]	[.53, .69]	[.76, .85]	[.64, .77]	[.51, .68]	[.87, .92]	[.83, .89]	[.86, .91]	[.77, .85]	[.55, .71]	[.72, .82]	[.75, .84]
	IC	.67	.48	.83	.69	.81	.86	.86	.79	.83	.71	.76	.79
		[.60, .74]	[.35, .58]	[.78, .86]	[.61, .75]	[.76, .84]	[.83, .89]	[.80, .87]	[.94, .83]	[.79, .86]	[.64, .77]	[.70, .80]	[.74, .83]

Table 7. Average intra-class correlations (ICCs) and 95% confidence intervals between Emotiv and Neuroscan waveforms at each electrode in each condition. Uni = Unimodal, Cross= cross-modal, C = Congruent, IC = Incongruent.

The N450 effect

The N450 is classified as a lower amplitude component in response to incongruent relative to congruent trials occurring around 400-500ms post stimulus onset (resulting in a negative difference wave). To identify electrode locations at which the N450 occurred two mass univariate analyses (one for each EEG system) were performed to compare the mean amplitude of congruent vs. incongruent ERPs 400-500ms post stimulus presentation at all electrodes (as outlined in Chapter 4, page 96, this mass univariate approach controlled for multiple comparisons between electrodes). Once electrodes of interest were identified, Bayesian one-sample t-tests were used to compare the amplitude of effects to 0 (to evaluate support for the null) and Bonferroni corrected paired t-tests were used to compare the mean amplitude of N450 effects between EEG systems at electrodes showing effects in both EEG systems. Table 8 and Table 9 show all statistics for the N450 effect under unimodal and cross-modal conditions respectively.

No electrodes showed N450 effects in either system for the cross-modal Stroop task. Electrodes showing N450 effects in the unimodal Stroop within this time-window are highlighted in Figure 34. A significant effect was seen at F8 in both Emotiv and Neuroscan systems. An effect was also found at FT8 in the Neuroscan system that was not seen in Emotiv. Notably, both of these effects were positive in polarity, indicating the response was more positive in incongruent vs. congruent conditions. This is because the N450 typically peaks at central/left parietal locations (shown in Chapters 4 and 5). As such, this results in an effect of reversed polarity over right frontal locations (this was seen in Chapter 4, where a positive difference was seen at F6). As such, although positive, this difference was still associated with the N450 effect.

These findings suggest that both the Emotiv and Neuroscan systems were able to detect the unimodal N450 effect, even with limited electrode coverage. However, the peak location of effects may differ between systems. Interestingly, the size of the effect at F8 was actually larger in Emotiv (*M* = .81uV, *SE* = .27) vs. Neuroscan (*M* = .43uV, *SE* = .13). However a Bonferroni comparison of this difference showed no significant difference between EEG systems (*t*(26) = 1.49, *p* = .148, Cohens *d* = .23, BF_{01} = 1.831), although Bayes factors also did not provide strong support for the null in this comparison, suggesting a lack of difference may be due to an underpowered comparison.

Unimodal N450 effect

	Emotiv				Neuroscan			
Electrode	t^{Max}	*p*	*d*	BF_{01}	t^{Max}	*p*	*d*	BF_{01}
AF3/F3	0.16	1	0.03	4.85	-0.8	0.982	-0.15	3.67
AF4/F4	-1.35	0.911	-0.26	2.18	0.42	1	0.08	4.53
F3/FC3	-1.15	0.972	-0.22	2.7	-1.73	0.511	-0.33	1.32
F4/FC4	-0.49	1	-0.1	4.39	0.24	1	0.05	4.78
F7/F7	0.7	1	0.14	3.92	0.15	1	0.03	4.85
F8/F8	3.06	0.046*	0.59	0.12	3.22	0.027*	0.62	0.09
FC5/FT7	0.9	0.995	0.17	3.39	1.28	0.822	0.25	2.35
FC6/FT8	0.35	1	0.07	4.64	4.38	<.001***	0.84	0.01
P7/P7	-2.03	0.437	-0.39	1.51	-1.94	0.381	-0.37	0.84
P8/P8	1.64	0.735	0.32	1.51	0	1	0	1.51
O1/O1	-1.97	0.481	-0.38	0.88	-2	0.35	-0.38	0.88
O2/O2	-0.91	0.995	-0.17	3.38	-0.57	0.998	-0.11	4.22

p* <= .05, ** *p* <= .01, * *p* <= .001

Table 8. Statistics for the amplitude of the N450 effect under unimodal conditions. t^{Max} = the t value from the mass univariate analysis (corrected for 12 comparisons), *d* = Cohens *d,* BF_{01} = Bayes factor indicating support for the null hypothesis that N450 amplitude = 0. Cohens *d* was calculated using Cohens d_z for correlated samples based on sample size and t statistics yielded from the tmax permutation test (i.e. $t\sqrt{n}$) (Lakens, 2013).

Exploratory analysis of cross-modal Stroop effect

The current study did not find cross-modal Stroop interference produced an N450 effect 400-500ms post stimulus onset at any electrodes in either EEG system. As I have proposed that cross-modal Stroop effects may occur earlier in processing, an exploratory analysis was conducted to compare congruent and incongruent ERPs at all time-points 0-800ms post stimulus presentation in all electrodes (using the mass-univariate approach to correct for multiple comparisons). This analysis did not show any difference between congruent and incongruent ERPs under cross-modal conditions in either EEG system (all *p* values >= .86 in Emotiv and >= .79 in Neuroscan).

Cross-modal N450 effect

	Emotiv				Neuroscan			
Electrode	t^{Max}	*p*	*d*	BF_{01}	t^{Max}	*p*	*d*	BF_{01}
AF3/F3	-0.09	1	-0.02	4.89	-0.97	.958	-0.19	3.2
AF4/F4	-0.64	1	-0.12	4.06	-0.33	1	-0.06	4.67
F3/FC3	-1.32	.898	-0.25	2.26	-1.57	.66	-0.3	1.66
F4/FC4	0.16	1	0.03	4.85	-0.31	1	-0.06	4.69
F7/F7	-0.78	.998	-0.15	3.72	-0.15	1	-0.03	4.86
F8/F8	-0.83	.996	-0.16	3.58	2.28	.246	0.44	0.55
FC5/FT7	0.34	1	0.06	4.66	-0.37	1	-0.07	4.61
FC6/FT8	0.16	1	0.03	4.85	1.64	.616	0.32	1.51
P7/P7	1.16	.952	0.22	2.68	0.54	.999	0.1	4.29
P8/P8	0.44	1	0.09	4.49	0.11	1	0.02	4.88
O1/O1	1.33	.892	0.26	2.23	-0.76	.99	-0.15	3.77
O2/O2	0.47	1	0.09	4.44	-0.3	1	-0.06	4.71

p* <= .05, ** *p* <= .01, * *p* <= .001

Table 9. Statistics for the amplitude of the N450 effect under cross-modal conditions. t^{Max} = the t value from the mass univariate analysis (corrected for 12 comparisons), *d* = Cohens *d*, BF_{01} = Bayes factor indicating support for the null hypothesis that N450 amplitude = 0.

Discussion

The current study aimed to evaluate whether the Emotiv EPOC© portable EEG system can be used to measure the N450 Stroop effect in a unimodal and cross-modal Stroop task. This research contributes towards the accumulating literature evaluating whether commercially available portable EEG can be used to study time-locked brain activity in real-world contexts. Validation of a low-cost portable EEG system such as Emotiv would open up the possibility to collect electrophysiological data from multiple participants of all age ranges outside of the laboratory (for example at public engagement events designed to collect data from large sample sizes, such as Summer Scientist Week, www.summerscientist.org).

The primary results of this study were; *a*) as with previous studies (Badcock et al., 2015) more trials were retained within the research grade EEG system compared with Emotiv, *b*) intra-class correlations (ICCs) showed a high degree of similarity between ERP waveforms in all conditions between systems, however similarity appeared lower at frontal electrodes, *c*) both EEG systems were able to detect an N450 effect in the unimodal Stroop task (neither detected a cross-modal Stroop effect). Together these results suggest the Emotiv system as a promising research tool to study ERPs. However, several challenges were also encountered that currently limit the use of this set-up. First, identified lag in the event-marking system indicates a need for calibration a protocol prior to data collection and approaches to "sanity check" events following data collection. Second, EOG channels from the Neuroscan system were used to manually reject artifacts from both systems. However, if the Emotiv set-up is to be used standalone, artifact rejection methods for use with this system must be developed. I shall now discuss the implications of the ERP results followed by a consideration of limitations and how they may be approached in future research.

Can the Emotiv system be used to measure ERPs?

As with previous validation research (Badcock et al., 2015) Emotiv and Neuroscan were compared on three metrics; number of trials retained, intra-class correlations (ICCs), and the amplitude of the N450 response. Although a high proportion of trials were retained in both EEG systems, significantly fewer epochs were accepted in the Emotiv compared with the Neuroscan system. This is a result that has been reported previously (Badcock et al., 2015; Barham et al., 2017) and may be attributed to the higher susceptibility of the Emotiv system to artifacts such as eye-movements. As shown in Figure 33, the Emotiv system appeared particularly susceptible to vertical eye artifacts (i.e. blinks), producing artifacts around twice the size of those seen in the research grade system. One explanation of this is the more frontal positioning of the AF3/AF4 electrodes in the Emotiv system compared with the F3/F4 electrodes in the Neuroscan. This explanation is supported by the results of the ICC analyses. The ICCs in the current study were relatively high, and comparable to those previously reported. However, the lowest ICCs were seen at AF3/F3 and AF4/F4. It is possible that higher susceptibility to eye-blinks at these electrodes in the Emotiv system distorted ERPs (even following artifact rejection), resulting in lower similarity between systems. Together these issues highlight the need for optimised artifact rejection methods if the Emotiv EEG system is to be used in research.

This study is the first to use the Emotiv EEG to study the N450 Stroop effect. Furthermore, this is the first to also implement a cross-modal paradigm. Both the Emotiv and Neuroscan EEG systems were able to detect an N450 effect under unimodal conditions. This effect was also seen at the same electrode in both systems (F8), although the Neuroscan system also detected an effect at FT8 that was not identified in the Emotiv system. Remarkably, the effect at F8 appeared larger in the Emotiv compared with the Neuroscan (although this difference did not reach significance). This is contrary to previous reports, that have shown the ERPs in Emotiv appear smaller in amplitude (Barham et al., 2017; Petit et al., in prep). One explanation of this might be that F8 was not used for comparisons in these studies, and this location may be more optimal for Emotiv. For example, as shown in Figure 31a, the “arm” connecting F8 to the headband is shorter than the arm connecting other electrodes, likely resulting in less movement. This highlights the benefits of including all electrodes in ERP comparisons (using a mass univariate approach) to identify the optimal location of effects. A second explanation might be that EOG channels were used to remove eye-movements, thus achieving a higher signal-to-noise ratio. One issue that must be emphasised is that the effect seen at FT8 in the Neuroscan appeared strong and significant, however no effect was seen at the Emotiv electrode corresponding to this location (FC6). This suggests that Emotiv may not be sensitive to all experimental effects identified in the research grade EEG.

The fact that neither system found N450 effects in the cross-modal

Stroop task may be due to two factors. Firstly, there may genuinely have been no effect of cross-modal congruency at the level of the N450 effect. Secondly, the cross-modal N450 effect may have been in a location that could not be detected with the available electrode coverage. In Chapter 5 all electrodes in the Neuroscan system were used to compare unimodal and cross-modal Stroop effects. Qualitatively, it was evident that both effects peaked around centro-parietal areas (see Figure 27, page 122). These regions had no electrode coverage in the current study. Furthermore, under unimodal conditions the effects showed stronger left lateralisation compared with cross-modal effects. It is possible that this left lateralisation under unimodal conditions makes it possible to detect the opposing dipole of this effect at right frontal locations under unimodal conditions. However, as cross-model effects did not appear lateralised they may be less detectable at these electrode locations. In Chapter 5 cross-modal effects were also investigated using LRPs. These potentials are calculated using electrodes positioned over the motor cortex (C3/C4). However, using the electrode array in its intended form means calculation of these effects is not possible. Given the high ICCs between waveforms under both unimodal and cross-modal conditions, it is highly likely that the Emotiv system has the potential to detect both unimodal and cross-modal effects. However, the results of the current study highlight the importance of electrode location in identifying effects.

Limitations and Future Directions

The primary limitation in the current study was the delay in the event-marking system. This delay was surprising given the success of this system in previous studies (Badcock et al., 2013, 2015; de Lissa et al., 2015; Petit et al., in prep). As shown in Figure 31c, the event marking system required several wires to be connected to the Emotiv headset. Throughout testing, it was noted by the experimenter that these wires were susceptible to loosening between sessions and required reconnecting on several occasions. The fragility and need to reconnect the set-up may thus explain some inter-subject differences in event marking delay. A simple solution to this problem is therefore to ensure that the wiring is robust and stable throughout testing. Furthermore, it might be useful to employ calibration protocol prior to testing to ensure event markers are aligned. The current study used eye-movement artifacts between a research grade system and Emotiv to perform this calibration post-hoc. High ICCs between systems suggest this realignment procedure was successful. Although this may be less helpful for studies wishing to use standalone Emotiv EEG, the current protocol represents one approach to checking the timing of event markers prior to testing within a single pilot participant. This would then serve as a useful check prior to data collection and allow for correction of event timings post-hoc if necessary. However, if the current set-up is susceptible to wear and tear over time, this questions its appropriateness for use with large sample sizes with many testing sessions. It would therefore be ideal to characterise the stability of the event marking system between headsets and over multiple testing sessions.

A second challenge encountered was identifying approaches to reject artifacts from the Emotiv EEG. This caveat has been noted by previous studies, Petit et al (in prep) and Badcock et al (2015) both found ICA did not isolate artifacts clearly associated with eye blinks. These studies therefore did not reject eye-blink artifacts from the Emotiv EEG. Interestingly, these studies also suggest that eye blinks may not have been strong enough in the Emotiv data to be detected, whilst the current study suggests eye blink artifacts were larger in the Emotiv EEG. Originally, I attempted to use an automated artifact protocol to facilitate future replication and make pre-processing less time-consuming. However, these automated pipelines required either correlation with EOG channels or the topography of artifacts to identify eye blinks. The Emotiv headset does not contain EOG channels and the distribution of electrodes is mainly frontal. As such it is not possible to directly measure eye movements, and the majority of components will likely show frontal distribution (characteristic of eye movements) or a focus on a single or few electrodes (usually used to characterise electrode "pop-off"). One solution to this might be to have participants deliberately make a series of blinks and horizontal eye-movements prior to testing, thus allowing clearer characterisation of eye-movements for artifact rejection. For example, many participants in the current studies made clear horizontal eye movements associated with reading during the break. This was highly useful in the identification and rejection of artifacts during pre-processing. Another approach to enable artifact rejection in the Emotiv set-up would be to integrate EOG channels into the Emotiv system, allowing correlation with these channels.

Conclusion

In sum, the current findings suggest the Emotiv EPOC© portable EEG system is capable of measuring ERPs associated with unimodal Stroop interference (i.e. the N450 response). Significantly fewer trials were retained in the Emotiv system compared with the research grade EEG following pre-processing. Intra-class correlations (ICC) suggested strong correspondence between the two EEG systems. However, correspondence was lower at the most frontal electrodes and this may have been due to higher susceptibility to eye-movement artifacts. Although the current findings suggest Emotiv can be used as a low-cost method of gathering ERP data, two core caveats must be addressed. First, it is advised that studies conduct an initial calibration period to ensure the reliability of the wireless event-marking system. Second, steps must be taken to optimise the rejection of eye-movement artifacts. Potential avenues for this include inducing deliberate eye-movement artifacts to aid characterisation for manual rejection and incorporating EOG channels into the Emotiv headset.

Interim summary

You may recall that in this thesis I aimed to address three questions. In order of priority, these were:

1. Do unimodal and cross-modal interference arise from the same or different mechanisms?
2. Does a shift in sensory dominance occur across development?
3. Is sensory dominance associated with cross-modal interference?

The experimental chapters presented so far were aimed at addressing the first of these questions. However, in my introduction I also speculated a link between sensory dominance and interference control. This was because previous literature found children were more susceptible to auditory distraction when focusing on vision (Hanauer & Brooks, 2003, 2005), but not vice versa (Thomas et al., 2017). Conversely, in young adults (Donohue, Appelbaum, et al., 2013; Yuval-Greenberg & Deouell, 2009) and older adults (Guerreiro et al., 2010; Guerreiro & Van Gerven, 2011) vision influenced audition more than vice versa. In the remaining experimental chapters I addressed whether a shift in sensory dominance occurs across development, and whether individual differences in sensory dominance might be related to cross-modal interference. The next chapter provides a thorough systematic literature review to assess one measure of cross-modal interference, the Colavita effect. Following this, I present a study in which I exploited audio-visual illusions (the McGurk effect) to further test whether sensory dominance changes across development. Finally, I present an exploratory study in which I attempted to correlate sensory dominance measures with unimodal and cross-modal Stroop performance. All experimental chapters and their implication for future research will then be discussed in a general discussion

Chapter 7: Vision dominates audition in adults but not children: A meta-analysis of the Colavita effect

The Colavita effect occurs when participants respond only to the visual element of an audio-visual stimulus. This visual dominance effect is proposed to arise from asymmetric facilitation and inhibition between modalities. It has also been proposed that, unlike adults, children appear predisposed to auditory information. In this chapter, I provide the first quantitative synthesis of studies exploring the Colavita effect, combining data from 70 experiments across 14 studies. A mixed-meta-regression model was applied to assess whether the Colavita effect is influenced by methodological factors and age group tested. Studies reporting response time data were used to test for the presence of asymmetrical facilitation between modalities. Studies with adult participants yielded a medium, approaching large, effect size. Studies exploring the Colavita effect in children yielded no Colavita effect. Across adult and child studies, no methodological factors influenced the effect. Contrary to asymmetrical facilitation, response time data suggested a general slowing under bimodal conditions. These findings suggest that whilst vision dominates in adults, this effect is absent in childhood.

Introduction

Our world is perceived through multiple senses, but it is unclear whether information from all senses is treated equally. Whilst reading this thesis, are you more likely to be distracted by the sight of an email pop-up on your screen, or the sound of your phone ringing? Furthermore if your phone rings and an email pops-up simultaneously, which do you respond to first? The answer to these questions may lie with sensory dominance.

Colavita (1974; 1976) reported that when participants were presented with an auditory and a visual stimulus simultaneously they responded as though only the visual stimulus had occurred, and frequently reported having not perceived the auditory stimulus at all. This *Colavita effect* was found even when the auditory stimulus (a tone) was presented at twice the subjective intensity of the visual stimulus (a light), ruling out a simple explanation of physical inequality between the two modalities (Colavita, 1974). A *Colavita error* is defined as occurring when participants respond only to the visual element of a bimodal, in this case audio-visual, target. This effect has been used to imply a hierarchy of sensory processing in which visual information is given precedence.

Multiple studies have since replicated the Colavita effect, although the extent of the effect does appear to depend on the specific instructions given to participants. Studies conducted in the decade following the original study used two response keys and instructed participants to “make a response appropriate to the signal recognised first” (Colavita, 1982; Colavita & Weisberg, 1979; T. L. Johnson & Shapiro, 1989; Shapiro, Egerman, & Klein, 1984). These studies found Colavita “errors” to occur on a relatively large number of bimodal trials ranging from 38-98%. In these studies, however, participants were instructed to make only one response (to that which was recognised first) but it is possible that the participants still perceived both auditory and visual signals. More recent studies (Koppen & Spence, 2007a, 2007b, 2007c, 2007d) instructed participants to press both keys on bimodal trials. Although the number of visual-only responses was smaller in these studies (0.9-12.1%) these error rates remained significantly higher than auditory-only responses, thus demonstrating the Colavita effect.

In contrast, variations in other task manipulations do not appear to influence the Colavita effect. Qualitative reviews of the literature exploring visual precedence in adults (Spence, 2009; Spence, Parise, & Chen, 2012) have concluded the Colavita effect to be relatively insensitive to manipulations of stimulus intensity (Colavita, 1974; Shapiro & Johnson, 1987), attention bias to one or other modality created by the experiment

(Egeth & Sager, 1977; Koppen & Spence, 2007a, 2007c; Sinnett, Spence, & Soto-Faraco, 2007), response demands (Egeth & Sager, 1977; Hecht & Reiner, 2009; Koppen & Spence, 2007c; Sinnett et al., 2007) and stimulus complexity (Koppen, Alsius, & Spence, 2008; Sinnett et al., 2007). This suggests that visual precedence may have an origin beyond simply response bias. However, since the previous review was descriptive, and over ten large studies have been published since, a quantitative update of the review is essential. Therefore, the primary aim of the current study was to quantify how robust the Colavita effect is, and, furthermore, whether it can be manipulated by task demands or age group tested.

The additional factor of age may be of particular importance to the sensory dominance literature. Robinson and Sloutsky (2004) and Barnhart, Rivera, and Robinson (2018) assessed sensory dominance in 4 year olds and 5-12 year olds respectively. Findings from these studies suggested that visual dominance may develop across the lifespan and that children may be auditory dominant. Wille and Ebersbach (2016) suggest a shift occurring around 9 years of age, as they found 9-year-olds showed Colavita effects, albeit weaker than the effects seen in adults. Indeed, the auditory system undergoes substantial development in utero (Graven & Browne, 2008a) whereas the visual cortex undergoes lengthy, protracted development throughout childhood (Graven & Browne, 2008b). Consequently, children may rely less upon vision, and more upon audition, early in life. In line with this, it has been shown that young children struggle to ignore auditory information when focusing upon visual stimuli (Hanauer & Brooks, 2003) and children manifest smaller, sometimes reverse, Colavita effects (Nava & Pavani, 2013; Wille & Ebersbach, 2016). Given this, a comparison of the Colavita effect across studies using different age groups is of great theoretical interest.

A further aim of the current study was to explore the mechanisms underpinning the Colavita effect. Sinnett et al. (2008) proposed that the appearance of visual precedence is due to an asymmetrical inhibitory-facilitatory relationship between vision and audition (Sinnett, Soto-Faraco, &

Spence, 2008). Sinnett et al. (2008) report that, in simple detection tasks (using a single key), presenting auditory and visual stimuli together facilitated response times. Conversely, in discrimination tasks (using multiple keys), presenting auditory and visual stimuli together impeded response times. In a second experiment, using a simple detection task, they found that auditory stimuli facilitated response times to visual targets whilst visual stimuli impaired response times to auditory targets. These opposing effects have been used to infer an asymmetrical inhibitory-facilitatory relationship between audition and vision.

Sinnett et al. (2008) propose that this asymmetrical relationship might result in Colavita errors, as when participants are presented with bimodal targets the ‘internal threshold’ for responding to visual targets is reached sooner than auditory targets (Spence, 2009). Thus visual processing interferes with, and delays, auditory target detection and speeded responses are most likely to be visual-only responses (Spence, 2009). This hypothesis is supported by event-related potential (ERP) data showing ERPs to audio-visual stimuli occur at an increased latency relative to auditory only ERPs and a decreased latency relative to visual only ERPs (Molholm et al., 2002).

On the other hand, previous literature has suggested vision facilitates audition and vice versa. In simple response time tasks (using one response key) response times to bimodal targets are typically faster than unimodal targets (the redundant target effect; Diederich & Colonius, 2004; Forster, Cavina-Pratesi, Aglioti, & Berlucchi, 2002; Gondan, Niederhaus, Rösler, & Röder, 2005; Sinnett et al., 2008). Furthermore, detection thresholds for luminance appear lower (Frassinetti et al., 2002), and the saliency (Noesselt, Bergmann, Hake, Heinze, & Fendrich, 2008) and perceived brightness (Odgaard, Arieh, & Marks, 2003) of visual events increases with simultaneous sound. Similarly, irrelevant visual stimuli can enhance auditory detection (Lovelace, Stein, & Wallace, 2003) and increase the perceived loudness of simultaneously presented sounds (Odgaard, Arieh, & Marks, 2004). However, Odgaard, Arieh and Marks (2004) suggest different processes may underpin facilitation between modalities, as the effect of audition upon vision might arise from decisional processes, whilst the effect of vision upon audition may hold sensory origin.

A general, symmetrical, model of multisensory facilitation is consistent with additivity, whereby neural responses elicited from bimodal targets are greater than responses to unimodal elements (Meredith & Stein, 1986). However, asymmetrical effects upon response times are not necessarily incompatible with additivity. For example, although visual and auditory evoked ERPs are asymmetrically influenced by one another with respect to latency, the amplitude of ERPs to audio-visual stimuli are greater than the sum of both unimodal auditory and unimodal visual responses (Molholm et al., 2002). However, it has yet to be established how physiological models of multisensory integration can accommodate asymmetries in cross-modal influences.

Given the mixed literature regarding symmetrical vs. asymmetrical inhibition and facilitation between vision and audition, I aimed to test this within the existing Colavita literature. The hypothesis of Sinnett and colleagues (2008) is based upon findings from a simple detection task (using one response key). In contrast to this, many Colavita studies have utilised multiple response keys. Sinnett and colleagues note that with multiple response keys slowing can be observed. As such, it was assessed whether asymmetrical response time effects are observed within the wider Colavita literature, in which multiple response keys were sometimes used.

This study provides the first quantitative synthesis of literature exploring the Colavita effect. The primary objectives were to *a)* quantify how robust the Colavita effect is (i.e. making a unimodal visual response when bimodal stimuli are presented), *b)* test whether the Colavita effect is sensitive to experimental manipulations and age, and *c)* use available response time data to assess the presence of symmetrical vs. asymmetrical facilitation between audition and vision. Given the specific predictions provided by Sinnett and colleagues regarding auditory vs. visual modalities, and the audio-visual nature of the Colavita effect in original reports (Colavita, 1974), I focused on studies comparing auditory vs. visual modalities. Nevertheless it should be noted that the Colavita effect has been extended to the visual-tactile domain (Hartcher-O'Brien, Levitan, & Spence, 2010; Hecht & Reiner, 2009; Occelli, Hartcher-O'Brien, Spence, & Zampini, 2010). By including data from multiple studies this analysis overcomes some of the limitations of individual studies. Small sample sizes have been used in many cases and effect sizes vary. For instance, Colavita's early (1974; 1976; 1979) experiments contained very few participants (n=10) and trials (35 trials per participant, 5 bimodal).

To allow comparison between the present quantitative review and the qualitative review by Spence (2009) variables highlighted by Spence (2009) were included as potential moderator variables. Specifically, it was predicted that the Colavita effect would be insensitive to manipulations of:

- Number of response keys (2 or 3). Note that studies including only a single response key were considered for the response time analysis only as Colavita errors cannot be made with a single response key.
- Ratio of visual, auditory and bimodal targets (and in one case no target present[26]).
- Attentional manipulation: was attention biased towards the visual or auditory modality either through arousal, cueing, perceptual biasing (if the light was twice the subjective intensity of the sound), or via instructional manipulation (participants asked to attend to or respond only to auditory information).
- Stimulus category: simple (i.e. tones and lights) vs. complex (i.e. pictures/videos and natural sounds).
- Whether auditory and visual stimuli were perceptually matched in intensity (either subjectively or based upon thresholds).
- Stimulus congruency: A stimulus could be "congruent" semantically, e.g. a picture of a cat and the sound of a cat, or spatially, e.g. a visual stimulus on the left and a sound on the left.

---

[26] Koppen, Levitan and Spence (2009) had four trial types using a ratio of 25 visual: 25 auditory: 25 bimodal: 25 no target present. On the no target present trials no key should have been pressed.

Furthermore, the comparisons were extended to include:

- Age group: child vs. adult. A reduced Colavita effect was predicted in children.
- Asymmetric facilitation and inhibition. Studies were included that used Colavita tasks and also reported response times to test the prediction of Sinnett et al. (2008); that response times to visual stimuli are faster under bimodal conditions, whilst response times to auditory stimuli are slower under bimodal conditions.

Method

Search and inclusion criteria

Studies were retrieved and selected using the guidelines outlined in PRISMA (Moher, Liberati, Tetzlaff, & Altman, 2009). Figure 36 outlines the search strategy used. Studies were found by searching the electronic databases Scopus, PubMed and Web of Science (July 2016- August 2017) and reviewing the references of studies sourced. Initial search terms included: *Colavita effect* (64 hits across all data-bases), *Colavita* (362 hits across all data-bases) and *sensory dominance* (256 hits across all data-bases). The following inclusion criteria were then applied:

- Studies using a choice response time task to compare responses to unimodal and bimodal stimuli in humans (Figure 36; box b).
- Studies comparing responses to auditory, visual and audio-visual targets (Figure 36; box c).
- Studies available to the author in English (Figure 36; box c).
- Sources in which full text could be sourced (i.e. meeting abstracts and posters excluded - Figure 36; box c).
- Studies where error data and/or response time data for bimodal (audio-visual) stimuli could be sourced (either within the paper or via personal communication with the author - Figure 36; box d). Notably, because response time analyses were performed to examine the effect of vision on audition and vice versa, response time data needed to be available for unimodal visual targets, unimodal auditory targets,

visual targets in the presence of auditory stimuli and/or auditory targets in the presence of visual stimuli. Thus, response times for bimodal targets were not used for analyses unless participants were asked to make separate responses for visual and auditory elements of the target (i.e. press both keys).

- Studies conducted upon healthy participants (children and adults). For example in two cases data was sought from the healthy control group of larger studies (Moro & Steeves, 2012, 2013).

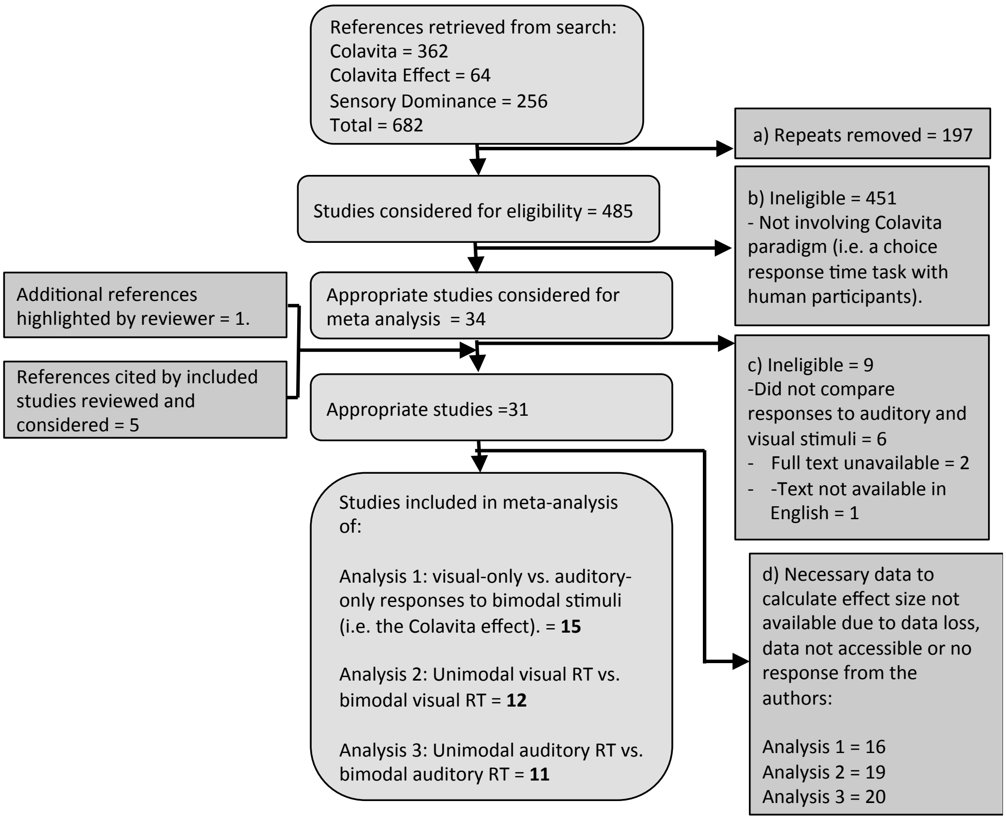


Figure 36. Flow diagram illustrating search strategy and exclusion criteria used to isolate studies to be included in the meta-analysis.

Many of the studies sourced included multiple experiments, each containing its own conditions/comparisons. For example, Wille and Ebersbach (2016) conducted three experiments each containing three age groups, in which three levels of congruency were explored – thus providing 27 experiments for the purposes of the current analysis. By breaking down each study into its component experiments a total of 125 experiments were available for analysis. Details of these studies can be found in Table 10.

Of the studies and experiments available, only those that provided sufficient information for the calculation of effect size data were included to explore the following dependant variables:

1. The overall Colavita effect as defined in Equation 8, where Vb refers to the percentage of visual-only responses made on bimodal trials and Ab refers to percentage auditory-only responses made on bimodal trials (15 studies, 71 experiments). Note that ratio scores were used in order to place the effects observed in all studies on the same scale (i.e. a study yielding 60% “visual only” responses and 20% “auditory only” responses shows the same level of visual dominance over audition as a study with 6% “visual only” vs. 2% “auditory only”).

Equation 8

$$Colavita\ effect = Vb/Ab$$

2. Response times to unimodal visual targets vs. visual targets paired with an auditory stimulus (12 studies, 28 experiments).
3. Response times to unimodal auditory targets vs. auditory targets paired with a visual stimulus (11 studies, 25 experiments).

Author	Exp.	n	Age		Stimuli			Keys (n)	Instructions on bimodal trials	Attentional Manipulation	Congruency	Possible to calculate effect sizes for:		
												Colavita	Unimodal vs. Bimodal RT	
			Group	M	Type	Matched	Ratio						Vis.	Aud.
Osborn, Sheldon, & Baker, 1963	1	41	Adult	20.2	Simple	N	33:33:33	2	Press both keys	-	-	✖	✖	✖
Colavita, 1974	1	10	Adult	-	Simple	Y	43:43:14	2	No specific instructions	-	-	✖	✖	✖
	2	22				N				A (tone twice subjective intensity of light)		✖	✖	✖
	3	10				Y			Press whichever key appropriate for the signal recognised first	-		✖	✖	✖
	4	10				Y			Press the tone key	A (respond with tone key)		✖	✔	✔
Colavita, Tomko, & Weisberg, 1976	1	10	Adult	-	Simple	N	43:43:14	2	Press whichever key appropriate for the signal recognised first	-	-	✖	✖	✖
	2	10								A (fixation moved away from light source)		✖	✖	✖
Colavita & Weisberg, 1979	1	10	Adult	-	Simple	Y	38:38:23	2	Press the tone key when the tone terminates and the light key if the light terminates	-	-	✖	✔	✔
Egeth & Sager, 1977	1	10	Adult	-	Simple	Y	43:43:14	2	Press both keys	-	-	✖	✖	✖
	2	16						1	Press the tone key whenever you hear a	A (respond with tone	-	✖	✖	✖

Author	Exp.	n	Age		Stimuli			Keys (n)	Instructions on bimodal trials	Attentional Manipulation	Congruency	Possible to calculate effect sizes for:		
												Colavita	Unimodal vs. Bimodal RT	
			Group	M	Type	Matched	Ratio						Vis.	Aud.
									tone	key)				
Egeth & Sager, 1977 (continued)	3.1	10	Adult	-	Simple	Y	20:40:40	1		A (respond with tone key, and more auditory targets)	-	✗	✗	✗
	3.2	10						2	Respond only to the tone	A (respond with tone key, and more auditory targets)	-	✗	✗	✗
	3.3	10					40:40:20	1	Press the tone key whenever you hear a tone	A (respond with tone key)	-	✗	✗	✗
	3.4	10						2	Respond only to the tone	A (respond with tone key)	-	✗	✗	✗
	4.1	10							Press the light key	V (respond with light key)	-	✗	✓	✗
	4.2						20:40:40				-	✗	✓	✗
	5.1	10					40:40:20		Press the tone key	A (respond with tone key)	-	✗	✗	✓
	5.2									A (instructed to attend to auditory info)	-	✗	✗	✓

Author	Exp.	n	Age		Stimuli			Keys (n)	Instructions on bimodal trials	Attentional Manipulation	Congruency	Possible to calculate effect sizes for:		
												Colavita	Unimodal vs. Bimodal RT	
			Group	M	Type	Matched	Ratio						Vis.	Aud.
Egeth & Sager, 1977 (continued)	6.1	12	Adult	-	Simple	Y	40:40:20	2		A (respond with tone key)	-	✖	✖	✔
	6.2					N (light half subjective intensity)	40:40:20	2		A (light half intensity of tone)	-	✖	✖	✔
	6.3					N (light twice subjective intensity)				V (light twice intensity of tone)	-	✖	✖	✔
Colavita, 1982	1	10	Adult		Simple	Y	39:39:22	2	Press whichever key is appropriate to the signal you recognise first	-	-	✖	✖	✖
Shapiro, Egerman, & Klein, 1984	1.1	16	Adult	-	Simple	N	40:40:20	2	Respond to the stimulus first perceived- informed that shock would occur if too fast or slow.	-	-	✖	✖	✖
	1.2	16							Respond to the stimulus first perceived	-	-	✖	✖	✖
	2.1	16					45:45:10		Respond to the stimulus first perceived- informed that shock would occur if too fast or slow.	-	-	✖	✖	✖
	2.2	16							Respond to the	-	-	✖	✖	✖

Author	Exp.	n	Age		Stimuli			Keys (n)	Instructions on bimodal trials	Attentional Manipulation	Congruency	Possible to calculate effect sizes for:		
												Colavita	Unimodal vs. Bimodal RT	
			Group	M	Type	Matched	Ratio						Vis.	Aud.
									stimulus first perceived					
Shapiro, Egerman, & Klein, 1984 (continued)	3.1	16	Adult	-	Simple	N	45:45:10	2	Respond to the stimulus first perceived- informed that shock would occur if too fast or slow. Participants also received additional tactile stimulus with shock.	-	-	✖	✖	✖
	3.2	16							Respond to the stimulus first perceived	-	-	✖	✖	✖
Johnson & Shapiro, 1989	1.1	12	Adult	-	Simple	N	40:40:20	2	Respond to the stimulus you detect first-Visual stimulus appeared in random location.	-	-	✖	✖	✖
	1.2	10							Respond to the stimulus you detect first- Visual stimulus appeared in same location on all trials.	-	-	✖	✖	✖
	1.3	12							Respond to the stimulus you detect first- Subjects received random, infrequent shocks following trials. Visual		-	✖	✖	✖

Author	Exp.	n	Age		Stimuli			Keys (n)	Instructions on bimodal trials	Attentional Manipulation	Congruency	Possible to calculate effect sizes for:		
												Colavita	Unimodal vs. Bimodal RT	
			Group	M	Type	Matched	Ratio						Vis.	Aud.
									stimulus appeared in random location.					
Johnson & Shapiro, 1989 (continued)	1.4	10	Adult	-	Simple	N	40:40:20	2	Respond to the stimulus you detect first- subjects received random, infrequent shocks following trials Visual stimulus appeared in same location.	-	-	✖	✖	✖
Koppen & Spence, 2007a	1	14	Adult	24	Simple	N	40:40:20	2	Press both keys	-	-	✔	✔	✔
	2	14		20			33:33:33	2		-	-	✔	✔	✔
	3	12		25			40:40:20	3	Press separate key	-	-	✔	✖	✔
	4.1	20		24				2	Press both keys	Total - visual cues and auditory cues	-	✖	✖	✖
	4.2									V (visual cue)	-	✔	✔	✔
	4.3									A (auditory cue)	-	✔	✔	✔
Sinnett, Spence, & Soto-Faraco, 2007	1	24	Adult	-	Complex	N	40:40:20	3	Press a separate key	-	-	✖	✖	✖
	2	54								-	-	✖	✖	✖
	4	24					33:33:33	3		-	-	✖	✖	✖

Author	Exp.	n	Age		Stimuli			Keys (n)	Instructions on bimodal trials	Attentional Manipulation	Congruency	Possible to calculate effect sizes for:		
												Colavita	Unimodal vs. Bimodal RT	
			Group	M	Type	Matched	Ratio						Vis.	Aud.
Sinnett, Spence, & Soto-Faraco, 2007 (continued)	5.1	18	Adult	-	Complex	N	60:20:20			V (more visual targets)	-	✖	✖	✖
	5.2	18					20:60:20			A (more auditory targets)	-	✖	✖	✖
	6.1	18					40:40:20			A (low auditory load)	-	✖	✖	✖
	6.2	18								V (low visual load)	-	✖	✖	✖
Koppen & Spence, 2007b	1	18	Adult	22	Simple	N	20:20:60	2	Press both keys	-	-	✔	✔	✖
	2.1	12					5:5:90			-	-	✔	✔	✖
	2.2						25:25:50			-	-	✔	✔	✖
	2.3						45:45:10			-	-	✔	✔	✖
Koppen & Spence, 2007c	1.1	36	Adult	23	Simple	N	40:40:20	2	Press both keys	-	C (same position (13°))	✔	✔	✔
	1.2									-	I (different position (13°))	✔	✔	✔
	1.3									-	C (same position (26°))	✔	✔	✔
	1.4									-	I (different	✔	✔	✔

Author	Exp.	n	Age		Stimuli			Keys (n)	Instructions on bimodal trials	Attentional Manipulation	Congruency	Possible to calculate effect sizes for:		
												Colavita	Unimodal vs. Bimodal RT	
			Group	M	Type	Matched	Ratio						Vis.	Aud.
											position (26°))			
Koppen & Spence, 2007d	1	22	Adult	23	Simple	N	40:40:20	2	Press both keys – stimuli presented at various stimulus onset asynchrony's-overall errors considered for analysis.	-	-	✖	✖	✖
Koppen, Alsius, & Spence, 2008	1.1	12	Adult	21	Complex	N	40:40:20	2	Press both keys	-	C	✓	✓	✓
	1.2									-	I	✓	✖	✖
	2.1	30		21				3	Press separate key	-	C	✓	✖	✖
	2.2									-	I	✓	✖	✖
	3.1	15		23	Complex					-	C	✓	✖	✖
	3.2									-	I	✓	✖	✖
Sinnett, Soto-Faraco, & Spence, 2008	1.1	22	Adult	-	Complex	N	40:40:20	3	Press separate key	-	-	✖	✖	✖
	1.2							1	Press single key to all targets			✖	✖	✖
	2	20							Press single key to all targets	-	-	✖	✓	✓
Koppen, Levitan, & Spence, 2009	1	22	Adult	20	Simple	Y	25:25:25:25(no target)	2	Press both keys	-	-	✖	✖	✖
Hecht & Reiner, 2009	1	12	Adult	24.6	Simple	N	40:40:20	3	Press separate key	-	-	✖	✖	✖
Van Damme, Crombez, &	1.1	20	Adult	19.5	Simple	N	40:40:20	3	Press separate key	Visual threat baseline	-	✓	✖	✖

Author	Exp.	n	Age		Stimuli			Keys (n)	Instructions on bimodal trials	Attentional Manipulation	Congruency	Possible to calculate effect sizes for:		
			Group	M	Type	Matched	Ratio					Colavita	Unimodal vs. Bimodal RT	
													Vis.	Aud.
Spence, 2009														
Van Damme, Crombez, & Spence, 2009 (continued)	1.2	20	Adult	19.5	Simple	N	40:40:20	3	Press separate key	V (visual threat – received infrequent shocks following visual stimuli)	-	✓	✖	✖
	1.3	20								Auditory threat baseline	-	✓	✖	✖
	1.4									A (auditory threat- received infrequent shocks following auditory stimuli)	-	✓	✖	✖
	1.5	17								Control baseline	-	✓	✖	✖
	1.6									Control	-	✓	✖	✖
Ngo, Sinnett, Soto-Faraco, & Spence, 2010	1	24	Adult	28	Complex	N	40:40:20	3	Press separate key- detect immediate repetition	-	-	✓	✖	✖
Ngo et al.,	1	20	Adult	28	Comple	N	40:40:20	3	Press separate key-	-	-	✓	✖	✖

Author	Exp.	n	Age		Stimuli			Keys (n)	Instructions on bimodal trials	Attentional Manipulation	Congruency	Possible to calculate effect sizes for:		
												Colavita	Unimodal vs. Bimodal RT	
			Group	M	Type	Matched	Ratio						Vis.	Aud.
2011					x				respond to repeated (n-1) targets					
Ngo et al., 2011 (continued)	2	18	Adult	27	Complex	N	40:40:20	3	Press separate key-respond to repeated (n-1) targets	-	-	✓	✗	✗
	3	16		24						-	-	✓	✗	✗
	4	24		25					Press separate key-detect immediate repetition	-	-	✓	✗	✗
Moro & Steeves, 2012	1	11*	Adult	24.6	Complex	N	40:40:20	1	Press one key to all targets	-	-	✗	✗	✗
	2							3	Press separate key for bimodal targets	-	-	✓	✗	✗
Moro & Steeves, 2013	1	11*	Adult	28.3	Complex	N	40:40:20	1	Press one key to all repeats	-	-	✗	✗	✗
	2							3	Press separate key for bimodal repeats	-	-	✓	✗	✗
Stubblefield, Jacobs, Kim, & Goolkasian, 2013	2.1	35	Adult	24	Complex	N	40:40:20	1	Press single key to predefined visual target – Incongruent auditory distractor present.	V	-	✗	✓	✗
	2.2								Press single key to predefined auditory target- Incongruent visual distractor present.	A	-	✗	✗	✓

Author	Exp.	n	Age		Stimuli			Keys (n)	Instructions on bimodal trials	Attentional Manipulation	Congruency	Possible to calculate effect sizes for:		
												Colavita	Unimodal vs. Bimodal RT	
			Group	M	Type	Matched	Ratio						Vis.	Aud.
Stubblefield, Jacobs, Kim, & Goolkasian, 2013 (continued)	4.1	31		20	Complex	N	40:40:20	1	Press single key to predefined visual target-Incongruent auditory distractor present.	V	-	✖	✔	✖
	4.2								Press single key to predefined auditory target-Incongruent visual distractor present.	A	-	✖	✖	✔
Nava & Pavani, 2013	1.1.1	14	Child	6.8	Simple	N	40:40:20	2	Press both keys	-	-	✔	✔	✔
	1.1.2	14		9.5						-	-	✔	✔	✔
	1.1.3	13		11.7						-	-	✔	✔	✔
	1.2	10		6.9			33:33:33			-	-	✔	✔	✔
Tsukiko & Desmarais, 2014	1	31	Adult	19	Complex	N	33:33:33	3	Press separate Key	-	-	✖	✖	✖
Yue, Jiang, Li, Wang, & Chen, 2015	1.1	28	Adult	22.7	Simple	N	40:40:20	2	Press both keys	-	C (both near)	✔	✔	✔
	1.2									-	C (both far)	✔	✔	✔
	2.1	20		22.6						V	I (auditory far - visual	✔	✔	✔

Author	Exp.	n	Age		Stimuli			Keys (n)	Instructions on bimodal trials	Attentional Manipulation	Congruency	Possible to calculate effect sizes for:		
												Colavita	Unimodal vs. Bimodal RT	
			Group	M	Type	Matched	Ratio						Vis.	Aud.
											near)			
Yue, Jiang, Li, Wang, & Chen, 2015 (continued)	2.2	20	Adult	22.6	Simple	N	40:40:20	2	Press both keys	A	I (Visual far - auditory near)	✓	✓	✓
Wille & Ebersbach, 2016	1.1.1	28	Child	6.39	Complex N		33:33:33	3	Press a separate key	-	Total	✓	✖	✖
	1.1.2									-	I	✓	✖	✖
	1.1.3									-	C	✓	✖	✖
	1.2.1	30		8.87						-	Total	✓	✖	✖
	1.2.2									-	I	✓	✖	✖
	1.2.3									-	C	✓	✖	✖
	1.3.1	28	Adult	25.96						-	Total	✓	✖	✖
	1.3.2									-	I	✓	✖	✖
	1.3.3									-	C	✓	✖	✖
	2.1.1	19	Child	6.42			25:25:50			-	Total	✓	✖	✖
	2.1.2									-	I	✓	✖	✖
	2.1.3									-	C	✓	✖	✖

Author	Exp.	n	Age		Stimuli			Keys (n)	Instructions on bimodal trials	Attentional Manipulation	Congruency	Possible to calculate effect sizes for:		
												Colavita	Unimodal vs. Bimodal RT	
			Group	M	Type	Matched	Ratio						Vis.	Aud.
Wille & Ebersbach, 2016 (continued)	2.2.1	17	Child	9.06	Complex	N	25:25:50	3	Press a separate key	-	Total	✓	✗	✗
	2.2.2									-	I	✓	✗	✗
	2.2.3									-	C	✓	✗	✗
	2.3.1	23	Adult	27.61						-	Total	✓	✗	✗
	2.3.2									-	I	✓	✗	✗
	2.3.3									-	C	✓	✗	✗
	3.1.1	18	Child	6.39						-	Total	✓	✗	✗
	3.1.2									-	I	✓	✗	✗
	3.1.3									-	C	✓	✗	✗
	3.2.1	24		9.13						-	Total	✓	✗	✗
	3.2.2									-	I	✓	✗	✗
	3.2.3									-	C	✓	✗	✗
	3.3.1	20	Adult	28.7						-	Total	✓	✗	✗
	3.3.2									-	I	✓	✗	✗

Author	Exp.	n	Age		Stimuli			Keys (n)	Instructions on bimodal trials	Attentional Manipulation	Congruency	Possible to calculate effect sizes for:		
												Colavita	Unimodal vs. Bimodal RT	
			Group	M	Type	Matched	Ratio						Vis.	Aud.
Wille & Ebersbach, 2016 (continued)	3.3.3	20	Adult	28.7	Complex	N	25:25:50	3	Press a separate key	-	C	✓	✗	✗
Monem & Fillmore, 2016	1.1	25	Adult	25.4	Complex	N	33:33:33	1	Press the key whenever you see a visual, auditory or audio-visual target (alcohol)	-	-	✗	✓	✗
	1.2								Press the key whenever you see a visual, auditory or audio-visual target (office supplies)			✗	✓	✗
	2.1							3	Press a separate key whenever you see a target (office supplies)			✓	✗	✗
	2.2								Press a separate key whenever you see a target (alcohol)			✓	✗	✗
Stekelenburg & Keetels, 2016	1	20	Adult	21.1	Simple	N	33:33:33	3	Press separate key	-	-	✗	✗	✗
Huang et al., 2015	1	20	Adult	-	Simple	N	40:40:20	2	Press both keys	-	-	✗	✗	✗
	2	24								-	-	✗	✗	✗
	3	17								-	-	✗	✗	✗
Li et al., 2017	1	18	Adult	-	Simple	N	40:40:20	2	Press both keys	-	-	✓	✓	✓

Table 10. Details of experiments considered for analysis broken down by experiment and condition. Tick boxes indicate whether details necessary for the calculation of Cohen's d_{av} were available (i.e. sample size, mean and standard deviation or standard error). Abbreviations within the "Attentional manipulation" and "Congruency" columns are as follows; C=Congruent, I=Incongruent, V=visual, A=auditory. If nothing is stated then this was either not manipulated or not reported within the obtained article. * Value indicates n for healthy control condition. For example, Moro and Steeves (2012; 2013) both included 11 participants who had undergone monocular enucleation, these participants were not included.

Statistical Analyses

Effect sizes were calculated for the percentage visual-only vs. auditory-only errors on bimodal trials (Colavita and reverse Colavita effects) as well as response times under unimodal visual vs. bimodal visual and unimodal auditory vs. bimodal auditory conditions.

Calculation[27] of weighted effect sizes (see below) and model fitting was conducted using the metafor package in R (Viechtbauer, 2010). Cohen's guidelines of 0.2, 0.5, and 0.8 were used to define small, medium and large effect sizes for descriptive purposes. Given the wide range of contexts under which the Colavita effect has been explored, a *random effects* rather than a *fixed effects* meta-regression model was applied (Thompson & Higgins, 2002). Furthermore, the majority of studies included reported a range of differences in experimental procedure. As such these factors were held as moderator variables to explore whether they could account for the variance of effect size between studies.

Outliers

In line with the guidelines outlined by Cook and Weisberg (1984) and Viechtbauer and Cheung (2010), outliers and influential cases were identified and examined if:

a) The absolute DFFITS value was larger than $3\sqrt{p/(k-p)}$ where p is the number of model coefficients and k the number of studies, suggesting the average effect size to be influenced by inclusion of *i*th study.
b) Cook's distance exceeded $X^2_{p,0.5}$, indicating the mahalanobis distance between studies to be decreased following the deletion of *i*th study.
c) The study was shown to have considerable leverage upon the fit of the model based upon a hat value larger than $3(p/k)$.

---

[27] Script available at https://osf.io/d7b3d/.

For further information on these parameters see Viechtbauer and Cheung (2010). Combined effect sizes are shown including and excluding influential studies. These studies were not included within the modelling of moderator variables.

Calculation of effect sizes

Measures of effect size were calculated using Hedges (g_{av}), derived using Cohen's d_{av} where the average standard deviation of both sets of observations (S_{av}) is used as a standardizer (Cumming, 2012; Cumming & Calin-Jagerman, 2017; Lakens, 2013)[28].

Equation 9

$$Cohens\ d_{av} = M_{diff} / \sqrt{(}\frac{SD_1^2 + SD_2^2}{2})$$

It is acknowledged that this is not the optimal measure of effect size for studying within-subject phenomena. Alternative effect size measures, such as Cohen's d_{rm} (see Lakens, 2013) take into account the correlation (*r*) between measures. However, although *r* is typically reported for clinical pre-post test designs, *r* is not always reported in experimental designs where trials are intermixed and correlation is not of primary interest (Dunlap, Jose, Vaslow, & Burke, 1996). Thus, unless raw data can be obtained, *r* is not always available. Few solutions to this problem have been suggested. Borenstein et al. (2009) suggested estimating the correlation based upon related studies and performing sensitivity analyses with a range of plausible correlations. Alternatively, *r* can be estimated from available *t* and *f* statistics (Hullett & Levine, 2003). However if these exact statistics are also unavailable one may need to estimate effect size directly from the means and standard deviations (Dunlap et al., 1996). Cohen's d_{av} provides a convenient solution to this problem.

A further issue occurs, however, when calculating the variance around Cohen's d_{av}. Cumming (2012) proposes Algina and Keselman's (2003)

---

[28] The equation used here is taken from Cumming and Calin-Jagerman (2017) but is also referred to as the common language effect size (*Z*) (Lakens, 2013; McGraw & Wong, 1992).

approximate method for the calculation of confidence intervals (Equation 10), and subsequently variance (Equation 11), for Cohen's d_{av}. This method still requires knowledge of *r*. [29]

Equation 10

$$Cohens\ d_{av_{CI}} = d_{av} \pm t_{(1-\frac{a}{2},n-1)} \sqrt{\frac{(2(SD_1^2 + SD_2^2 - 2r)}{n(SD_1^2 + SD_2^2)}}$$

Equation 11

$$Vd_{av} = (\frac{CIup = CIlow}{2*1.96})^2$$

Thus if the researcher is unable to derive *r* from the available information similar problems are faced when calculating the variance of Cohen's d_{av}.

To resolve this problem I utilised a method adapted from the calculation of variance for Cohen's d for independent samples (Equation 12) where n_1 and n_2 signify the number of observations contributing towards M_{diff}.

Equation 12

$$Vd_{av} = (\frac{n_1 + n_2}{n_1 n_2}) + (\frac{d_{av}^2}{2(n_1 + n_2)})$$

Note this is a conservative method yielding marginally wider confidence intervals, relative to Algina and Keselman's (2003) approximate method, and thus assuming slightly greater variance. Where possible, Vd_{av} was also calculated using Equation 11 to estimate the true extent of the effect. For experiments studying the Colavita effect only 26 of the 71 experiments to be included contained sufficient information for calculation of *r*. In all of these cases the proposed method proved to be more conservative; the mean variance was 0.114 (*SD* = 0.05) when calculated using Equation 12 vs. 0.073 (*SD* = 0.03) when calculated using the approximate method outlined in Equation 11 with knowledge of *r*.

---

[29] Notation used by Algina and Keselman (2003) changed to be in line with current notation.

Whilst Cohen's d_{av} is the most appropriate method for sample estimates, it may be positively biased for population estimates. For this reason a corrected Cohen's d_{av}, Hedges g_{av} was calculated using Equation 6. Whilst the differences between d_{av} and g_{av} are very small, g_{av} provides an unbiased estimate of effect size (see Cumming, 2012).

Equation 13

$$Hedges\ \mathrm{g}_{av} = d_{av} * (1 - (\frac{3}{4 * n - 1} - 1))$$

To summarise, Hedges g_{av} (Equation 13) was used as the effect size measure within the current analysis. The variance of g_{av} was calculated using Equation 12, in which d_{av} was substituted with g_{av}.[30]

Moderator variables

Given the range of contexts in which the Colavita effect has been explored the studies included in this meta-analyses were heterogeneous in terms of the methods used. As such the following 8 factors were explored by including them as moderator variables within a mixed-effects model of the data:

- Number of response keys (2 or 3). Note that studies including only a single response key were considered for the response time analysis only, as Colavita errors cannot be made with a single response key.
- Ratio of visual, auditory and bimodal targets (and in one case no target present).
- Age group: child vs. adult.
- Stimulus category: simple (i.e. tones and lights) vs. complex (i.e. pictures/videos and natural sounds).
- Whether auditory and visual stimuli were perceptually matched in intensity (either subjectively or based on thresholds).
- Stimulus congruency: stimuli could be "congruent" semantically, e.g. picture of a cat and the sound of a cat, or spatially, e.g. a visual

---

[30] Spreadsheet allowing replication of effect size calculation available at https://osf.io/d7b3d/.

stimulus on the left and a sound on the left. Likewise stimuli could be "incongruent" semantically, e.g. a picture of a cat and sound of a dog, or spatially, e.g. visual stimulus on the left auditory stimulus on the right.

- Attentional manipulation: was attention biased towards the visual or auditory modality either through arousal, cueing, perceptual biasing (e.g. if the light was twice the subjective intensity of the sound) or via instructional manipulation (e.g. participants asked to attend to or respond only to auditory information).

Results

Error data analyses: The Colavita effect

Figure 37 illustrates the effect size of the Colavita effect in each experiment within each study. Positive effect sizes indicate more "visual only" responses on bimodal trials. Conversely, experiments with negative effect sizes found more "auditory only" responses on bimodal trials. The combined effect size estimate reached Cohen's standard for a small effect size, 0.44 (*SE* = 0.1), but was significant ($p < .001$). This suggests that participants made more visual-only responses under bimodal stimulus presentation than auditory-only responses. One experiment (Monem & Filmore, 2016, Experiment 1.2.1) was identified as an influential case. Removal of this experiment decreased the overall effect size to 0.4 (*SE* = 0.09), however this was still significant ($p < .001$).

To explore the effects of moderator variables a mixed meta-regression model was conducted in which the intercept was set to reflect the effect size of studies using the most frequently used experimental parameters (adult participants, simple stimuli that were neutral in congruency and attentional manipulation, a trial ratio of 40 (visual): 40 (auditory): 20 (bimodal), 2 response keys). All studies included in this analysis presented stimuli at fixed intensities.

The estimated amount of residual heterogeneity in this meta-regression model (tau^2 = 0.23, *SE* = 0.06), suggested that the included moderator variables accounted for 42.54 % of the variability. This was significant based upon an omnibus test (QM(12) = 47.46, $p < .001$). The intercept significantly differed from 0 ($p < .001$) with an effect size estimate of 0.79 (*SE* = 0.15). Only one factor, age group, significantly influenced this effect size estimate ($p < .001$) suggesting that experiments with child participants (aged 6-12 years) decreased this effect size by 0.89 (*SE* = 0.18). Six separate ANOVAs were then conducted to clarify the effect of each factor upon the intercept. These ANOVAs supported the mixed model indicating that only age group influenced the effect size of the Colavita effect (see Table 11). It should be noted, however, that a test for residual heterogeneity was also significant (QE(56) = 211.66, $p < .001$), suggesting other factors not accounted for in this model are also likely to be important.

Factor	**df**	**QM**	***p***
Ratio	5	02.856	0.722
Response keys	1	00.008	0.9297
Stimulus category	1	00.243	0.6176
Congruency	2	00.287	0.8664
Attentional manipulation	2	00.085	0.9583
Age group	1	23.609	<.0001

Table 11. Statistics resulting from additional analyses of variance (ANOVAs) exploring the effect of each factor upon the intercept of the mixed model (i.e. the overall effect size of the Colavita effect). One factor, age group, significantly influenced the effect size of the Colavita effect. df = degrees of freedom, QM = omnibus test statistic.

Effect of age group

A further model was fitted to directly compare the effect sizes of studies using adult and child participants (regardless of other factors). For details of studies included in this comparison see Table 10, column 4 labelled Age group. Unlike the model described above, here studies using all types of ratio and stimuli (rather than only “typical” parameters) were included. This model indicated that the effect size significantly differed from zero in adults (*M* = 0.76, *SE* = 0.09; $p < .001$) but not children (*M* = -0.26, *SE* = 0.13; *ns*). Experiments investigating the Colavita effect in children yielded an effect size that did not differ from 0 and was significantly smaller than the effect size found for experiments investigating the Colavita effect in adults ($p < .001$).

Thus, although children appeared to show a small reverse Colavita effect, this did not reach significance.

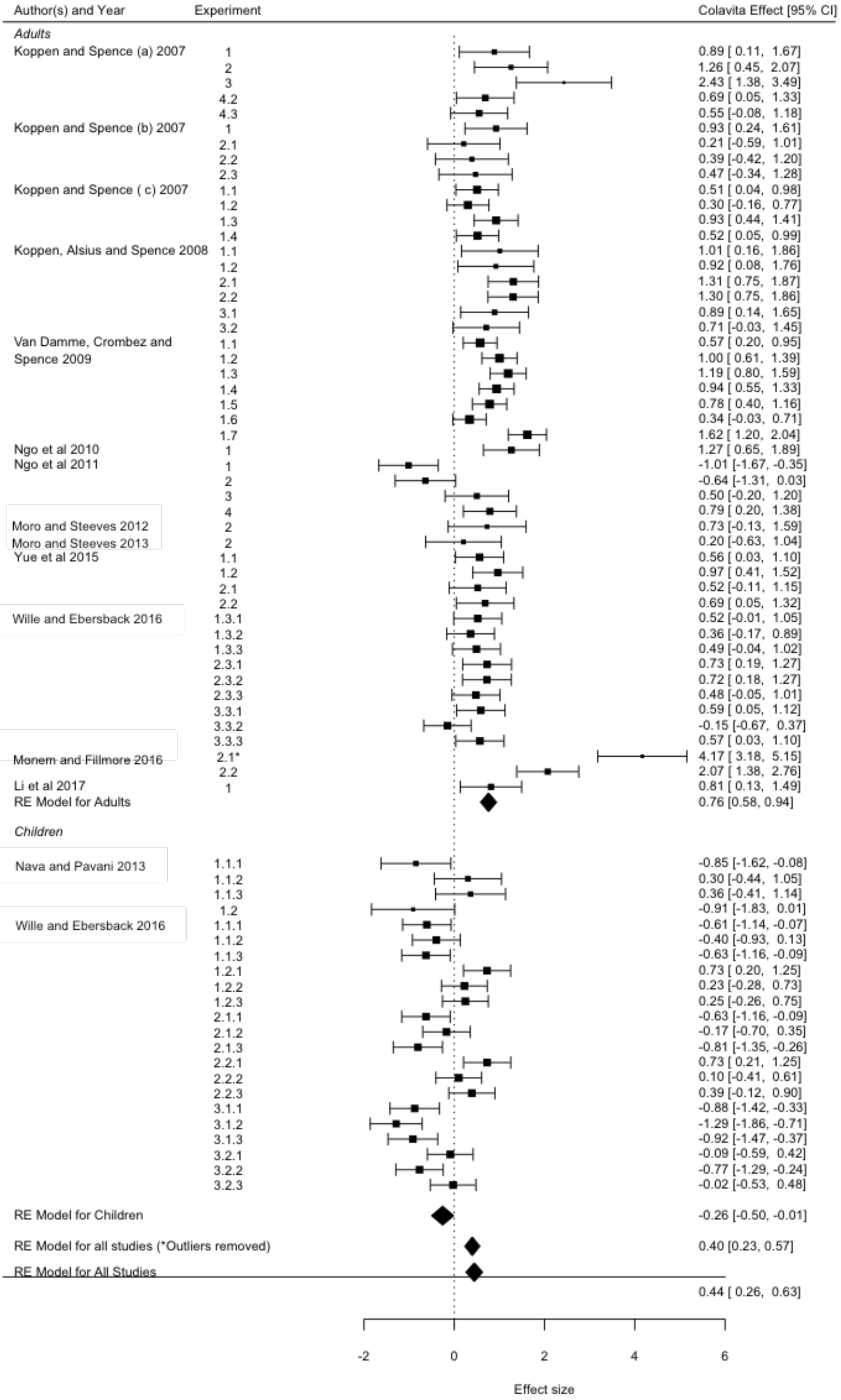


Figure 37. Effect sizes and 95% confidence intervals of studies reporting “visual only” responses on bimodal trials (the Colavita effect) and “auditory only” responses on bimodal trials. Symbol size reflects sample size. Weighted effect sizes are shown for all studies, all studies excluding outliers (asterisked experiments) and studies examining children and adults separately. Positive effect sizes indicate more “visual only” responses on bimodal trials. Negative effect sizes indicate more “auditory only” responses on bimodal trials.

Publication bias

To evaluate the presence of publication bias, data from studies included in model 1 (analysing the Colavita effect) were plotted as a funnel plot (Figure 38). The amount of scatter around the true effect should decrease with decreased sampling variance/increased sample size, thus producing a classic “funnel” shape (Macaskill, Walter, & Irwig, 2001). Publication bias is associated with funnel plot asymmetry (Egger, Davey Smith, Schneider, & Minder, 1997), whereby studies with large sampling variance/smaller sample size cluster to the left or right of the true effect. To quantify asymmetry a meta-analytic mixed effects regression analysis was performed, holding sample size as a predictor variable. This test indicated no significant asymmetry ($z = 1.04$, $p = .3$), suggesting the reported findings were not influenced by publication bias.

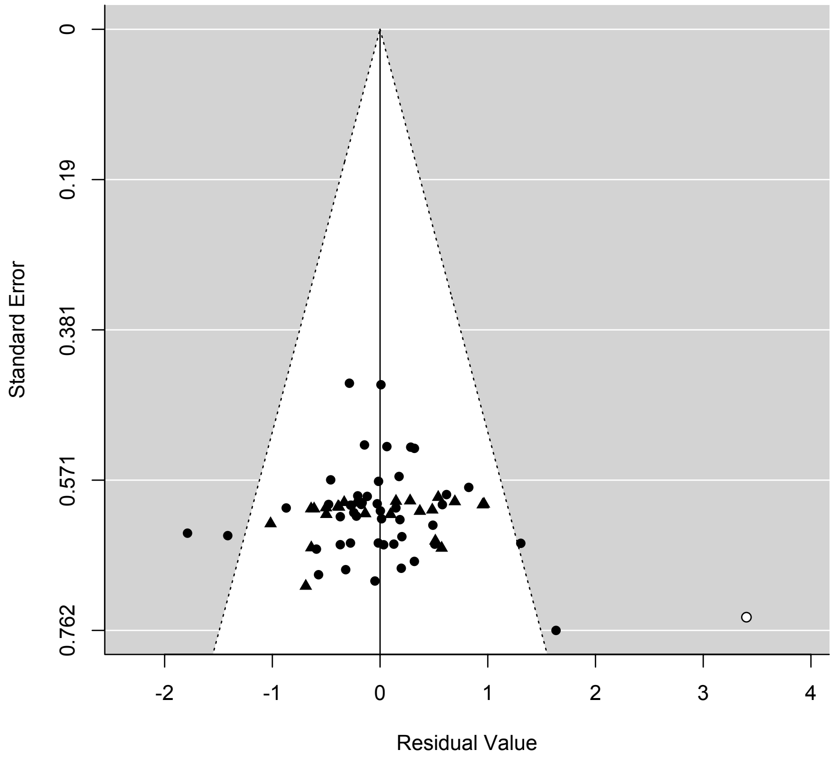


Figure 38. Funnel plot signifying the symmetrical distribution of effect size residuals (relative to the effect size of all studies) against standard error for studies reporting the Colavita effect. This symmetrical distribution suggests the amount of scatter around the true effect decreases with reduced standard error/increased sample size, suggesting no publication bias. Circles = adult studies, triangles = child studies, white circle = outlier/influential case.

Asymmetrical Facilitation: Response Time analyses

Studies reporting response times to auditory and visual stimuli under unimodal and bimodal conditions were used to investigate whether the Colavita effect occurs due to asymmetrical facilitation and inhibition (Sinnett et al., 2008). The first analysis compared response times to visual stimuli presented with an auditory stimulus (i.e. bimodal) to response times to unimodal visual targets. This asks if auditory stimuli facilitate response times to visual targets. The second analysis compared response times to auditory stimuli presented with a visual stimulus (i.e. bimodal) to response times to unimodal auditory targets. This asks if visual stimuli impede response times to auditory targets. Across both sets of analyses positive effect size values would indicate response times were faster to the target under bimodal conditions. Conversely, negative effect sizes would indicate response times were faster to the target in unimodal conditions. As in the analysis of Colavita errors, the effect of moderators was also explored in both sets of analyses to investigate if response time effects were modulated by; ratio, response keys (1 verses 2 as response time data were not available for any study using three keys), stimulus category, congruency, attentional manipulation, age group and whether stimuli were matched in intensity. This latter factor could only be included for the effect of audition on response times to visual targets, as all studies comparing unimodal and bimodal visual response times matched stimulus intensity.

Comparing response times to visual stimuli presented unimodally and bimodally

The combined effect size resulting from comparing response times to visual stimuli under unimodal vs. bimodal conditions was -0.26 (*SE* = 0.17) and non-significant (Figure 39). Two experiments (Egeth & Sager, 1977, experiment 4.2; Koppen & Spence 2007b, experiment 2.1) were identified as influential outliers. Removal of these studies resulted in an effect size of -0.43 (*SE* = 0.13), which significantly differed from 0 ($p < .001$). Contrary to Sinnett and colleagues' (2008) predictions of asymmetrical facilitation, response times were *slower* for visual stimuli accompanied by auditory stimuli compared to when they were presented alone.

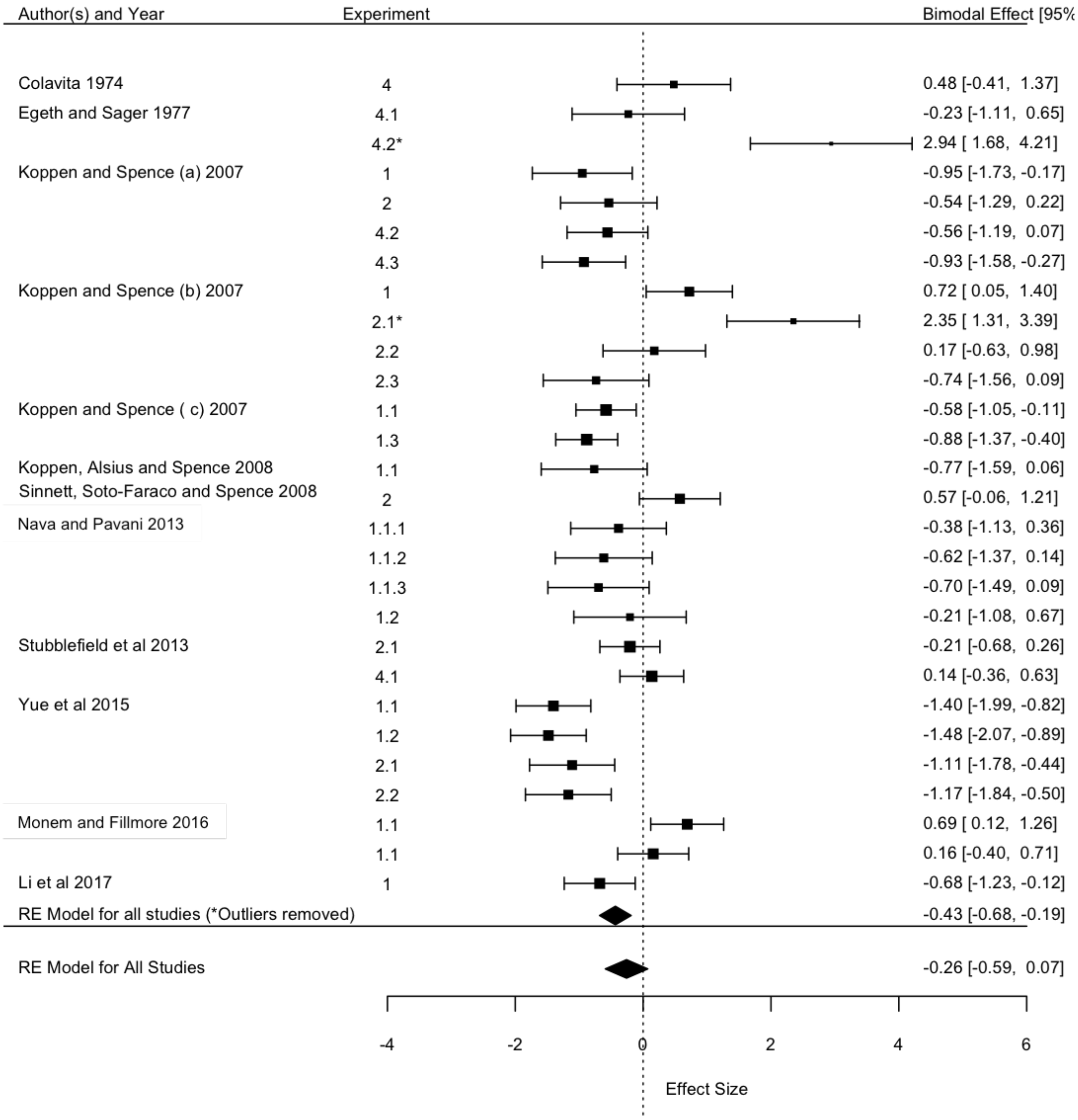


Figure 39. Effect sizes and 95% confidence intervals for studies/experiments reporting response times (RT) for visual targets under unimodal and bimodal conditions. Symbol size reflects sample size. Positive effect sizes indicate RT was faster under bimodal versus unimodal conditions. Negative effect sizes indicate RT was faster under unimodal versus bimodal conditions.

To explore the effects of moderator variables a mixed meta-regression model was conducted in which the intercept (reference) was set to reflect the effect size of studies using the most frequently used experimental parameters, as above. This model indicated that 96.74% of the residual heterogeneity ($tau^2 = 0.01$, *SE* = 0.04) was accounted for by the inclusion of moderator variables (QM(12)=75.25, $p < .001$). The effect size estimate of the intercept was large (-0.95, *SE* = 0.12), and decreased in studies using ratios in which bimodal stimuli were more frequent (20:20:60, 25:25:50 and 33:33:33; yielding estimated changes of 1.67 (*SE* = 0.38, *p* < .001), 1.13 (*SE* = 0.44, *p* < .01) and 0.39 (*SE* = 0.18, *p* = .0277) respectively). Thus when bimodal trials were infrequent (20%) response times were slower to visual targets under bimodal conditions. However, when bimodal targets were more frequent (33%, 50%, or 60%), this effect was decreased. The effect size was also decreased by 1.53 (*SE* = 0.36, *p* < .001) in studies using complex stimuli and increased by 1.34 (*SE* = 0.55, *p* =.0148) in experiments using congruent stimuli. In line with this, post-hoc ANOVAS showed a significant overall effect of ratio, stimulus category, and congruency upon the intercept whilst other factors did not yield a significant overall effect (Table 12). A test of residual heterogeneity was non-significant (QE(11) = 11.95, *p* = .37), suggesting there was no further heterogeneity not accounted for within the model.

Given the significant effect of ratio (i.e. the balance of audio-visual, unimodal visual and unimodal auditory trials) and stimulus category (i.e. simple stimuli such as flashes and tones vs. complex stimuli such as images and naturalistic sounds) found above, two further models were fitted to directly compare the effect size of multisensory facilitation/interference of studies using different ratios and stimulus categories regardless of other factors. A further model was not fitted to explore the effect of congruency as this had only been manipulated in one study.

Factor	**df**	**QM**	***p***
Ratio	5	28.460	<.0001
Response keys	1	01.931	0.1647
Stimulus category	1	18.307	<.0001
Congruency	2	07.049	0.0295
Attentional manipulation	2	01.176	0.5556
Age group	1	02.683	0.1014

Table 12. Statistics resulting from additional analyses of variance (ANOVAs) exploring the effect of each factor upon the intercept of the mixed model (i.e. the overall effect size for the effect of auditory stimuli on visual target detection). Three factors, ratio, stimulus category and congruency, significantly influenced the effect of auditory stimuli upon visual target detection. df = degrees of freedom, QM= omnibus test statistic.

The model for ratio indicated that only studies using the ratios 40:40:20 yielded effect sizes that significantly differed from 0 (*p* < .001). This suggested that when bimodal trials were infrequent (20%) response times to visual stimuli were slower under bimodal conditions. However when bimodal trials were more frequent (33%, 50% or 60%) response times were not significantly affected by auditory stimuli.

The model addressing stimulus category (simple vs. complex) revealed that only experiments using simple stimuli yielded an effect size that significantly differed from 0 ($p < .001$). This suggested that participants were slower to respond to visual stimuli paired with auditory stimuli but only when simple stimuli were used.

Overall, these findings were not consistent with the hypothesis that response times to visual targets would be faster under bimodal vs. unimodal conditions. Rather, these findings suggested response times were slower to visual targets paired with auditory stimuli particularly when the frequency of bimodal targets was low and when simple stimuli were used.

Comparing response times to auditory stimuli presented unimodally and bimodally

The combined effect size for unimodal auditory vs. bimodal auditory response times was medium (-0.57, *SE* = 0.08), and significant ($p < .001$). No experiments were identified as outliers.

A mixed meta-regression model was fitted for this effect in which studies using the parameters outlined as standard (see above) were used as the intercept. This model revealed no significant remaining heterogeneity ($tau^2 = 0$, *SE* = 0.04, QE(14) = 8.35, $p = .8701$) and a significant effect of moderators (QM(9)=20.5, $p = .0248$). However, post-hoc ANOVAs did not indicate any of the moderator variables to significantly influence the intercept (Table 13). From this it was concluded that participants were slower for auditory targets paired with visual stimuli compared with unimodal targets, as can be seen in Figure 40, and this was not modulated by experimental parameters.

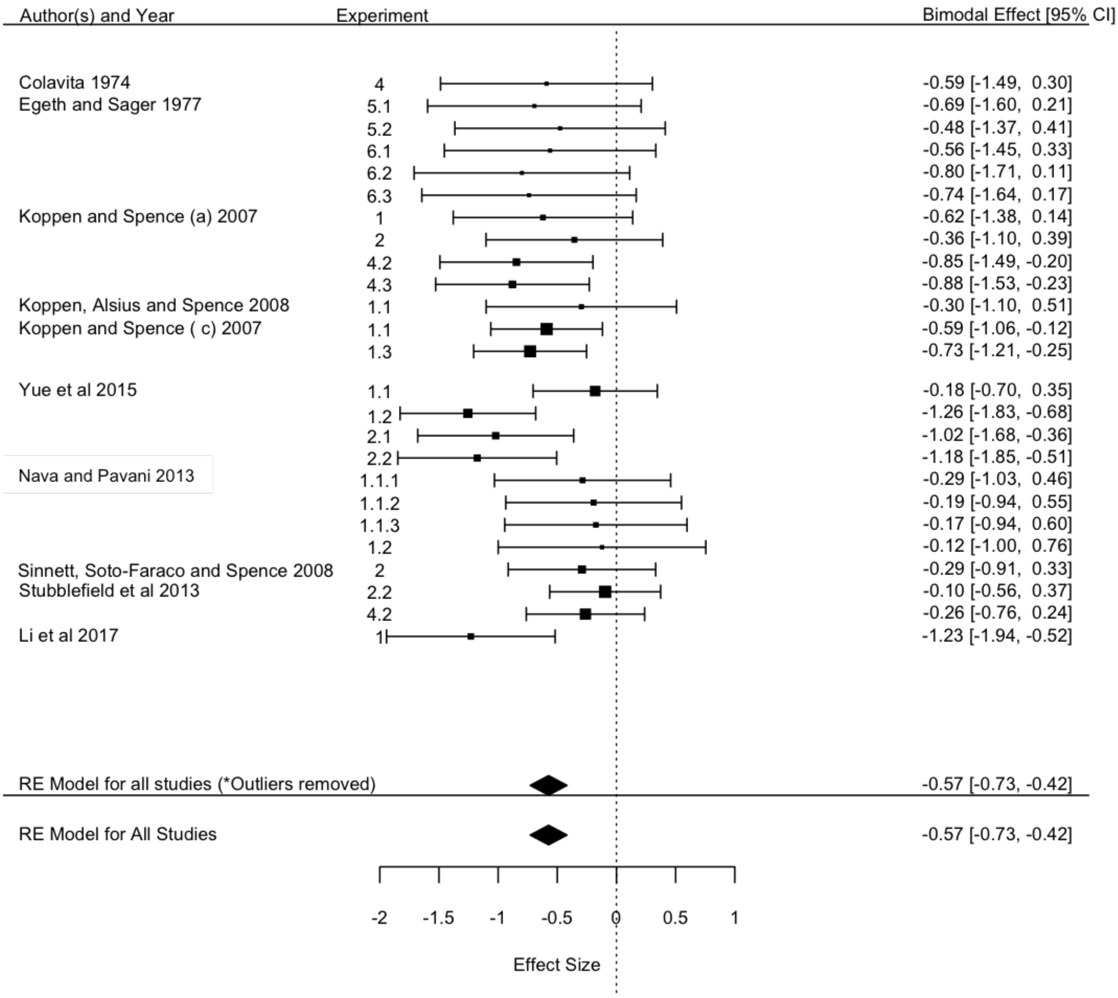


Figure 40. Effect sizes and 95% confidence intervals for studies/experiments reporting response times (RT) for auditory targets under unimodal and bimodal conditions. Symbol size reflects sample size. Positive effect sizes indicate RT was faster under bimodal versus unimodal conditions. Negative effect sizes indicate RT was faster under unimodal versus bimodal conditions

Factor	**df**	**QM**	***p***
Ratio	5	1.020	0.6005
Response keys	1	2.285	0.1307
Stimulus category	1	1.639	0.2004
Congruency	2	1.132	0.2874
Attentional manipulation	2	3.264	0.1956
Age group	1	3.103	0.0781
Matched	1	2.577	0.1084

Table 13. Statistics resulting from additional analyses of variance (ANOVAs) exploring the effect of each factor upon the intercept of the mixed model (i.e. the overall effect size for the effect of visual stimuli on auditory target detection). df =degrees of freedom, QM = omnibus test statistic.

Is the bimodal slowing effect between vision and audition symmetrical?

Contrary to the prediction based on the hypothesis of Sinnett et al. (2008) the current analysis showed vision slowed response times to auditory targets and vice versa. Robinson, Chandra and Sinnett (2016) noted that this might occur when multiple response keys are used, and conceptualised sensory dominance via the relative extent to which one sense slows another. They found that, when a single response key was used, visual stimuli slowed auditory response times more than auditory stimuli slowed visual response times. Moreover, when separate response options were available, auditory stimuli *also* slowed response times to visual stimuli. The authors interpret the extent to which one sense slowed the other as a measure of sensory dominance. To test whether vision slowed response times to auditory targets more than audition slowed response times to visual targets, a final model was fitted to directly compare the effect sizes yielded in the former two comparisons. No significant difference was found, suggesting visual and auditory stimuli slowed response times to the opposing modality to a similar extent.

Discussion

The current study quantitatively demonstrates that Colavita errors, whereby participants report only the visual element of an audio-visual target, are a robust experimental phenomenon. Mixed-effects analyses also corroborated the suggestion that Colavita errors are relatively insensitive to response demands, attentional manipulation, stimulus ratio, stimulus complexity, and congruency. However, residual heterogeneity did remain within the model, therefore, it should be noted that other factors not accounted for in this model likely influence the effect size of the Colavita effect.

Furthermore, the Colavita effect was moderated by age, in that it is smaller, perhaps even reversed, in childhood. Although the current analysis includes only 2 childhood studies, these studies include data from a relatively large sample of 187 children aged between 6 and 12 years (Nava & Pavani, 2013, n = 51; Wille & Ebersbach, 2016, n = 136). If the tentative finding of a reversed Colavita effect in children appears in further studies this would be in line with evidence suggesting an auditory preference in childhood (Napolitano & Sloutsky, 2004; Robinson et al., 2016b; Robinson & Sloutsky, 2004, 2010; Sloutsky & Napolitano, 2003) and difficulty ignoring auditory distractions in childhood (Hanauer & Brooks, 2003). It should be noted that in this analysis, I implemented a binary categorisation of age group (“adult” or “child”). Wille and Ebersbach (2016), however, reported a transition towards visual dominance around 9 years of age. As such, it must be considered that the size of the Colavita effect reported in children here likely differs between younger and older children. These previous findings together with the current data make an interesting case for the fluctuation of sensory dominance across the lifespan and highlight this as a field warranting further investigation.

The analysis of response times suggested that responses were slower for both visual and auditory stimuli when participants responded under bimodal rather than unimodal conditions and the effects of vision on audition and vice versa were not significantly different. The current study therefore does not suggest an asymmetrical relationship between vision and audition as proposed by Sinnett et al. (2008). They hypothesised a co-occurrence of multisensory facilitation and inhibition whereby auditory stimuli facilitate visual detection whilst visual stimuli inhibit auditory detection. This asymmetry was proposed to lead to the Colavita effect, since a visual response would be more likely to occur first on bimodal trials. An alternative, symmetrical, prediction is that response times are always faster under bimodal conditions. This would be expected based upon the known principles of multisensory integration, whereby neural responses elicited from bimodal targets are greater than unimodal targets (i.e. additive; see Stanford, Wallace, Vaughan, & Jiang, 2004). However, these findings suggest that response times were in fact *slower* under bimodal conditions. This finding is contrary to both asymmetric and symmetric models of multisensory facilitation.

One likely explanation for slowing on bimodal trials is that most studies used at least two response keys, whereas previous literature finding multisensory facilitation (faster responses on bimodal trials) has used one response key (Forster et al., 2002; Gondan et al., 2005; Sinnett et al., 2008). Moreover, most Colavita studies traditionally present response time data only for correct trials. If multisensory facilitation does contribute to the Colavita errors, the beneficial effects of audition upon visual response times might be more evident within incorrect trials. For example, in order to respond to a bimodal target correctly (i.e. with both buttons) it may be that participants must first suppress the automatic tendency to respond towards only the visual target and then make the correct, bimodal, response. Thus, response times on correct trials would be slower due to the need to suppress automatic responses. This explanation is at present tentative.

This analysis indicated that slowing of responses to visual targets by auditory stimuli was decreased in studies using fewer bimodal trials. This contradicts previous findings by Sinnett et al. (2007, Experiment 3), who found that the frequency of bimodal targets did not influence reaction times. Thus, although the influence of stimulus ratio on response times was not revealed at the single study level, combining across several studies did yield this effect. It is possible that a more equal distribution of unimodal and bimodal target types (33% visual, 33% auditory and 33% audio-visual) produces equivalent response times across targets by limiting effects such as novelty.

Only one adult study included in the analysis of Colavita errors yielded a clear reverse Colavita effect (Ngo et al., 2011). This study utilised a repetition detection variant of the Colavita paradigm. Participants were required to detect (n-1) repetitions in auditory, visual and audio-visual information. The temporal demands of this task, however, were predicted to introduce auditory dominance (Welch & Warren, 1980). Ngo and colleagues also predicted that this would be exaggerated by the longer lasting nature of echoic vs. iconic short-term memory. The reversal of the Colavita effect in this study is therefore attributed to arise from a greater visual masking of targets by intervening irrelevant items under visual vs. auditory conditions. In line with this, if the intervening item was semantically meaningless (a pattern mask/ burst of white noise), neither auditory nor visual dominance was observed.

Finally, it is notable that Colavita errors are not the only method by which sensory dominance has been operationalized, and other methods have not consistently inferred visual dominance in adults. As outlined in the final analysis of response times, Robinson et al. (2016) propose that, in adults, when a single response key is used, auditory stimuli slow response times to visual targets more than vice versa (suggesting auditory dominance). Conversely, when multiple separate responses are required to visual, auditory, and bimodal targets (as in many of the included Colavita studies) visual dominance is seen. Interestingly, Barnhart et al. (2018) recently demonstrated that although auditory dominance effects (operationalized via response times) occurred in children and young adults, the reverse occurred in older adults. This indicates a shift in sensory dominance across the lifespan and enhanced visual dominance in later life. In the current analysis, the extent to which vision slowed audition and vice versa did not differ, and this did not differ between adults and children. Nevertheless, this may have also been influenced by response times being based on correct trials (if slower responses were needed to make a correct response) and the limited number of child experiments included for analysis.

Conclusions

The current study provides an updated synthesis of literature surrounding the Colavita effect. The Colavita effect appears to be a robust phenomenon with medium effect size in adults, although not in children. The Colavita effect also appears insensitive to many experimental manipulations although it may be reversed under some designs (Ngo et al., 2011). This study highlights a need to examine the Colavita effect across the lifespan and suggests that visual dominance over audition may be weaker, or even reversed, in childhood.

Following this, and in answer to our original postulation, if you are an adult reading this thesis you may be more distracted by an email pop-up vs. your phone ringing. Furthermore, if your phone rings at the same time you see an email pop-up you may not answer (or hear) the phone at all. For this, you can blame sensory dominance.

Chapter 8: The threshold for the McGurk effect in audio-visual noise decreases with development

Chapter 7 suggested vision increasingly influences audio-visual perception across development. In this Chapter, I assessed whether this is evident in multisensory illusions in which vision can alter auditory perception (i.e. the McGurk effect). This study assessed the effects of manipulating the clarity of the heard and seen signal upon the McGurk effect in children aged 3 - 6 (n=29), 7 - 9 (n=32) and 10 – 12 (n=29) years, and adults aged 20 - 35 years (n=32). Auditory noise increased, and visual blur decreased, the likelihood of vision changing auditory perception. Based upon a proposed developmental shift from auditory to visual dominance I predicted that younger children would be less susceptible to McGurk responses, and that adults would continue to be influenced by vision in higher levels of visual noise and with less auditory noise. Susceptibility to the McGurk effect was higher in adults compared with 3-6-year-olds and 7-9-year-olds but not 10-12-year-olds. Younger children required more auditory noise, and less visual noise, than adults to eliminate McGurk responses (i.e. adults and older children were more easily influenced by vision). Reduced susceptibility in childhood supports the theory that sensory dominance shifts across development and reaches adult-like levels by 10 years of age.

Author contributions statement

The experimental work conducted in this Chapter was a collaborative project lead by myself. I developed the study concept in collaboration with Jemaine Stacey (Nottingham Trent University) and under the supervision of our supervisors, Lucy Cragg, Harriet Allen and Paula Stacey. Stimuli were designed and developed by Jemaine Stacey. I collected and analysed the data and data were interpreted by myself and Jemaine Stacey. I then prepared the manuscript, from which this Chapter is adapted, for publication.

Introduction

The ability to combine auditory and visual information (audio-visual integration) in noisy environments is essential in every day life, such as when holding a conversation. Individuals may differ in the extent to which one sense (or modality) alters or is *dominant* over another. The *McGurk effect* exemplifies this, as hearing a voice say “Ba” and seeing a face say “Ga” often results in perception of an alternative syllable such as “Da” or “Tha” (McGurk & MacDonald, 1976). Sensory dominance, or different weighting between sensory modalities results in different perceptual effects. Visual dominance results in a response to the seen mouth movement, auditory dominance a response to the sound and equal weighting results in a fusion of the two. Alternatively, fusion responses have also been interpreted as visual dominance (Lüttke, Ekman, Van Gerven, & De Lange, 2016), where the seen mouth movement alters the sound reported by participants.

Modality dominance might be flexible depending on context (Robinson, Chandra, & Sinnett, 2016a; Welch & Warren, 1986) and age (Barnhart et al., 2018; Diaconescu et al., 2013; Nava & Pavani, 2013). This was demonstrated in Chapter 7, in which the analysis of studies using the Colavita effect indicated age to be the primary moderator of sensory dominance measured using the Colavita effect.

The perceptual consequences of auditory dominance in childhood are evident in multisensory illusions. Children are more susceptible to illusions in which auditory information modulates vision. In the flash-beep illusion, participants perceive a single flash as two flashes when presented with two concurrent beeps. Innes-Brown et al. (2011) found 8-17-year-olds were more susceptible to the flash-beep illusion compared with adults (Innes-Brown et al., 2011; however see Parker & Robinson, 2018). Conversely, children appear less susceptible to the McGurk effect, maintaining veridical perception of sound despite incongruent visual information (Narinesingh et al., 2015; Tremblay et al., 2007). Tremblay et al. (2007) found that 5-9-year-olds made correct auditory “Ba” responses on ~60% of incongruent McGurk trials. This dropped to ~20-30% in 10-19-year-olds, suggesting older children are more susceptible to the McGurk effect.

One explanation of developmental differences is that children manifest delayed development of multisensory integration processes (Ernst, 2008) and thus are less susceptible to multisensory illusions. Nevertheless, children experience other multisensory illusions, such as the flash-beep illusion (Innes-Brown et al., 2011), and susceptibility to the McGurk effect is modulated by sensory weighting in childhood. Children who experienced early visual and hearing impairments are respectively less and more susceptible to the McGurk effect (Narinesingh et al., 2015; Schorr et al., 2005). Furthermore, lip reading ability in childhood is correlated with the size of visual contribution in speech perception (Massaro, Thompson, Barron, & Laren, 1986). These findings provide compelling evidence for a theory of an experience-based shift in sensory dominance, from audition to vision, that may be modulated by the learnt reliability of visual and auditory input.

In line with a role of sensory reliability, the presence of noise in one or both modalities influences sensory dominance and thus, multisensory integration. Everyday environments are inherently noisy, and this influences which sense drives audio-visual integration. During conversation, the listener may utilise both vision and audition to understand speech. However, in noisy environments, visual information may be more informative. If visual information becomes unclear through factors such as impaired sight or poor viewing conditions, then audition may be particularly salient. In line with this, current theory suggests that the brain weights sensory inputs according to their relative reliability to derive the most accurate percept possible (Brooks et al., 2018; Ernst & Bülthoff, 2004; Fetsch, Deangelis, & Angelaki, 2013; Witten & Knudsen, 2005). Following this, adults are more susceptible to the McGurk effect in auditory noise (Hazan & Li, 2008; Sekiyama & Burnham, 2008), and less susceptible in visual noise (Fixmer & Hawkins, 1998; Hazan & Li, 2008) as audition becomes respectively less and more reliable. This also explains why children with early visual and hearing impairments are respectively less and more susceptible to the McGurk effect (Narinesingh et al., 2015; Schorr et al., 2005). Thus, sensory dominance can be modulated within an individual by manipulating the reliability of sensory information.

Recent findings show that when *both* vision and audition are degraded, the McGurk illusion persists. Stacey, Howard, Mitra, and Stacey, (2017) degraded visual and auditory information in McGurk stimuli through introducing blur and white-noise respectively. In line with previous findings, the McGurk effect increased in high levels of auditory noise and decreased when visual information was degraded. Interestingly, McGurk perception remained robust even when both visual and auditory information were degraded; participants still perceived the effect on 66% of trials.

Noisy environments have an everyday impact on audio-visual integration and perception at every stage of life. Yet the effect of *combined* visual and auditory noise upon the McGurk effect in children remains unexplored. To my knowledge, no studies have explored the influence of visual noise on the McGurk effect in children, and only one study has examined auditory noise. Sekiyama and Burnham, (2008) tested the McGurk effect in 6-, 8- and11-year-olds and adults using four levels of auditory noise. Children were less susceptible to the effect, nevertheless auditory noise increased the effect in both children and adults.

Multiple studies have examined the effect of noise on the McGurk effect in adults (Fixmer & Hawkins, 1998; Hazan & Li, 2008; Sekiyama & Burnham, 2008; Sekiyama & Tohkura, 1991; Stacey et al., 2017). However, to my knowledge, none have exploited the effect of manipulating stimulus clarity to derive a threshold for the McGurk effect. A psychophysical approach to measuring sensory weighting in the McGurk effect is informed by computational models of McGurk perception. The Noisy Encoding of Disparity (NED) Model proposed by Magnotti and colleagues (Magnotti & Beauchamp, 2015) proposes that individual differences in McGurk perception may be accounted for by differences in sensory disparity, sensory noise and individual “disparity threshold”, a point at which noise in one modality becomes high enough to prevent fused percepts. An implication of this model is that manipulating sensory noise in one or the other modality may provide an indication of individual differences in thresholds for the effect. This has a benefit over previous approaches, which have used group means, as it provides an indication of how weighting between vision and audition may differ between individuals, change with development, and produce differing thresholds for audio-visual illusions.

This study explored the effect of auditory and visual signal quality on McGurk responses across development to derive thresholds for McGurk responses. The threshold was defined as the noise level inducing incorrect, non-auditory “Ba” responses 50% of the time – reflecting the point at which vision prevents correct auditory perception.

Specifically, hypotheses were:

- Adults would show more McGurk responses than children (regardless of noise level).
- The frequency of McGurk responses would increase across development.
- Although auditory and visual noise were expected to increase and decrease the McGurk effect respectively in adults and children, it was hypothesised that adults would show a lower threshold for the McGurk effect compared with children (i.e. require more visual noise to abolish the effect and less auditory noise to induce the effect or, in other words, would require less auditory noise to prevent correct auditory, “Ba”, perception and would show visually influenced, non-auditory, responses even with higher visual noise).
- The threshold for the McGurk effect would progressively decrease across childhood.

Method

Participants

To accurately judge the sample size required to detect an effect of noise on the McGurk effect (required for calculating thresholds) an a priori power analysis was conducted in G*power v3.1 to detect a Cohen's d of 0.8 in a 2 (sensory condition) x 5 (noise level) ANOVA (see supplementary material). This effect size was used based on the large effect sizes reported in the literature for the effect of noise on McGurk responses (Fixmer & Hawkins, 1998; Hazan & Li, 2008). This analysis governed the size of the adult sample (n = 32). The child sample size was based on opportunity (data were gathered at a large public engagement event and all children had the opportunity to participate). Following data collection, the sample size and age distribution of the child sample permitted a separation of children into three age groups, 3-6-year-olds (n=29), 7-9-year-olds (n=32) and 10-12-year-olds (n=29), enabling a more thorough comparison between different stages of childhood and adulthood.

Thirty-two young adults (Mean age = 26.66 years; range = 20-35 years; 19 female; 31 right handed) were recruited. Participants were staff and students of the University of Nottingham. They reported having normal or corrected to normal vision and hearing and were fluent English speakers (28 reported English first language, 1 Portuguese, 1 Icelandic, 1 Chinese and 1 Catalan).

Ninety-six children (Mean age = 8.1 years; range = 3.92-12 years; 47 female) were recruited via Summer Scientist Week (www.summerscientist.org), a public engagement event at the University of Nottingham. Following data collection, children were grouped into three evenly distributed groups for analysis; 3-6-year-olds, 7-9-year-olds and 10-12-year-olds. All participants were fluent English speakers (84 reported that English was the primary language used at home, 4 reported that English was used at home alongside a second language and 8 reported that another language was used at home (1 Russian, 1 Japanese, 1 Chinese, 1 Portuguese, 2 Telugu, 2 Tamil). Children were rewarded for their time with tokens to take part in other activities. Four 3-6-year-olds were excluded because they did not complete the task. Two 7-9-year-olds were also excluded as parents reported sensory processing difficulties (a perforated ear-drum and sensory processing disorder). Thus, a final sample of 90 children was available for analysis; 29 3-6-year-olds (14 female, Mean age 5.6 years; range 3.92-6.92 years), 32 7-9-year-olds (18 female, Mean age 8.3 years; range 7-9.75 years), 29 10-12-year-olds (14 female, Mean age 10.97 years; range 10-12.08 years) and 32 adults.

Equipment

Visual stimuli were presented via a Macbook Air on a Lenovo LT2423 24" LED Backlit LCD monitor (resolution 1920x1080 @ 60Hz) presented at a viewing distance of ~57cm. Auditory stimuli were presented via Senheiser eH150 headphones. A Targus numerical response pad was used to gather responses.

Stimuli

Stimuli were created by splicing together auditory and visual components using Adobe Premiere Pro. Stimuli consisted of videos of a single female speaker vocalising one of three syllables; “Ba”, “Ga” or “Da”. On congruent trials (75% of trials) congruent auditory stimuli were presented (25% “Ba”, 25% “Ga”, 25% “Da”). On incongruent trials (25% of trials), a visual “Ga” and an auditory “Ba” were presented. The proportion of incongruent trials used were comparable, if slightly higher, than those used in previous studies (Narinesingh et al., 2015; Schorr et al., 2005; Tremblay et al., 2007). The same female speaker was used for all test trials and two different female speakers were used for the practice trials. Videos displayed the head and shoulders of the speaker against a plain white background (size 40 x 21cm, 2 seconds duration, audio = 41000 Hz, 16 bit). Five levels of visual noise were created via Premiere Pro using the Gaussian blur function (0%, 30%, 40%, 50% and 60% blur). For the purposes of this Chapter both auditory noise and visual blur are referred to as noise, although blurring is a reduction in quality of the signal rather than strictly added ‘noise’. Syllables were presented either without noise or alongside white noise at 4 Signal-to-Noise-Ratios (SNRs; -2dB, -8dB, -14dB and -20dB). All stimuli were presented at the same sound level (50dB) determined using an artificial ear (Brüel & Kjær Type 4153). This intensity was clearly audible for all participants as accuracy for syllables in the absence of noise was persistently high (>80% see supplementary material). The five levels of auditory and visual noise were combined to produce 25 levels of stimulus quality per syllable (see Figure 1a). There were therefore 100 trials, 25 stimuli per condition (one per each possible noise level). On 10% of trials (see below) a pink cartoon monster (4 x 3.5cm) appeared covering the mouth, alongside a laughter sound effect. One catch trial was presented randomly within each 10 trial block.

Procedure

Adult participants completed the task in a quiet testing lab at the University. Child participants completed the task in a quiet room at the University alongside other studies taking place.

Within each trial, a video was presented followed by an on-screen message asking “What did you hear?” (Figure 41b) after which participants could respond using three counterbalanced response keys (“Ba”, “Ga” or “Da”/”Tha” – Figure 41). “Da” and “Tha” were mapped to the same response option in line with previous literature (Mallick, Magnotti, & Beuchamp, 2015; Stacey et al., 2017). If children could not read the labels they were asked to vocalise their responses and the experimenter would press the button. Once a response had been made the next trial began immediately. Previous research with children has used up to 6 response options (Narinesingh et al., 2015) or allowed an open-ended vocalised response (Schorr et al., 2005; Tremblay et al., 2007), thus, three response keys were judged to be appropriate.

Participants first completed five, randomly selected, congruent practice trials in which the spoken syllable was presented in the absence of any noise. Practice trials were followed by 10 blocks of 10 trials in a randomised order. Following each block participants clicked on one of ten treasure chests on the screen, revealing a clue to where a reward token was hidden.

Participants were instructed to focus on the mouth at all times. To ensure attention was maintained upon the mouth, a cartoon monster appeared in the mouth region once per block (Figure 41c). When the participant saw the monster they pressed a red button, on the same response pad. The trial would not move on until the participant had pressed the red button. All participants included in the analyses successfully completed all 100 trials.

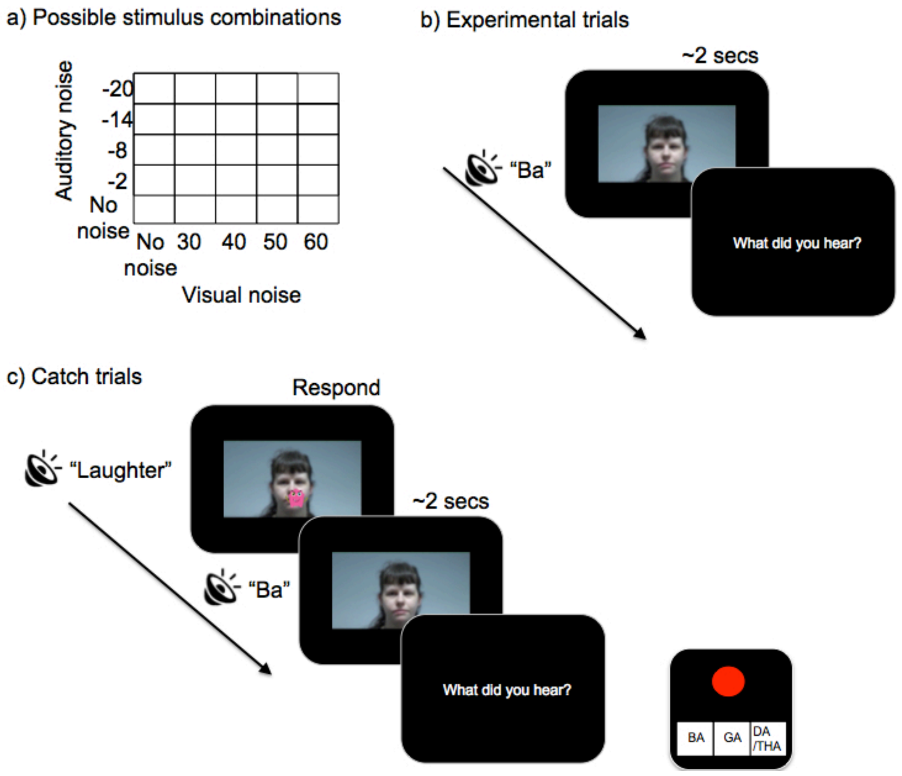


Figure 41. a) Possible stimulus combinations. There were 25 congruent “Ba”, 25 congruent “Da”, 25 congruent “Ga” and 25 incongruent auditory “Ba” visual “Ga” trials. b) Experimental trials: Participants watched a 2 second video accompanied with either a congruent or incongruent sound. When asked “what did you hear?” they responded using one of three buttons. c) Catch trials: The video froze at the beginning and the monster cartoon/laughter sound effect was presented. Participants pressed the red button as fast as possible and then the trial continued as normal.

Analysis and Results

First, I provide a summary of responses to congruent trials. Then, I focus on responses to incongruent trials. Following this, I present a threshold analysis to identify the 50% threshold for the McGurk effect (i.e. the point at which McGurk responses were made 50% of the time) in auditory noise (collapsed across visual noise conditions), visual noise (collapsed across auditory noise) and combined audio-visual noise. In line with the preceeding experimental chapters (with the exception of Chapter 3), I refer to *p* values <=.05 as significant in this analysis.

Responses to congruent trials

A full analysis of responses to congruent trials is available in appendix, page 326. As responses to congruent stimuli were not the main focus of this study, I here provide a brief overview of these results.

Accuracy on congruent trials was higher for congruent “Ba” and “Da” stimuli compared with “Ga” stimuli (Table 14). Across groups, participants frequently made “Da/Tha” errors in response to congruent “Ga” stimuli. Increasing visual and auditory noise also lowered accuracy for congruent trials. Interestingly, the effect of auditory, but not visual noise, interacted with age group. Younger children were less accurate than older children and adults only when there was no auditory noise or the highest levels of auditory noise.

Stimulus	Visual Ba /Auditory Ba			Visual Ga/Auditory Ga			Visual Da/Auditory Da			Visual Ga /Auditory Ba		
										Aud	Vis	Fus
Response	“Ba”	“Ga”	“Da”/”Tha”	“Ba”	“Ga”	“Da”/”Tha”	“Ba”	“Ga”	“Da”/”Tha”	“Ba”	“Ga”	“Da”/”Tha”
3 - 7	75.45	7.45	17.1	22.21	34.9	42.9	14.07	10.07	75.86	52.96	15.59	31.49
	(2.59)	(1.22)	(2.11)	(1.78)	(3.42)	(3.49)	(1.4)	(1.3)	(1.96)	(2.38)	(2.19)	(3.04)
7 - 10	84.5	5.50	10.0	17.00	42.75	40.25	9.125	8.88	82.00	48.75	20.5	30.75
	(2.47)	(1.16)	(2.01)	(1.7)	(3.26)	(3.32)	(1.34)	(1.24)	(1.87)	(2.27)	(2.09)	(2.9)
10 - 12	84.69	4.28	11.03	10.62	39.45	49.93	6.76	6.21	87.03	45.1	17.93	36.97
	(2.59)	(1.22)	(2.11)	(1.78)	(3.42)	(3.49)	(1.4)	(1.3)	(1.96)	(2.38)	(2.19)	(3.04)
Adults	89.00	2.63	7.75	5.25	42.25	52.13	4.13	4.25	91.63	37	19.88	42.88
	(2.47)	(1.16)	(2.01)	(1.7)	(3.26)	(3.32)	(1.34)	(1.24)	(1.87)	(2.27)	(2.09)	(2.9)

Table 14. Mean percentage of each response made to each stimulus (collapsed across all noise conditions) in each age groups. Standard error of the mean shown in brackets. Far right columns represent responses made on incongruent trials. Aud= Auditory, Vis = Visual, Fus = Fusion.

Development of the McGurk effect

Throughout analyses of incongruent trials four definitions are used to consider separate effects; “Visual”, “Auditory”, “Fusion” and “McGurk” responses, defined as follows:

- *Visual responses* - “Ga” response to incongruent McGurk stimuli; reflecting a response to the visually presented information.
- *Auditory responses* - “Ba” response to incongruent McGurk stimuli; reflecting a response to the auditory information.
- *Fusion* responses - “Da”/”Tha” response to incongruent McGurk stimuli - participants fuse auditory and visual information to report a syllable different from both the visual (“Ga”) and auditory (“Ba”; McGurk & McDonald, 1976).
- *McGurk responses* - both visual (“Ga”) and fusion (“Da”/”Tha”) responses; reflecting the point at which visual information influences or prevents veridical perception of auditory information.

Incongruent trials were first analysed by assessing mean visual, auditory and fusion responses, regardless of noise level, between age groups. Note that the proportions of these responses are not independent, since participants can make any of these responses to an incongruent trial.

Do adults make more McGurk responses than children?

A 4 (age group: 3-6-year-olds, 7-9-year-olds, 10-12-year-olds and adults) x 3 (response type: “Ba”, “Ga”, “Da/Tha”) ANOVA was used to compare responses made on incongruent trials (Table 14). This showed a main effect of response type (*F(*1.7, 201.29) = 81.861, *p* < .001, η^2 = .38). Participants made more “Ba” responses (*M* = 45.96%, *SE* = 1.16) compared with “Da”/”Tha” (*M* = 18.47% *SE* = 1.07) and “Ga” (*M* = 35.51% *SE* = 1.48), and more “Ga” than “Da”/”Tha” responses (*p* <.001 for all comparisons).

There was no main effect of age group (*F(*1, 3) = 1.93, *p* = .128, η^2 = .05), but an interaction between response type and age group (*F(*15.12, 201.29) = 4.49, *p* =.001, η^2 = .09). This occurred because the effect of age was significant for fusion ("Da"/"Tha") responses (*F*(3, 118) = 3.73, *p* =.013, η^2 = .09) and auditory ("Ba") responses (*F*(3, 118) = 8.61, *p* <.001, η^2 = .18) but not visual ("Ga") responses (*F*(3, 118) = 1.06, *p* =.368, η^2 = .03). As shown in Figure 42, adults made more fusion responses than 3-6-year-olds (*p* = .045) and 7-9-year-olds (*p* = .022). However adults did not differ from 10-12-year-olds (*p* = .972), and child groups did not significantly differ from one another (all *p* values >.849).

Correct auditory "Ba" responses were higher in 3-6-year-olds and 7-9-year-olds vs. adults (*p* < .001 and *p* = .002 respectively). However, 3-6-year-olds and 7-9-year-olds did not significantly differ from one another (*p* = 1). 10–12-year-olds, did not significantly differ from adults or other child groups (all *p* values > .09).

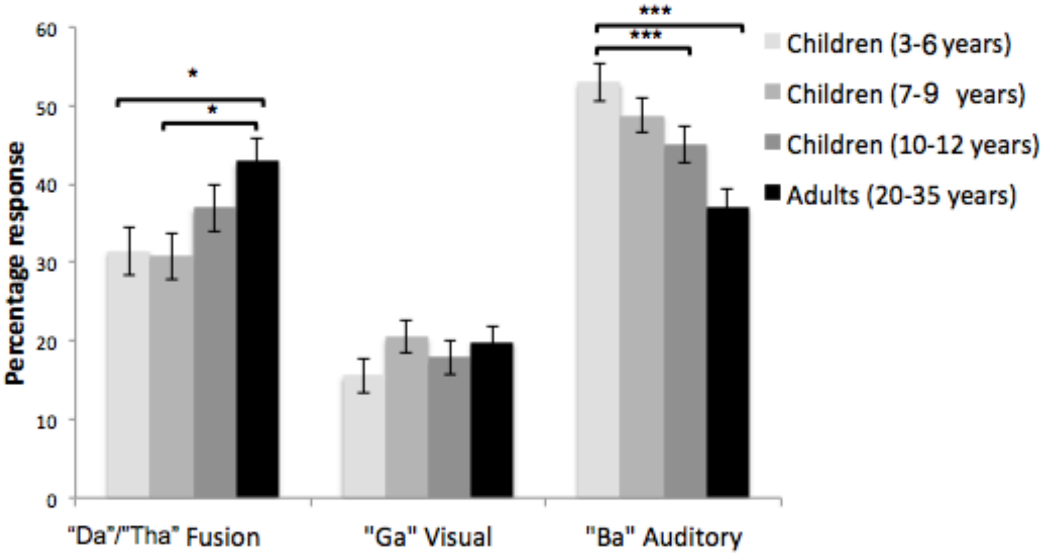


Figure 42. Mean number of "Da"/"Tha", "Ga" and "Ba" responses made on incongruent trials (visual "GA" paired with auditory "BA") in 3-6-year-olds, 7-9-year-olds, 10-12-year-olds and adults. "Da"/"Tha" responses indicate a fusion response (i.e. a response different from either the visual or auditory stimulus presented), "Ga" responses indicate a response to the visual stimulus and "Ba" responses indicate a response to the auditory stimulus. Error bars indicate standard error of the mean.

Do McGurk responses increase with development?

To assess whether responses made to incongruent trials could be predicted by age across childhood, three linear regression models were fitted to explore whether the percentage of fusion ("Da/Tha"), visual ("Ga") and correct auditory ("Ba") responses on incongruent McGurk trials was predicted by age (Figure 43). These models found no relationship between age and visual responses ($F(1, 89) = 1.07$, $p = .303$, $R = .11$, $R^2 = .01$) or fusion responses ($F(1, 89) = 1.88$, $p = .17$, $R = .15$, $R^2 = .02$), but a negative relationship between age and correct auditory ("Ba") responses ($F(1, 89) = 6.64$, $p = .012$, $R = .27$, $R^2 = .07$). Correct auditory responses decreased by 1.57% ($SE = .06$) with every year of age ($t(89) = 11.81$, $p < .001$).

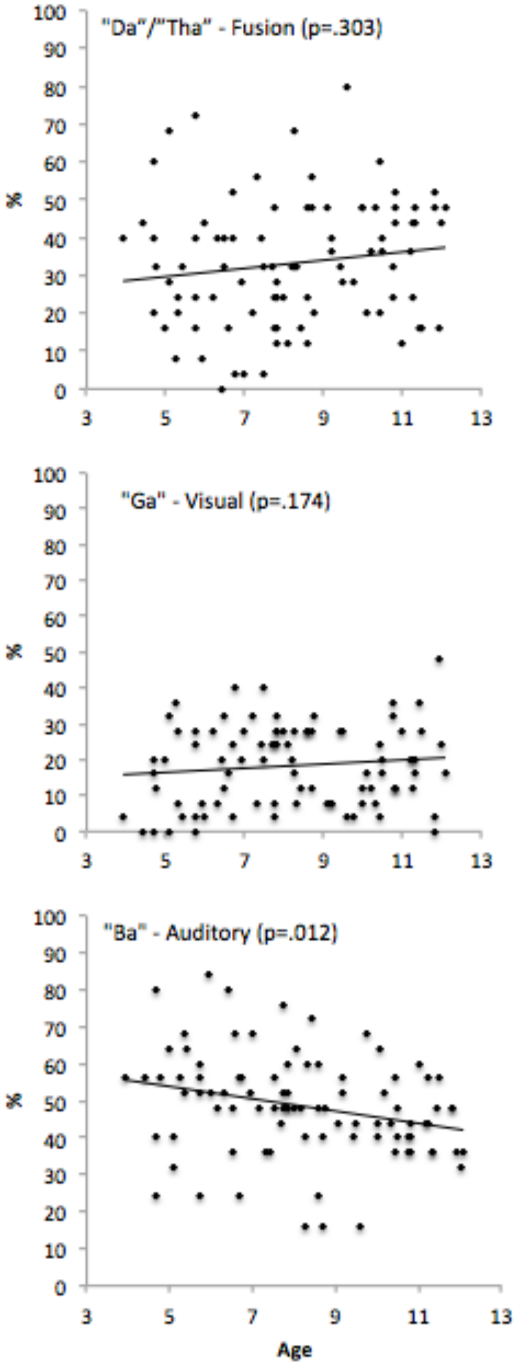


Figure 43. Correlation plots between age and fusion (a), visual (b), and auditory (c) responses to incongruent trials. R values indicate Pearsons correlations.

Effect of degrading sensory information on McGurk responses in adults and children

Thresholds were defined as the noise level inducing McGurk responses (i.e. “Ga”/“Da/Tha” responses) on 50% of trials, reflecting the point at which vision prevents correct auditory perception.

Three thresholds were identified for each participant:

1. McGurk responses in auditory noise (collapsed across visual noise conditions; i.e. the y axis of Figure 41a). Auditory noise would be expected to increase visually-driven responses, so this threshold reflects resistance to visual interference.
2. McGurk responses in visual noise (collapsed across auditory noise conditions; i.e. the x axis of Figure 41a). Visual noise would be expected to reduce the influence of the visual signal, so this threshold reflects dominance of the visual signal.
3. McGurk responses in combined auditory and visual noise. As only one data point per stimulus level was available for incongruent trials in each participant, a three-dimensional psychometric plane was fitted to data-points, and the threshold was identified as the centroid (mean) coordinate of coordinates yielding 50% accuracy. The change in position of this centroid reflects the audio-visual bias or dominance.

Participants were only included in threshold analyses if their threshold occurred within the range of noise presented. This left twenty two 3-6-year-olds, twenty six 7-9-year-olds, twenty six 10-12-year-olds and 21 adults available to compare thresholds in auditory noise, nineteen 3-6-year-olds, twenty seven 7-9-year-olds, twenty six 10-12-year-olds and 24 adults to compare thresholds in visual noise and twenty four 3-6-year-olds, thirty two 7-10-year-olds, twenty nine 10-12-year-olds and 32 adults to compare thresholds in combined noise. To identify the impact these exclusions had upon the probability of detecting an effect ($1-\beta err\ prob$, post hoc analyses were performed in G*power v 3.1 to assess the likelihood of detecting an effect given a critical alpha of .05, the available sample sizes and the observed effect sizes for each comparison (Faul et al., 2007). To aid interpretation I report $1-\beta err\ prob$ and $F_{critical}$ statistics alongside results. Critically, exclusion of these participants would not have biased results (i.e. the excluded participants did not show strong effects in the opposing direction to those reported here -see appendix page 334).

Table 15 and Figure 44 show thresholds for McGurk responses in visual and auditory noise *separately*. An ANOVA comparing thresholds for McGurk responses in auditory noise showed a significant effect of age group ($F(3, 91) = 6.55$, $p <.001$, $\eta^2 = .09$, $1-\beta err\ prob = .71$, $F_{critical} = 2.7$). Three to six year-olds and 7-9-year-olds did not significantly differ from one another ($p = 1$), both groups required significantly more noise to induce McGurk responses compared with adults (both comparisons $p =.003$). 10-12-year-olds did not significantly differ from 3-6-year-olds ($p = .076$), 7-9-year-olds ($p = .105$) or adults ($p = 1$).

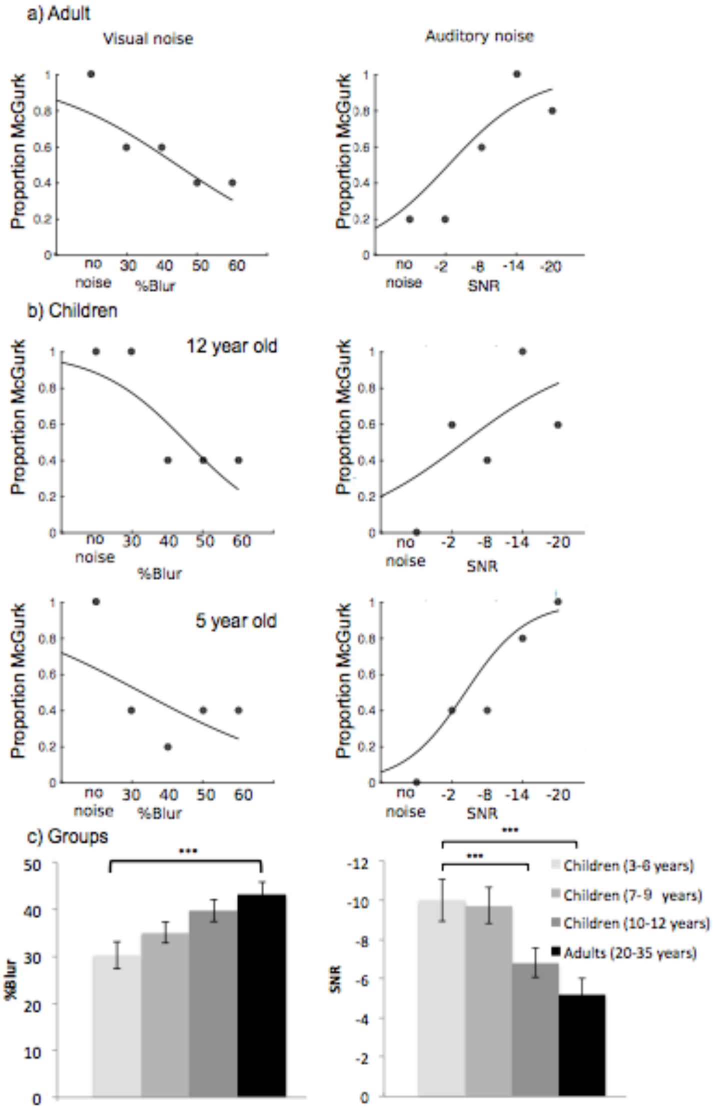


Figure 44. Example psychometric functions showing McGurk responses in (a) an adult participant and (b) children participants, a 12-year-old and a 5-year-old, and (c) average amount of auditory (right) and visual (left) noise required to induce McGurk responses in adults and children (aged 3-6, 7-9 and 10-12 years) with responses collapsed across the opposing noise level. Error bars indicate standard error of the mean.

Age group	**Visual Threshold (%Blur)**			**Auditory Threshold (SNR)**		
	M (SE)	*Lower CI*	*Upper CI*	*M (SE)*	*Lower CI*	*Upper CI*
3 – 6y	30.24 (2.83)	24.29	36.18	-10 (1.08)	-12.26	-7.75
7 – 9y	35.09 (2.2)	30.57	39.6	-9.72 (.9)	-11.59	-7.86
10 – 12y	39.84 (2.38)	34.94	44.74	-6.8 (.75)	-8.33	-5.26
Adults	43.15 (2.81)	37.33	48.96	-5.15 (.84)	-6.91	-3.4

Table 15. Mean thresholds for McGurk responses in visual and auditory noise (with noise conditions collapsed across the other modality) in each age group. *M* = mean; *SE* = standard error of the mean; *SNR* = signal-to-noise ratio; *y* = years; *CI* = 95% confidence interval.

A separate ANOVA comparing thresholds for McGurk responses in visual noise also showed a significant effect of age group (F(3, 92) = 4.48, p =.006, η^2 = .06, $1-\beta err\ prob$ = .54, $F_{critical}$ = 2.7) . Adults required more visual noise to eliminate McGurk responses compared with 3-6-year-olds (p = .006), but did not significantly differ from 7-9-year-old (p = .137) or 10-12-year-old (p = 1) groups. 10-12-year-olds also did not significantly differ from 7-9-year-olds (p = .988) or 3-6-year-olds (p = .072) and 7-9-year-olds did not differ from 3-6-year-olds (p = 1).

Developmental trajectory analysis

Two linear regression models assessing whether thresholds for McGurk responses could be predicted by age (Figure 45) showed age accounted for a significant proportion of variablity in McGurk responses in auditory noise (*F*(1, 73) = 7.68, *p* = .007, R = .31, R^2 = .10). The auditory noise level required to induce McGurk responses decreased by .64 SNR (SE = .23) per year (*t*(73) = -7.02, *p* <.001). Age also accounted for a significant proportion of variablity in McGurk responses in visual noise (*F*(1, 71) = 7.46, *p* = .008, R = .31, R^2 = .10). An increase in 1.74 (% blur, *SE* = .64) was required to eliminate McGurk responses per year (*t*(72) = 3.67, *p* < .001).

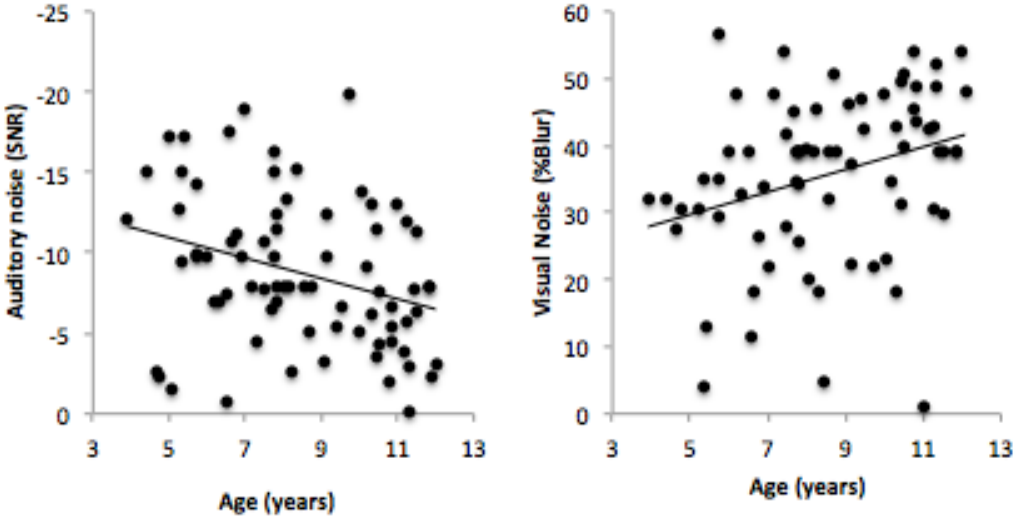


Figure 45. Correlations between age and the amount of auditory noise required to induce McGurk responses (left) and age and amount of visual noise required to prevent McGurk responses (right) in children. Younger children showed correct auditory responses even in higher levels of auditory noise and lower levels of visual noise compared with older children.

The effect of degrading both visual and auditory information

Table 16 and Figure 46 show thresholds for McGurk responses in *combined* visual and auditory noise. Two ANOVAs compared whether thresholds differed along the auditory and visual noise axes between age groups. Significant effects of age group occurred for the amount of visual noise (*F(*3, 91) = 5.52, *p* =.001, η^2 = .13, $1-\beta err\ prob$ = .95, $F_{critical}$ = 2.69) and auditory noise (*F(*3, 91) = 3.81 , *p* = .012, η^2 = .09, $1-\beta err\ prob$ = .81, $F_{critical}$ = 2.69) required to eliminate and induce McGurk responses.

Age	**Visual Axis (%Blur)**			**Auditory Axis (SNR)**		
	M (SE)	**Lower CI**	**Upper**	**M (SE)**	**Lower CI**	**Upper CI**
3 – 6 *y*	25.80 (1.81)	22.06	29.55	-12.05 (0.78)	-13.67	-10.44
7 – 9*y*	27.85 (1.63)	24.52	31.19	-11.42 (0.66)	-12.77	-10.07
10 – 12*y*	29.74 (0.86)	27.98	31.49	-10.95 (0.49)	-11.95	-9.95
Adults	33.72 (1.30)	31.07	36.38	-9.20 (0.59)	-10.4	-8.00

Table 16. Mean threshold positions for McGurk responses in combined visual and auditory noise in each age group. *M* = mean; *SE* = standard error of the mean; *SNR* = signal-to-noise ratio; *y* = years; *CI* = 95% confidence interval.

In combined noise, adults required significantly more visual noise to prevent McGurk responses compared with 3-6-year-olds (*p* = .001) and 7-9-year-olds (*p* = .018) but did not significantly differ from 10-12-year-olds (*p* = .278). Thresholds did not significantly differ between 10-12-year-olds and 3-6-year-olds (*p* = .405), 10-12-year-olds and 7-9-year-olds (*p* = 1) or 7-9-year-olds and 3-6-year-olds (*p* = 1). Adults also required less auditory noise to induce McGurk responses compared with 3-6-year-olds (*p* = .014) but did not differ from 7-9-year-olds (*p* = .063) or 10–12-year-olds (*p* = .284). Thresholds did not significantly differ between any of the child groups (*p* = 1 for all comparisons).

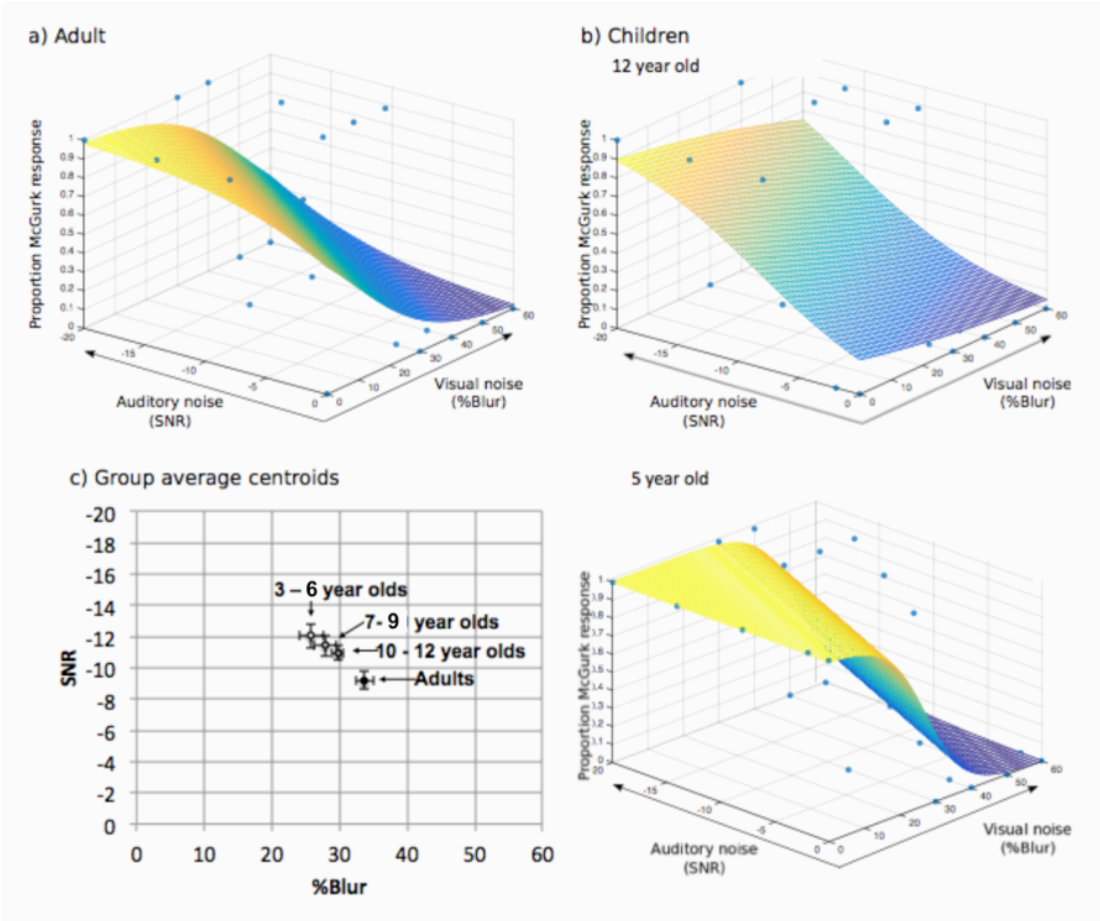


Figure 46. Example psychometric planes fitted for (a) a single adult and (b) two child participants (aged 12 and 5). The threshold was taken as the centroid co-ordinate of points crossing the 50% threshold for McGurk responses. The average centroid co-ordinates for children and adults are shown in (c). Error bars indicate standard error of the mean. There was a significant difference in centroids between groups; adults required more visual noise to prevent McGurk responses and less auditory noise to induce McGurk responses compared with children (i.e. children showed correct auditory responses even in lower levels of visual noise and higher levels of auditory noise compared with adults).

Discussion

This study was the first to use a threshold approach, inspired by computational models of McGurk perception (Magnotti & Beauchamp, 2015) and sensory weighting (Brooks et al., 2018; Ernst & Bülthoff, 2004; Fetsch et al., 2013; Witten & Knudsen, 2005), and combine audio-visual noise to examine developmental shifts in susceptibility to the McGurk effect. This approach exploits the effect of degrading signal quality on McGurk responses to gain a precise measure of sensory weighting, whilst also limiting the number of statistical comparisons (i.e. one threshold vs. comparison of means at each noise level).

It was hypothesised that adults would show more McGurk responses than younger children and that McGurk responses would increase with development. It was also predicted that McGurk responses would be influenced by visual and auditory noise in both adults and children (Sekiyama & Burnham, 2008), but that thresholds for McGurk responses would decrease (i.e. less auditory noise, more visual noise) through childhood into adulthood.

McGurk responses increase with development

Findings support a developmental shift in sensory dominance. Adults made more fusion responses and fewer correct auditory responses compared with 3-6-year-olds and 7-9-year-olds. However, 10-12-year-olds did not significantly differ from adults. Thus, in line with existing literature (Tremblay et al., 2007), these findings show the influence of vision over audition increases across development, reaching adult-like dominance by 10-12 years. An alternative explanation to sensory dominance is that younger children were poor at integrating auditory and visual information (Ernst, 2008). However, these data show that McGurk responses could be induced in young children depending upon the weighting of auditory and visual clarity (discussed below). Thus, I propose a role of sensory weighting (dominance) in influencing McGurk responses across development. Notably, this is not incompatible with an additional role of reduced multisensory integration in childhood.

Interestingly, contrary to a theoretical increase in visual dominance, the frequency of visual (“Ga”) responses to incongruent stimuli did not differ between age groups. One explanation of this is that participants erroneously identified “Ga” stimuli as “Da/Tha”, as shown in the analysis of congruent trials. Therefore, some “Da/Tha” responses may have actually reflected (incorrect) visual responses, rather than a fused percept. Nevertheless, increased fusion responses in adults and older children still indicates that visual information was more likely to alter auditory perception in adults compared with younger children.

Notably the range of fusion (“Da”/”Tha”) responses made on incongruent McGurk trials was highly varied in adults (12-76%) and children (0-80%). Such variance has been reported in adults (Mallick et al., 2015), the current findings extend this observed variability to childhood. Individual differences in adults have been attributed to variability in fronto-temporal connectivity required for integration (Keil, Müller, Ihssen, & Weisz, 2012). Connectivity differences also likely contribute towards developmental changes, as the underlying neural circuits supporting multisensory integration develop (Ernst, 2008; Stein, Stanford, & Rowland, 2014).

The threshold for McGurk responses in auditory and visual noise decreases with development

When comparing thresholds for McGurk responses in auditory and visual noise separately, adults required less auditory noise to induce McGurk responses compared with 3-6-year-olds and 7-9-year-olds but not 10-12-year-olds. They also required more visual noise to eliminate McGurk responses compared with 3-6-year-olds. However, they did not significantly differ from 7-9-year-olds or 10-12-year-olds. Regression analyses also showed threshold shifts occurred progressively across childhood. Thus, the weighting of visual and auditory information (dominance) shifts across development, such that vision influences auditory perception even under higher noise in adults and older children.

Interestingly, when comparing effects in combined noise, adults did not significantly differ from 7-9-year-olds in the auditory noise level inducing

McGurk responses, but did significantly differ from 7-9-year-olds as well as 3-6-year-olds in the amount of visual blur required to eliminate McGurk responses. Thus, when auditory and visual signals are both unreliable visual dominance appears immature in 7-9-year-olds (therefore a clearer signal is required for vision to dominate) whilst the influence of audition may be similar to adults (therefore similar auditory noise levels will prevent correct auditory responses). Nevertheless, differences between separate and combined noise comparisons may partially be explained by increased power retained by the combined vs. separate comparison (see limitations page 224).

Visual noise reduces, and auditory noise increases, the McGurk effect in both adults and children

McGurk responses were modulated by stimulus clarity in both adults and children. Increasing visual blur increased the amount of correct auditory responses. Increasing auditory noise decreased the amount of correct responses (Stacey et al., 2017). Thus, sensory reliability influenced audio-visual integration across age groups.

It might have been expected that children would be more susceptible to auditory noise (given auditory dominance) and therefore require less auditory noise to induce McGurk responses. Conversely, adults might be more susceptible to visual noise (given visual dominance) and require less visual noise to prevent McGurk responses. As this didn’t occur, dominance may map onto an ability to identify a relevant signal (i.e. speech sound or lip movement) within the dominant modality rather than general susceptibility to noise in that modality. This proposal also appears in line with the NED model of McGurk effects, as low sensory disparity (i.e. better detection of signal in noise) in the visual modality relative to the auditory modality predicts a higher proportion of McGurk responses. This hypothesis is also supported by findings showing children are more sensitive to change in auditory information whilst adults are more sensitive to change in visual information (Sloutsky & Napolitano, 2003), and findings showing lip reading ability (i.e. detecting a visual signal) predicts a higher influence of vision in speech perception (Massaro et al., 1986).

Limitations

The current study had several limitations that should be considered. Primarily, the task was limited in the number of trials presented (100 trials, 25 of which were incongruent). This limitation was due to the maximum time available for testing each child at Summer Scientist week (15 minutes per child) and was necessary to maintain young children’s attention throughout the task (all children included in analyses completed all 100 trials). The number of trials used was, however, comparable to previous studies assessing the McGurk effect in children (Narinesingh, Wan, Goltz, Chandrakumar, & Wong, 2014; Schorr et al., 2005). Nevertheless, gathering data from more trials over multiple testing sessions would enable fitting of separate two-dimensional psychometric functions to derive thresholds at each level of combined audio-visual noise (through holding noise constant in one modality and varying noise in the other). The findings from the current study provide strong justification for a more in depth investigation of developmental shifts in the McGurk effect using such an approach.

A second limitation to consider is that not all participants could be included in threshold comparisons. This was because derived thresholds fell outside the range of noise presented. Post-hoc power was therefore used to infer the likelihood of detecting effects with the remaining sample size. These statistics indicated that the primary comparison hindered was the effect of visual noise (collapsed across auditory noise conditions) as the test was limited to a 53% likelihood of rejecting the null. This comparison showed a significant difference between the youngest child group and adults, whilst other child groups did not significantly differ from adults. However, given the reduced sensitivity of this test, these null findings should be interpreted with caution. Fortunately, a more sensitive insight is gained when observing the effect of visual noise in the combined noise comparison, which retained a 94% probability of detecting true effects. Interestingly, in this comparison both 3-6-year-olds and 7-9-year-olds differ from adults in the amount of visual noise required to eliminate McGurk responses.

Conclusions

This study showed the threshold for the McGurk effect in audio-visual noise was lower in adults compared with 3-6-year-olds and 7-9-year-olds, but not 10-12-year-olds. Visual noise reduced McGurk responses and auditory noise increased McGurk responses in both adults and children; however the threshold for McGurk responses was lower in adults compared with younger children. These results suggest that susceptibility to the McGurk effect progressively increases, supporting a shift from auditory dominance in childhood towards adult-like visual dominance by the age of 10-12 years. This is consistent with the conclusions drawn in Chapter 7, suggesting a developmental shift in visual dominance.

Chapter 9: Does sensory dominance influence cross-modal distractibility?

Chapters 7 and 8 both showed sensory dominance shifts with development. Furthermore, Chapter 7 showed the Colavita effect was a robust measure of sensory dominance across studies. This final experimental Chapter presents a study in which I explored whether sensory dominance, measured using the Colavita effect, was associated with susceptibility to distraction from auditory and visual modalities.

Twenty-nine young adults (aged 22 – 46 years) took part in a Colavita task, a unimodal Stroop task and a cross-modal Stroop task requiring focus on vision, whilst ignoring audition. Correlational analysis between Colavita errors and Stroop performance showed unimodal response interference increased with increasing visual dominance. However, no significant correlations were found between cross-modal distractibility and sensory dominance. I consider this link with respect to visual experience (such as reading) driving visual dominance, thus making written words more distracting to visually dominant individuals.

Introduction

It is an intuitive hypothesis that individuals would be more distracted by sensory information presented to their dominant modality. However, to my knowledge, no study has explored this. Research prior to this thesis illustrated that children and adults treat visual and auditory stimuli differently in cross-modal attention tasks. Thomas et al, (2017) found that children aged 6 years were not disadvantaged by incongruent visual information in a cross-modal Stroop task. Conversely, Donohue, Appelbaum et al, (2013) and Yuval-Greenberg and Deouell (2009) both found that visual information had a stronger, negative, effect on auditory detection than vice versa in young adults. Guerreiro et al, (2010) also reported that older adults were able to focus on visual information whilst ignoring audition, but might not be able to focus on audition and ignore vision. In parallel with this, findings had suggested that children showed less visual dominance (Nava & Pavani, 2013; Wille & Ebersbach, 2016) whilst adults (Colavita, 1974) and older adults (Diaconescu et al., 2013) were more visually dominant. Experimental Chapters 7 and 8 of this thesis provided evidence for a developmental shift from audition towards visual dominance. Although beyond the scope of the current thesis, it is possible that more pronounced visual dominance occurs in ageing (Diaconescu et al., 2013). For example, Barnhart, Rivera, and Robinson (2018) found that, in a choice response time task, presenting simultaneous auditory information attenuated accuracy and slowed response times to visual targets in children (consistent with auditory dominance) whilst the reverse occurred in older adults. The aim of the study presented in this final experimental Chapter was to test the possible link between cross-modal attention and sensory dominance in a group of young adults.

It should be noted, that several mixed findings and results from the current thesis suggest that a link between sensory dominance and cross-modal distractibility cannot be assumed. First, although Guerreiro et al. (2010) found that older adults struggled to focus on audition whilst ignoring vision (Guerreiro et al., 2012, 2013; Guerreiro, Anguera, et al., 2014; Guerreiro & Van Gerven, 2011, 2017), supporting a sensory dominance account rather than generalised preserved ability, this was not the observation I made in Chapter 3, in which older adults could ignore both visual and auditory information. Further to this, in Chapter 3 children struggled to ignore both vision and audition, also suggesting generalised maturation rather than a role of sensory dominance. Given these mixed findings, the relationship between sensory dominance and cross-modal control requires further investigation.

A secondary aim of this study was to assess whether asymmetrical facilitation between vision and audition would be evident on incorrect Colavita trials. As outlined in Chapter 7, Sinnett et al, (2008) suggest Colavita errors may occur because auditory stimuli facilitate response times to visual targets but not vice versa. Consequently, the internal threshold for making a “visual-only” response is reached sooner when auditory and visual stimuli are presented together. However, the meta-analysis of response times in Chapter 7 did not show asymmetrical effects, rather, a general slowing was observed. One interpretation of this was that most studies only reported response times for trials in which participants responded correctly. Thus, if asymmetrical response time effects result in Colavita errors response times on correct trials would not reveal asymmetrical effects. This study provided the opportunity to test the hypothesis that asymmetrical effects would occur on trials in which participants made incorrect Colavita responses.

The main aims of the current study (in order of priority) were therefore:

1- Use a correlational approach to explore whether a link exists between sensory dominance and unimodal/cross-modal distractibility.
2- Identify whether asymmetrical facilitation and inhibition between vision and audition are evident on incorrect Colavita trials.
3- Replicate the finding from Chapters 3 and 5; that unimodal interference arises at both stimulus- and response-interference whilst cross-modal interference may occur mainly due to stimulus-interference.

Notably, although this study was inspired by the development and ageing literature, a sample of young adults was employed to investigate this exploratory question. The justification for this was pragmatic. If a relationship was found in young adults this would lay the foundations for future research investigating lifespan changes.

Method

Participants

Twenty-nine participants aged 22-46 (*M* = 26.13 years, 18 female, 27 right handed) took part. This sample size was used based on the power analysis conducted for the experiments reported in (Chapter 3) and considering the effect sizes observed for the Colavita effect with adults in the meta-analysis reported in (Chapter 7). A wider age range was utilised to obtain larger variance in sensory dominance better suited to a correlational analysis. Participants reported normal or corrected to normal hearing and vision. Thirteen participants spoke English as a first language, 16 reported English as their second language (mean age of English acquisition 8.31 years; range 2-20 years). Supplementary analyses were conducted to ensure Stroop performance was not modulated by first language (see appendix page 338).

Participants were students at the University of Nottingham and were paid £10 inconvenience allowance for their time. Data from one participant was excluded (missed 72% of visual Colavita targets), and three further participants did not survive outlier exclusion. This left 25 participants for analysis[31].

Equipment

The equipment used to present visual and auditory stimuli was identical to those used in Chapter 3.

Stimuli

Colavita stimuli

The visual target was a centrally presented white 1.5° circle. The auditory target was a binaurally presented 110Hz tone produced via PsychoPy v1.82 (Peirce, 2007, 2009). All targets had a duration of 47ms, closely resembling the 50ms duration used in previous literature (Koppen et al., 2009; Koppen & Spence, 2007c; Nava & Pavani, 2013), within the restrictions of the monitor's frame-rate. 60dB Brown noise produced via Audacity v2.0.6.0 was presented throughout. This enabled accurate manipulation of auditory stimuli despite ambient noise. The intensity of visual and auditory targets was set to 10x (20dB above) each participant's 79% detection threshold.

---

[31] The age range of the remaining participants remained the same (M = 26.48).

Stroop Stimuli

Stroop stimuli were identical to those used in Chapter 3 Experiment 1. However, “Babble” stimuli were used instead of the word “Brown” on neutral trials.

Procedure

The procedure consisted of 4 parts. First, thresholds were obtained for visual and auditory Colavita stimuli. Second, participants completed a Colavita task. Third, thresholds were obtained for visual and auditory Stroop stimuli. Finally, participants completed a unimodal and cross-modal Stroop task. The procedure used to derive thresholds for Stroop stimuli and the Stroop task were identical to Chapter 3 Experiment 1. I therefore only outline details related to thresholds for Colavita Stimuli and the Colavita task below.

Thresholds for Colavita stimuli

A 2-interval forced choice staircase was used to isolate detection thresholds for visual and auditory Colavita stimuli. Visual thresholds were always obtained first so that the experimenter (who was present throughout) could observe that the task was being performed correctly.

Participants were shown a “1” for 235ms followed by a blank screen for 1082ms. A “2” then signalled the start of the second 1028ms interval followed by a question mark. When the question mark was presented participants identified whether the target occurred in the first or second interval using the ‘A’ and ‘L’ keys, respectively. The target (a white circle or a tone – depending on threshold being derived) was presented randomly within the first 0-541ms of the interval followed by a 494ms break.

The intensity of the stimulus (i.e. opacity/dB) decreased following three correct responses and increased following a single incorrect response (visual starting intensity = 100% opacity, step sizes (% opacity) = [11.76, 7.84, 3.92, 3.92, 1.96, 0.39, 0.10, 0.04], auditory starting intensity = 60dB step size (dB)= 20,15,10,5,3,2,1). This converged upon the 79% threshold for detecting stimuli. Each staircase terminated when 8 reversals had been reached. Thresholds were taken as the average of the final 6 reversal values.

Colavita paradigm

The Colavita paradigm was designed using the most common parameters identified in Chapter 7. Figure 47 illustrates the Colavita task used. Using the 'A' and 'L' keys on a QWERTY keyboard, participants were instructed to identify visual, auditory and audio-visual targets by pressing one key in response to visual targets, the other in response to auditory targets and both in response to audio-visual targets. Participants completed four 100-trial blocks each containing 40 auditory targets, 40 visual targets and 20 audio-visual targets in a randomised order. Each block was preceded by 10 practice trials containing 4 auditory, 4 visual and 2 audio-visual targets in a randomised order. Within each trial, stimuli were presented with a variable onset of 0-494ms and presented for 47ms followed by an interval of 2000ms. The next stimulus was always presented following this interval of 2000-2494ms even if no response was made.

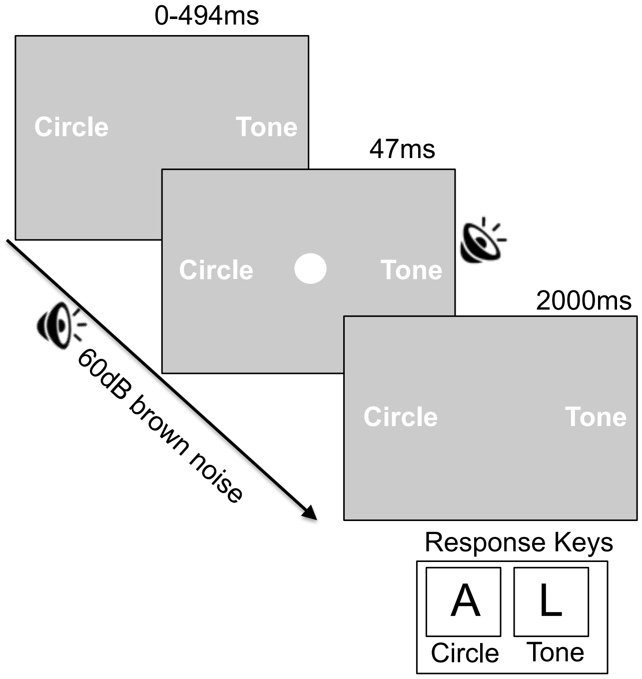


Figure 47. Schematic illustration of a bimodal trial in the Colavita task. Cues were presented indicating the correct buttons to press for visual and auditory targets (these switched following each 100 trial block). Following a 0-494ms interval participants were presented with either a visual, auditory or bimodal (audio-visual) target. Participants were instructed to press one key to visual targets, the other to auditory targets and both keys on bimodal trials. Targets were followed by a 2000ms break. The next trial always began following this break. Stimuli not to scale.

To prevent an effect of handedness, auditory and visual response buttons swapped following each 100-trial block. Participants were given 10 additional practice trials at the start of each block to accustom themselves with the new response mappings. The side to which each response was mapped was visible on the screen at all times (i.e. if visual targets were mapped to the ‘A’ key and auditory targets were mapped to the ‘L’ key the word ‘circle’ was displayed 12° to the left of fixation and the word ‘tone’ was displayed 12° to the right of fixation). The starting order of response mappings (i.e. visual to left and auditory to right or vice versa) was counterbalanced across participants.

Analysis

As in previous chapters, outliers in each individual’s RT data (for Colavita and Stroop tasks) were removed if they exceeded +/-3 times the absolute deviation from the median (Leys et al., 2013). Outlying participants were removed based upon the mahalanobis distance of each participant from group performance in the Stroop task based upon the relationship of accuracy and RT measures between congruent, stimulus-incongruent and response-incongruent conditions. Four participants were removed as outliers (final n = 25). Four sets of analyses were then performed:

1- Analysis of sensory dominance to establish whether a Colavita effect occurred
2- Analysis of Stroop performance to establish whether, in line with our previous experiments, unimodal interference occurs at both stimulus and response levels whilst cross-modal interference occurs at stimulus levels[32].
3- Correlation of sensory dominance with Stroop performance to establish whether sensory dominance is associated with cross-modal and unimodal distractibility.

---

[32] As with previous analyses, ratio scores were analysed. Analysis of all congruency conditions are shown in appendix, page 339.

4- Analysis of response times on Colavita task to establish whether response times were faster for Colavita errors vs. unimodal visual responses (in line with asymmetrical facilitation and inhibition).

As in the preceding experimental chapters (with the exception of Chapter 3) I refer to *p* values <=.05 as significant.

Results

Colavita errors

Figure 48 illustrates the errors made on unimodal visual, unimodal auditory and bimodal trials. A three-way ANOVA showed no significant effect of target type (*F*(2, 48) = .32, *p* = .728, η_p^2 = .01) suggesting a similar number of errors were made to unimodal visual (*M* = 4.07%, *SE* = .066), unimodal auditory (*M* = 4.13%, *SE* = .8) and bimodal (*M* = 9.3%, *SE* = 1.71) targets.

A paired samples t-test compared the percentage of visual-only responses made on bimodal trials (the Colavita effect) vs. the percentage of auditory-only responses made on audio-visual trials (the “reverse” Colavita effect; Figure 49). There was no significant difference between visual-only responses (*M* = 4.74%, *SE* = 1.06) vs. auditory-only responses (*M* = 4.44%, *SE* = 1.03) (*t*(24) = .244, *p* = .809, *Cohens d* = .05), thus a Colavita effect was not found when considering average scores across participants.

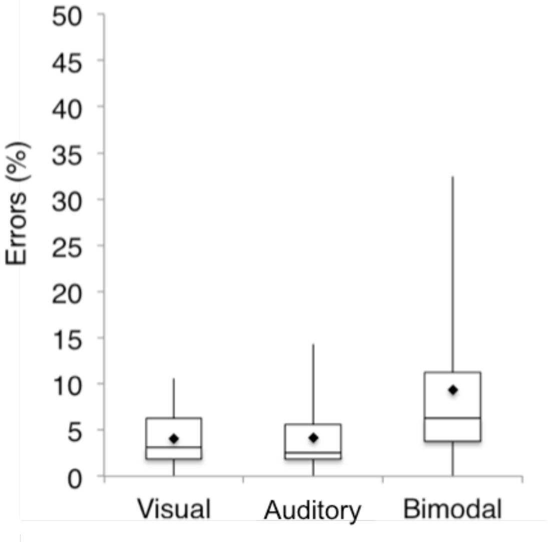


Figure 48. Errors made on unimodal visual, auditory and bimodal trials. Black diamonds indicate mean used for analysis. Central line indicates the median. Whiskers and boxes indicate interquartile range.

Colavita errors in the first 100 trials

Contrary to previous studies the present study required participants to switch hands between each block. This may have increased task difficulty and enhanced participants' awareness of response accuracy. To explore this, a paired samples t-test was conducted to compare "visual-only" responses to "auditory only" responses made to bimodal targets within the first 100 trials (i.e. before any change in response mapping occurred). This analysis showed that in the first 100 trials participants made significantly more "visual only" responses (*M* = 7.8%, *SE* = 1.7) than "auditory only" responses (*M* = 2.8%, *SE* = .9; *t*(24) = 3.61, *p* = .001, *Cohens d* = .73). This yielded an effect size similar to that seen across adult studies within the meta-analysis (Chapter 7). Thus, although there was a Colavita effect in the first 100 trials, this might have been attenuated by task difficulty following response switching.

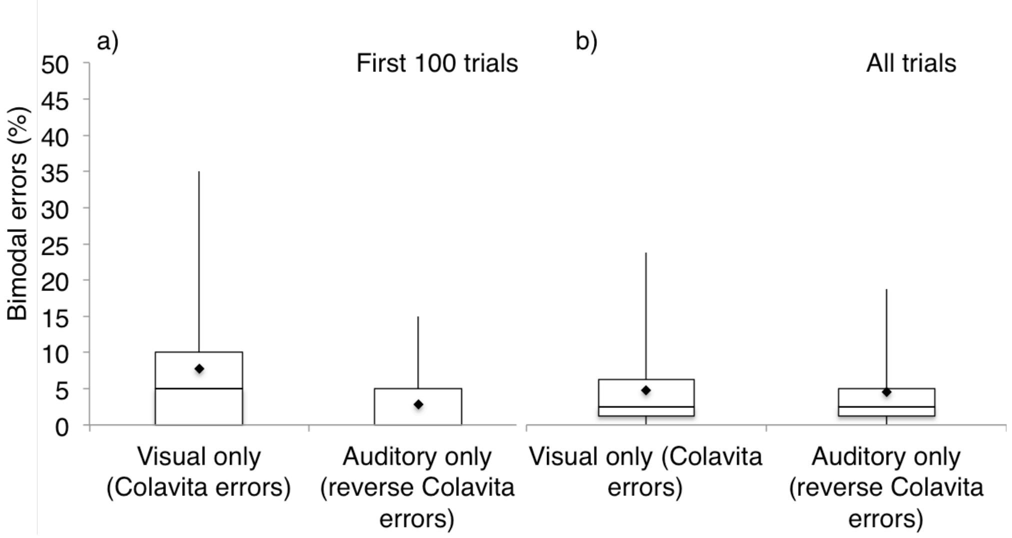


Figure 49. Bimodal errors resulting from visual (Colavita) errors and auditory (reverse Colavita) errors. These are shown across all trials and in the first 100 trials before response switching occurred. Means used for statistical analyses are indicated by the black diamond.

Stroop task

When analysing Stroop effects, one goal was to replicate the effects found in previous experiments (i.e. that unimodal interference results from stimulus- and response-interference whilst cross-modal interference arises mainly from stimulus-interference). Because this hypothesis specifically related to stimulus- and response-interference I do not present general interference ratios in this Chapter. A second aim was to identify whether individual differences in sensory dominance (i.e. Colavita errors vs. reverse Colavita errors) are related to distractibility from auditory and visual sources. Specifically, it was predicted that a negative correlation would be seen between visual dominance and cross-modal Stroop effects (i.e. auditory distraction). Thus, increased visual dominance should result in less distraction from auditory sources in the cross-modal Stroop task.

Replication of previous results

Figure 50 shows ratio scores for stimulus- and response-interference under unimodal and cross-modal conditions in terms of response time and accuracy.

Response times

A 2 (sensory condition) x 2 (interference type) ANOVA of response time ratios showed no main effect of sensory condition (F (1,24) = 3.85, p = .062, η_p^2 = .14), no main effect of interference type (F (1,24) = 4.18, p = .052, η_p^2 = .15) and no interaction between sensory condition and interference type (F (1,24) = .07, p = .794, η_p^2 <.01).

Accuracy

A 2x2 ANOVA of accuracy ratios showed no main effect of sensory condition (*F* (1,24) = .01, *p* = .911, η_p^2<.01) and no interaction between sensory condition and interference type (*F* (1,24) = .002, *p* = .967, η_p^2<.01) but a main effect of interference type (*F* (1,24) = 6.43, *p* = .018, η_p^2=.21). This effect occurred because accuracy was reduced by response-interference and not stimulus-interference (Figure 50). Interestingly, Bonferroni corrected t-tests comparing accuracy ratios to 1 showed that response-interference only significantly differed from 1 under unimodal conditions. This is consistent with the theory that response interference does not reduce accuracy under cross-modal conditions. However, given that the ANOVA showed no interaction between sensory condition and interference type for either accuracy of response time, this is not a complete replication of the previous reported results.

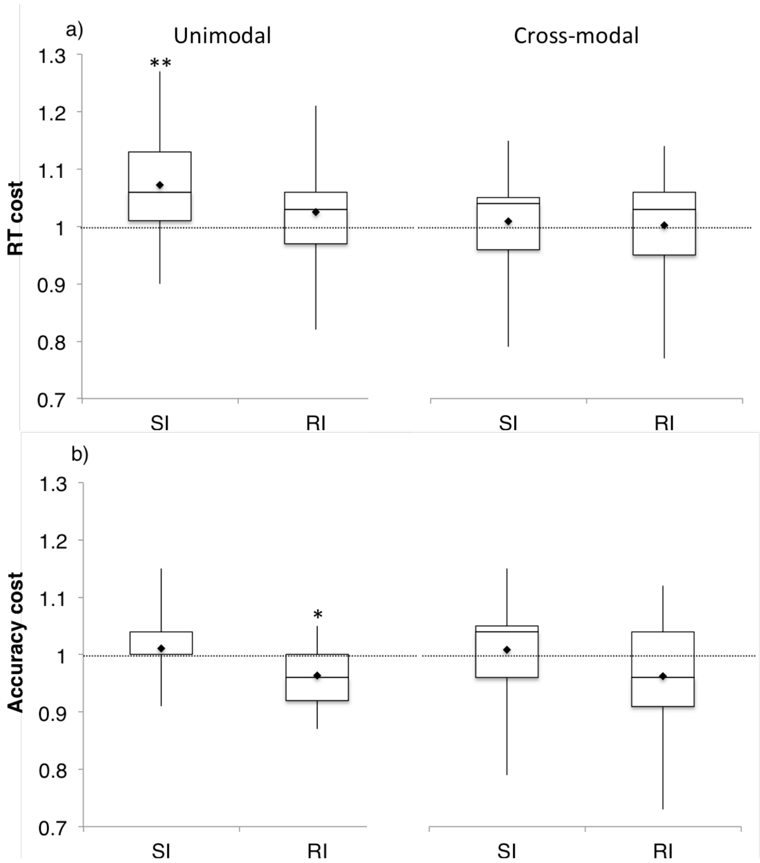


Figure 50. Response time (RT) and Accuracy costs associated with stimulus interference (SI) and response interference (RI) under unimodal (left) and cross-modal (right) conditions. For RT, scores higher than 1 indicate slowing on incongruent vs. congruent conditions. For Accuracy, scores lower than 1 indicate lower accuracy on incongruent vs. congruent conditions. Black diamonds indicate means. Asterisks show Bonferroni corrected t-tests (corrected for 4 comparisons) comparing ratios to 1; ***=<.001, **=<.01, *=<.05.

Is there a relationship between sensory dominance and cross-modal distractibility?

Sensory dominance was defined as:

Equation 14

$$\text{Sensory Dominance} = \text{Total Colavita Error} - \text{Total Reverse Colavita Error}$$

where a “Colavita Error” is a visual only response to a bimodal trial and a “Reverse Colavita error” is an auditory only response to a bimodal trial. A positive sensory dominance score therefore indicates visual dominance, whilst negative scores indicate auditory dominance. To avoid handedness effects, each participant’s sensory dominance was calculated based upon all trials.

Figure 51a shows the resulting Pearson correlations between sensory dominance scores and response time ratios. The only measure sensory dominance was significantly correlated with was unimodal response-interference in terms of reaction times. Figure 51b shows the resulting Pearson correlations between sensory dominance and accuracy ratios. Sensory dominance was not significantly correlated with accuracy costs in any condition.

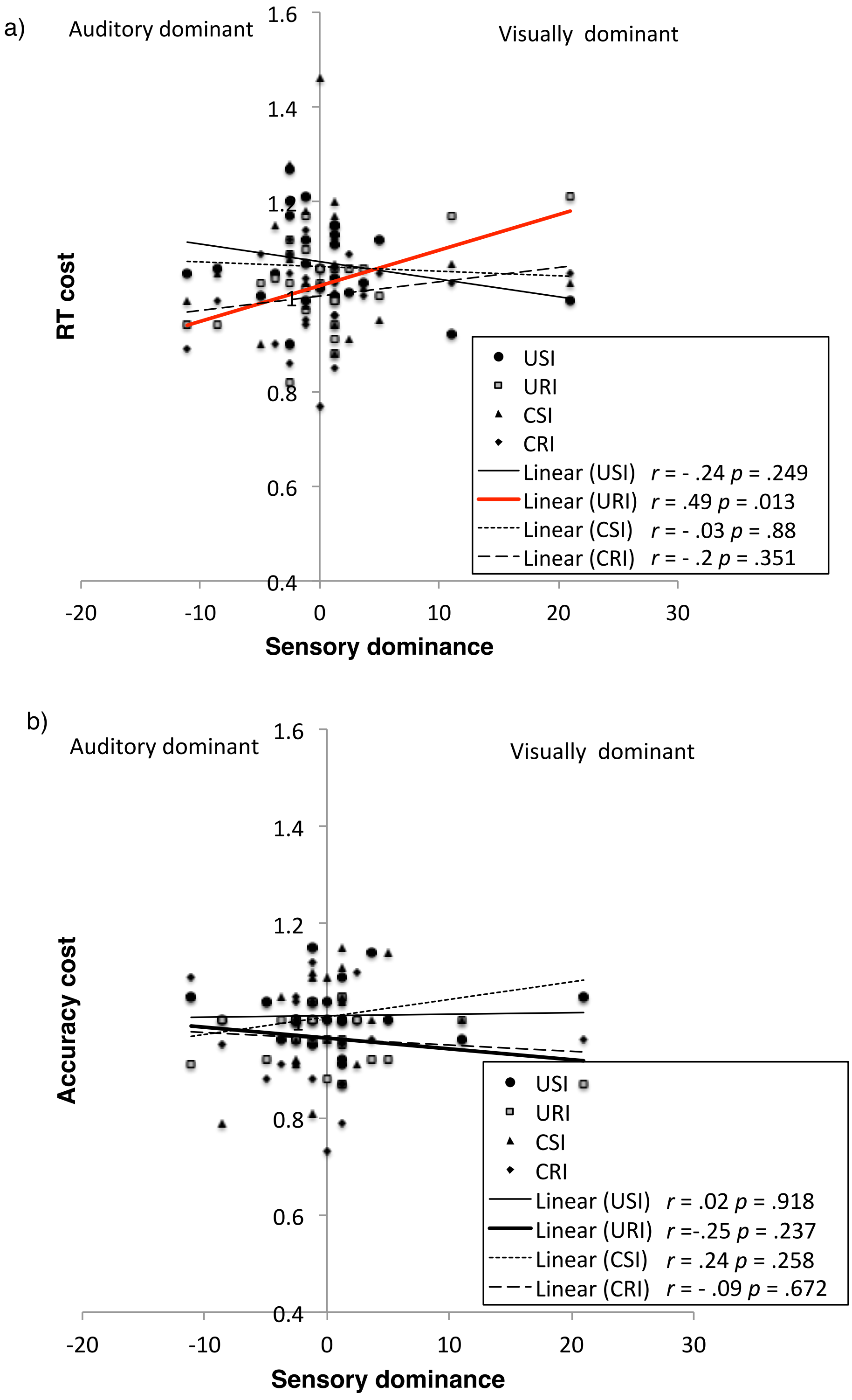


Figure 51. Pearsons correlations between Response time (RT) (a) and accuracy ratios (b) with sensory dominance. A negative sensory dominance score indicates auditory dominance, positive scores indicate visual dominance. Response time costs higher than 1 indicate slowing on incongruent vs. congruent conditions. Accuracy costs lower than 1 indicate response times were less accurate for incongruent vs. congruent. No correlations reached significance for accuracy. USI = unimodal stimulus interference, URI = unimodal response interference, CSI = cross-modal stimulus interference, CRI = cross-modal response interference.

Colavita response times

A final aim of the current study was to investigate whether Colavita errors occur due to asymmetrical inhibition and facilitation. To assess this, I compared response times to visual and auditory stimuli under unimodal and bimodal conditions but also separated trials in which correct responses were made and trials in which incorrect responses were made, each of these response times are shown in Table 17. The following were hypothesised:

1- When correct responses are made to bimodal targets, response times would be slower to visual and auditory stimuli compared with unimodal trials (as shown in Chapter 7).

2- When incorrect responses are made to bimodal targets, response times would be faster to visual stimuli compared with unimodal conditions (thus resulting in Colavita errors).

		Min	**Max**	***M***	***SE***
All trials					
Unimodal	Visual	499	959	667	23
	Auditory	500	873	706	18
Bimodal (correct)	Visual	537	923	720	18
	Auditory	545	1005	750	21
Bimodal (errors)	Visual	431	2271	738	86
	Auditory	404	1588	838	61
First 100 trials					
Unimodal	Visual	435	1082	659	26
	Auditory	531	1022	724	25
Bimodal (Correct)	Visual	506	955	719	23
	Auditory	583	1034	763	23
Bimodal (Errors)	Visual	474	2271	742	111
	Auditory	108	881	552	92

Table 17. Response times (ms) to visual and auditory targets under unimodal and bimodal conditions. Response times under bimodal conditions are shown for when participants correctly pressed both buttons and erroneously pressed only the visual or auditory response key. *M* = mean, *SE* = standard error

Correct responses

A 2 (sensory modality: visual vs. auditory) x 2 (sensory condition: unimodal vs. bimodal) ANOVA was conducted to compare response times to visual and auditory targets under unimodal and bimodal conditions. Response times were significantly faster to visual (*M* = 693ms, *SE* = 16) vs. auditory (*M* = 728ms, *SE* = 18) targets (*F*(1, 24) = 10.59, *p* = .003, η_p^2 = .31). Response times were also significantly faster under unimodal (*M* = 686ms, *SE* = 18) vs. bimodal (*M* = 734ms, *SE* = 18) conditions (*F* (1,24) = 10.75, *p* = .003, η_p^2 = .31). There was no significant interaction between sensory modality and sensory condition (*F* (1,24) = .46, *p* = .504, η_p^2 = .02). This finding was in line with the results of the meta-analysis in Chapter 7 and the hypothesis that, when correct responses were made, response times would be generally slower under bimodal vs. unimodal conditions.

Incorrect responses

A 2x2 ANOVA was conducted to compare correct responses to visual and auditory stimuli on unimodal trials to incorrect unimodal responses made to visual and auditory stimuli on bimodal trials (i.e. Colavita errors and reverse Colavita errors). The analysis showed no main effect of target modality (*F* (1,18[33]) = .35, *p* = .561, η_p^2 = .02), no main effect of whether targets were presented unimodally or cross-modally (*F* (1,18) = 1.91, *p* = .184, η_p^2 = .1) and no interaction between target modality and sensory condition (*F* (1,18) = .028, *p* = .869, η_p^2 < .01). This suggests that, contrary to hypotheses, Colavita errors did not occur significantly faster than correct unimodal responses to visual targets. Nevertheless, the effect of slowing under bimodal conditions was not evidenced within this analysis.

---

[33] Only 19 participants made errors, thus degrees of freedom = 18 for this comparison.

Discussion

The current study yielded several interesting findings. Firstly, when considering all 400 Colavita trials a Colavita effect, on average, was not found. However, an effect was evident in the first 100 trials. Secondly, in the Stroop task, a sensory condition x interference type interaction was not found. This was not in line with the previous finding that cross-modal interference arises from stimulus- but not response-interference whilst unimodal interference arises from both. Thirdly, a positive correlation was found between visual dominance and unimodal response interference, however no further significant correlations were found between Stroop performance and sensory dominance. Fourthly, and finally, there was no evidence to suggest asymmetrical response time effects resulted in Colavita errors.

With regards to Colavita errors, it is surprising that a strong Colavita effect was not seen, given the results of the meta-analysis in Chapter 7. This analysis showed a medium, approaching large, effect size in adults, and the current study used the parameters used most frequently across studies. However, the meta-analysis did show a significant proportion of unexplained variability even after considering multiple experimental manipulations as moderator variables. The main difference between the current study and those included within the meta-analysis was that participants switched response hands at the end of every 100 trials. This was designed to prevent confounding handedness and sensory dominance within each subject, and allow the use of sensory dominance scores as a correlational measure. However it is possible that this increased the task demands, resulting in fewer Colavita errors. Indeed, within the meta-analysis only two experiments showed reverse Colavita effects (Ngo et al., 2011, 2010). These experiments used an n-back version of the Colavita effect, and attribute observed auditory dominance due to the temporal demands of the task (Welch & Warren, 1986). However, it is also possible that this task was more difficult, and therefore, as with the current study, Colavita errors were not observed.

Nevertheless, these studies saw a reversal of the Colavita effect and this was not the case in the current study.

Regarding Stroop effects, the current study did not replicate the finding that cross-modal interference arises from stimulus- not response-interference whilst unimodal interference arises from both. This failure to replicate is seen in the lack of significant interactions within either the response time or accuracy data, suggesting a similar pattern of effects between unimodal and cross-modal conditions. Interestingly, however, Bonferroni corrected t-tests comparing ratios to 1 showed that only unimodal stimulus interference significantly slowed response times relative to baseline (i.e. response times on congruent trials). Furthermore, only unimodal response-interference reduced accuracy relative to baseline. Although this does not fully replicate the pattern of effects seen in the previous experiments of this thesis, this suggests that cross-modal interference may have been particularly small in the current study, and therefore significant effects were not observed.

The primary aim of the current investigation was to explore whether links exist between sensory dominance and cross-modal distractibility. In the introduction to this thesis it was speculated that increasing visual dominance might explain why previous literature had showed children were more suceptible to auditory distraction whilst adults were more suceptible to visual distraction. However, findings in this thesis (Chapter 3) suggested children were suceptable to both visual and auditory distraction, and older adults were also able to focus on audition and ignore vision. These findings suggest sensory dominance cannot fully account for preserved cross-modal interference control in ageing.

The current results showed no significant correlations between cross-modal distractibility and sensory dominance. Thus, I cannot conclude that increasing visual dominance reduces distraction from auditory sources (Guerreiro et al., 2010). However, several limitations must be considered. Primarily, error rate overall was very low and the range of sensory dominance scores was focused between -5 (made 5 more auditory only errors than visual only) and 5 (made 5 more visual only errors than auditory only). It is therefore possible that the current sample did not yield a broad enough range of sensory dominance scores to show links with cross-modal distractibility. Secondly, this study was an exploratory study with a young adult sample, it is possible that including older adults and children, who are typically more heterogeneous in nature, would have yielded a wider range of sensory dominance scores and addressed age-related hypotheses. Nevertheless, this was an exploratory investigation that provided key preliminary findings.

One unexpected result was that increasing visual dominance increases unimodal distractibility (i.e. distraction arising from written words). One interpretation of this is that sensory dominance emerges due to experiences such as learning to read. If this is the case, then written words (such as those presented in the unimodal Stroop task in this thesis) might be particularly distracting for visually dominant individuals. A similar, although subtly different, hypothesis is that written words are more distracting due to familiarity (rather than being visual per se). Evidence partly supporting this hypothesis comes from the finding that sensory dominance within a task can be influenced by object familiarity. For example, Setti and Chan (2011) found that in the sound-induced-flash-illusion, in which audition usually dominates, presenting a familiar visual object (such as a face) lowers this effect (swaying perception towards visual influences). It is possible that the relationship between visual dominance and the effect of the written word distractor arises due to object familiarity of written words. However, it is unclear why a similar level of familiarity would not also be seen for spoken words, given that humans are exposed to spoken language earlier in life compared with written language. Further research investigating the role of stimulus familiarity on auditory and visual distraction is needed.

Conclusions

In sum, although the results of the meta-analysis presented in Chapter 7 suggested the Colavita effect is empirically robust, the current study did not find a Colavita effect in adult participants. However, when considering the first set of trials (before response keys were switched) an effect was observed. This suggests task difficulty as a potential moderator of the Colavita effect. Further to this, the current findings do not suggest increased visual dominance results in reduced distraction from auditory sources. However, increasing visual dominance was correlated with unimodal response-interference. Future research should investigate whether visual dominance is associated with increased visual experience (and therefore increased susceptibility to visual distractors).

Chapter 10: General discussion

The aim of this thesis was to identify whether similar mechanisms underpin unimodal and cross-modal (audio-visual) interference control. This was addressed through a) comparing lifespan trajectories in unimodal and cross-modal interference, b) comparing the processing levels at which unimodal and cross-modal interference occur, and c) examining whether changes in sensory dominance occur with development and if this might influence distractibility from different senses. In this Chapter I discuss how the presented research answer each of the questions presented at the start of this thesis. I then discuss further gaps in the literature that this research fills and consider limitations and future directions.

Summary of findings

- Chapter 3
    - Unimodal interference improves with development and deteriorates with age. Cross-modal interference improves with development but does not deteriorate in old age (when ignoring both vision and audition).
    - Unimodal interference arises from stimulus- and response-interference. Cross-modal interference is mainly due to stimulus-interference.
- Chapter 4
    - Limiting colour information did not increase Stroop interference in young adults (in terms of stimulus- or response-interference).
    - The N450 effect was associated with stimulus-interference.
    - Stimulus- and response-locked lateralized readiness potentials (LPRs) did not differ between congruency conditions.
- Chapter 5
    - The N450 effect did not significantly differ between unimodal and cross-modal conditions.
    - The N450 effect was sensitive to response-interference.
    - The stimulus-locked LRP indicated a difference between cross-modal and unimodal response interference (such that more interference was seen under cross-modal conditions).
    - The response-locked LRP did not manifest congruency effects.
- Chapter 6

  - A portable EEG headset (Emotiv EPOC) was capable of measuring the N450 effect, suggesting this could be used to measure interference in the real world.
- Chapter 7
  - The Colavita effect was found to be a robust measure of sensory (visual) dominance.
  - Adults showed visual dominance whilst children did not.
  - Visual stimuli slowed response times to auditory targets and vice versa (contrary to the theory that audition should facilitate visual responses).
- Chapter 8
  - The threshold for the McGurk effect in audio-visual noise decreased with development, supporting a developmental shift from auditory to visual dominance.
- Chapter 9
  - Sensory dominance, measured using the Colavita effect, was not correlated with cross-modal (auditory) distraction, but was positively correlated with unimodal Stroop interference.

In my introduction to this thesis I specified three questions I aimed to answer:

1. Do unimodal and cross-modal interference arise from the same or different mechanisms?
2. Does a shift in sensory dominance occur across development?
3. Is sensory dominance associated with cross-modal interference?

In this general discussion, I first focus on the mechanisms underlying unimodal and cross-modal interference, then I discuss the implications of sensory dominance research, finally I outline additional gaps in the literature addressed by this thesis. Throughout, I will raise the limitations of experiments presented in this thesis and propose directions for future research. However, I provide a more in depth outline of limitations and future directions before concluding.

Do unimodal and cross-modal interference arise from the same or different mechanisms?

Several findings reported prior to this thesis suggested unimodal and cross-modal interference control occur via different mechanisms. Firstly, the limited cross-modal Stroop literature had shown cross-modal Stroop effects were lower in magnitude than unimodal effects (Elliott et al., 2014; Francis et al.,

2017) and peaked at different points in the response time distribution (Elliott et al., 2014). It had also been observed that cross-modal effects showed stronger facilitation on congruent trials, in contrast to unimodal results, showing equal facilitation and interference (Donohue, Appelbaum, et al., 2013). Furthermore, unimodal Stroop interference appeared correlated with working memory capacity, whilst the cross-modal Stroop effects did not (Morey et al., 2012). Neuroimaging literature had also suggested that a fronto-parietal network, including the lateral prefrontal cortex (LPFC) and anterior cingulate cortex (ACC) underpins unimodal interference control. Although some findings suggested these regions were involved in unimodal and cross-modal tasks (J. A. Johnson et al., 2007; Weissman et al., 2004), it had also been suggested that cross-modal interference control may be less reliant on prefrontal mechanisms (Guerreiro et al., 2010) and may be achieved at earlier stages of processing such as sensory gating (Haxby et al., 1994).

In this thesis, I used two approaches to identify whether similar or different mechanisms underpin unimodal and cross-modal interference control. First, I explored whether unimodal and cross-modal Stroop interference manifest similar patterns of developmental maturation and age-related decline. Second, I compared whether unimodal and cross-modal interference occur at similar levels of processing. I will discuss the findings from these approaches in turn.

Prior to this thesis, only three studies had investigated cross-modal Stroop effects in development and none in ageing. It was generally assumed that both unimodal and cross-modal interference control processes develop and deteriorate across the lifespan, in line with the maturation and deterioration of the prefrontal cortex (Comalli et al., 1962; Dempster & Vegas, 1992). However, in Chapter 3, I showed that unimodal and cross-modal interference control follow different lifespan trajectories, suggesting different underlying mechanisms. Maintained cross-modal control in ageing had been previously reported (Guerreiro et al., 2010), and attributed to lower reliance on prefrontal mechanisms. This might explain why working memory capacity, a function typically attributed to the dorsolateral prefrontal cortex, was found to correlate with unimodal, but not cross-modal attention (Morey et al., 2012).

The theory that cross-modal interference control can be implemented at earlier levels of processing is in line with several findings in this thesis. First, cross-modal interference was spared in ageing. Second cross-modal interference arose from stimulus-interference whilst unimodal interference was also caused by response-interference. Third, cross-modal interference showed congruency effects in early stimulus-locked lateralized readiness potentials (LRPs) associated with stimulus-interference (Smulders et al., 1995). If cross-modal control can be implemented at earlier stages of processing, i.e. via sensory gating, it makes sense that interfering stimuli are “blocked out” prior to the level of response selection, thus resulting in stimulus-interference but no response-interference.

However, if the cross-modal interference control is less reliant on prefrontal processes, several findings remain unclear. First, developmental and age-related changes in executive function have both been attributed to the development and decline of the prefrontal cortex (Dempster & Vegas, 1992). As such, it is unclear why cross-modal control would be maintained in ageing but poor in childhood. One explanation could be that gating mechanisms also undergo protracted development, but are less susceptible to ageing. This explanation would also account for why children showed cross-modal interference associated with response-interference, suggesting that they failed to “block out” cross-modal distraction prior to response-selection stages. However, in contrast to this account, it has been shown that sensory gating develops very early in infancy, between 1-4 months (Kisley, Polk, Ross, Levisohn, & Freedman, 2003), and measures of sensory gating appear consistent between 3 months to 4 years of age (Hunter, Gillow, & Ross, 2015). Thus, poor cross-modal control in childhood cannot be attributed to protracted development of gating.

It is possible that gating mechanisms are triggered by the anterior cingulate cortex (ACC) via cortico-thalamic connections, and that the sensitivity of the ACC to conflict differs between children and older adults. It is known that the dorsal ACC has excitatory connections with the mediodorsal thalamus (Ouhaz, Fleming, & Mitchell, 2018) an area that has been shown to gate the flow of information between cortical regions (for review see Gisiger & Boukadoum, 2011). For example, the mediodorsal thalamus has been shown to gate information flow from the hippocampus to the prefrontal cortex (Floresco & Grace, 2003). The sensitivity of the ACC to Stroop conflict has been shown to progressively increase between the ages of 6 and 24 years (Adleman, 2002). In contrast to this, in older adults, the ACC appears more sensitive to conflict in the Stroop task, and this has been attributed to increased susceptibility to error in ageing, and therefore increased need to monitor conflict (Milham et al., 2002). If the ACC acts to trigger sensory gating in cross-modal Stroop tasks, and the ACC remains sensitive to conflict in ageing, but shows protracted development, this might explain poor cross-modal control in childhood but maintained effects in ageing.

One finding in this thesis supporting the role of the ACC in cross-modal control is the similar N450 amplitude observed between sensory conditions (Chapter 5). The N450 is an ERP marker of conflict detection and has been localised to the dorsal ACC (West et al., 2004). In Chapter 4, I found that the N450 was associated with stimulus-interference, because, although there was a significant difference in ERP amplitude between congruent and stimulus-incongruent conditions, response conflict did not increase this difference. This conclusion supported previous results linking the N450 to stimulus-interference (Killikelly & Szűcs, 2013; Szűcs & Soltész, 2012) and the finding that the ACC is sensitive to non-response conflict (West et al., 2004). As such, similar N450 effects between unimodal and cross-modal tasks suggests ACC involvement in both tasks. I therefore speculate that the ACC acts to trigger sensory gating prior to response selection stages, thus resulting in no response-interference on cross-modal trials. However, it should be noted that in Chapter 5 the N450 was also sensitive to response conflict, and was larger under unimodal conditions (although differences did not reach significance). Further research is therefore needed to establish the circumstances under which the N450 may reflect stimulus- versus response-interference.

In sum, the findings in this thesis suggested unimodal and cross-modal interference occur via different mechanisms. Previous studies had reported that cross-modal interference may be maintained via sensory gating, and therefore be less susceptible to ageing. The findings in this thesis support the theory that cross-modal interference can be suppressed prior to the stage of response selection. I speculate that the ACC acts to detect sensory conflict, which then triggers thalamic gating. However, further research is needed to investigate this hypothesis.

Does a shift in sensory dominance occur across development, and is sensory dominance related to cross-modal interference?

A speculation I made at the start of this thesis was that sensory dominance might influence cross-modal distractibility. This speculation was based on findings in the Stroop literature suggesting visual and auditory information were treated differently in different age groups. Children appeared more susceptible to auditory distraction when focusing on vision (Hanauer & Brooks, 2003, 2005), yet vision did not interfere with audition in younger children (Thomas et al., 2017). In contrast, young adults showed asymmetrical interference, such that vision influences audition more than vice versa (Donohue, Appelbaum, et al., 2013; Yuval-Greenberg & Deouell, 2009). Furthermore, older adults appeared able to ignore audition whilst focusing on vision (Guerreiro et al., 2010; Guerreiro & Van Gerven, 2011). In parallel to this, children appeared to show less visual dominance in the Colavita task (Nava & Pavani, 2013; Wille & Ebersbach, 2016) whilst adults (Colavita, 1974) and older adults (Diaconescu et al., 2013) appeared more visually dominant. Given these parallels, I speculated that sensory dominance and cross-modal distractibility might be related. However, before exploring this link, I aimed to test the robustness of the Colavita task as a measure of sensory dominance (via a meta-analysis) and assess whether a developmental shift in sensory dominance was present (using the McGurk effect).

In Chapter 7 the Colavita effect was found to be a robust measure of sensory dominance, and age group (child vs. adult) was the only factor to moderate this effect. Unfortunately, there had been no studies investigating the Colavita effect in older adults, therefore exploring a shift in sensory dominance was constrained to childhood development. Chapter 8 confirmed the hypothesised developmental shift in sensory dominance, as children became progressively more influenced by vision in the McGurk effect. Furthermore, this Chapter indicated that sensory dominance appeared associated with an ability to detect signal in noise within the dominant modality, rather than a generalised increased sensitivity to information in that modality.

Contrary to hypotheses, in Chapter 9 I found no correlation between sensory dominance (measured with the Colavita task) and cross-modal interference. However, unimodal interference was positively correlated with visual dominance. As highlighted in Chapter 9, this result indicates a need to investigate whether sensory dominance may be associated with experiences (such as reading) and if this therefore influences the type of information people are distracted by.

In sum, it appears that a developmental shift in sensory dominance from audition to vision does occur. This shift is consistent with the observation from previous cross-modal Stroop studies, showing stronger influence of audition in children and vision in adults. However, the current thesis did not find vision and audition affected children and adults differently in the cross-modal Stroop task. Furthermore, in Chapter 9, there was no correlation between cross-modal distractibility and sensory dominance. However, the correlation between unimodal interference and sensory dominance does highlight the need for further research to understand the origin of sensory dominance shifts and why this might be associated with unimodal Stroop interference.

Gaps in literature addressed by this thesis

The research in this thesis contributes substantially to an area severely under-represented in the literature. Prior to this thesis only 24 studies were identified using cross-modal Stroop tasks. The majority of studies had investigated the effect of auditory distractors on visual attention, only five studies had investigated the effects of vision on audition (Appelbaum et al., 2013; Donohue, Appelbaum, et al., 2013; Donohue, Todisco, et al., 2013; Thomas et al., 2017; Yuval-Greenberg & Deouell, 2009), and, of these, only four had investigated the effect of vision on audition *and* vice versa (Appelbaum et al., 2013; Donohue, Appelbaum, et al., 2013; Donohue, Todisco, et al., 2013; Yuval-Greenberg & Deouell, 2009). Only one study had used what I describe to be a "fair" unimodal comparison task (Francis et al., 2017). Finally, only three studies had examined cross-modal effects in children (Hanauer & Brooks, 2003, 2005; Thomas et al., 2017) and none had studied effects in older adults. As I have already discussed findings relating to development and ageing, I will briefly outline the contributions of this thesis to ignoring vision and vice versa, and integrated vs. separate Stroop tasks.

Focusing on vision whilst ignoring audition and vice versa

A recent review by Van Gerven and Guerreiro (2016) highlighted the need for "fully-crossed" paradigms to compare unimodal and cross-modal attention. Figure 52 illustrates the fully crossed design, in which auditory and visual information can be either relevant or irrelevant and unimodal or cross-modal. This approach is important because the goal of studying cross-modal distraction is to understand how interference occurs in real-life, and, in real-life, distraction can occur within and across multiple senses.

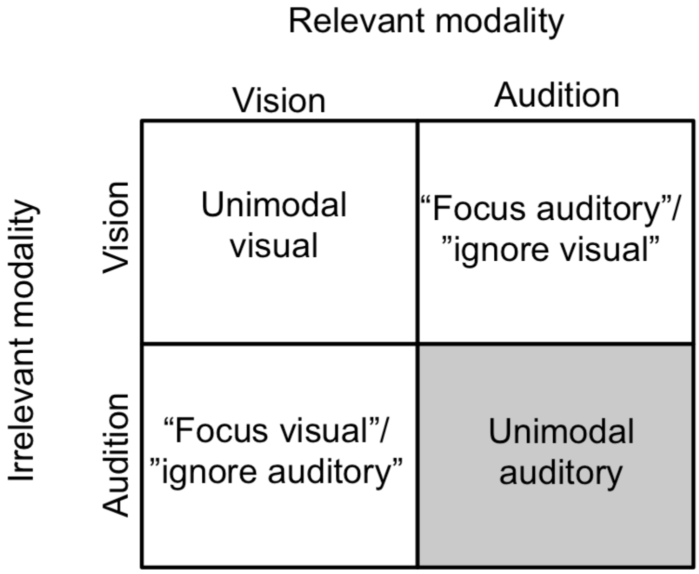


Figure 52. Figure adapted from Van Gerven and Guerreiro (2016) showing the "fully crossed" design. White boxes were studied in this thesis, the grey box was not.

In this thesis I included experiments investigating three of the four possible combinations of the fully-crossed design. Findings suggested that the differences in lifespan trajectories between unimodal and cross-modal interference occur whether the distractor is auditory or visual. However, I did not examine unimodal auditory change. Furthermore, it is possible that the "ignore visual" condition was easier than the "ignore auditory" condition. The "ignore visual" condition of Chapter 3, Experiment 2, involved ignoring visually presented colour and focusing on a spoken word, whilst the "ignore auditory" condition involved focusing on visually presented colour and ignoring a spoken word. The *word production architecture account* of Stroop interference proposes that Stroop interference occurs because colour naming is more demanding (and therefore slower) than word reading, requiring an additional stage to identify the associated word (Dyer, 1973; Elliott et al., 2014; Roelofs, 2005). As such, smaller interference effects would be expected in the "ignore visual" condition, in which the relevant stimulus (the spoken word) was more easily accessible. Therefore, the similarities between cross-modal conditions in Chapter 3 might be attributed to ease of access. To test this, future research could employ cross-modal paradigms in which auditory and visual information are more balanced in terms of accessibility. If the current findings can be attributed to ease of access, it would be expected that visual distractors would have a larger effect on audition than vice versa.

In sum, although the current research partially addresses the lack of studying the effect of vision on audition and vice versa, several gaps remain. Specifically, future research should aim to use fully-crossed paradigms, including a unimodal auditory condition, and it should be ensured that visual and auditory distractor conditions are balanced in terms of accessibility.

Integrated vs. separate Stroop tasks

Prior to this thesis, most studies comparing unimodal and cross-modal Stroop interference had not used a “fair” unimodal comparison task. This is because the unimodal task in most studies was the traditional colour-word Stroop task, in which colour and word were integrated into the same object. Conversely, in cross-modal Stroop tasks colour and word appeared as separate objects. Elliott et al (2014) propose that the integrated vs. separate nature of unimodal vs. cross-modal tasks respectively, might account for why cross-modal Stroop effects have appeared smaller compared with unimodal effects, as it is well established separated Stroop tasks result in smaller Stroop effects (Francis et al., 2017; Kahneman & Chajczyk, 1983; Macleod & Hodder, 1998).

Although the current thesis did not directly compare integrated vs. separate Stroop tasks, the Stroop paradigm used throughout employed a fairer unimodal comparison task in which word and colour could still be considered separate. A core observation throughout this thesis was that, overall, unimodal and cross-modal interference effects did not significantly differ. In Chapter 3, unimodal interference was significantly higher than cross-modal interference, but this was due to an effect seen in older adults that did not occur in young adults. In Chapter 5, with young adults, there was no significant difference in overall interference between unimodal and cross-modal conditions. Similarly in Chapter 9, there was no significant difference in interference between unimodal and cross-modal tasks. This finding supports the conclusion that overall differences in magnitude between unimodal and cross-modal Stroop effects may result from differences in integrated vs. separated tasks (Elliott et al., 2014).

One assumption that I made in this design, was that presenting auditory information over headphones would result in participants localising sound to the point of fixation (Stern et al., 2006). This approach was practical because research with children was to be conducted at a public engagement event in which testing took place in a room alongside other studies. However, this could be better controlled through presenting auditory information via a speaker at the same location as the visual stimulus. Furthermore, this approach would allow future research to directly compare integrated vs. separate unimodal and cross-modal tasks through moving visual and auditory distractors set distances away from the relevant colour information.

Limitations and future directions

There are several limitations of the studies contained within this thesis that should be considered. The specific limitations relating to each experiment are discussed within the relevant experimental chapters. Therefore, I will discuss the overarching limitations of this thesis and potential future directions.

First, although I aimed to understand unimodal and cross-modal interference in development and ageing, child and older adult populations were only included in Chapter 3, and the following studies sought to test hypotheses with young adults. For example, Chapter 4 investigated whether limiting colour vision could “simulate” ageing, and Chapters 5 and 6 were conducted with young adults to test hypotheses relating to the processing stages of unimodal vs. cross-modal control. Although these studies were critical in following up the hypotheses developed in Chapter 3, it cannot be firmly concluded that the effects observed in young adults in these chapters also transfer to children and older adults. Future work should therefore be conducted to assess whether these findings apply to young and old age groups. The work conducted in Chapter 6 lays the ground-work for studying ERP effects in child and adult populations using more convenient portable EEG. This work should identify whether children and older adults show the unimodal and cross-modal interference at the hypothesised processing stages. Specifically, it would be expected that children should show lower N450 sensitivity to cross-modal interference whilst this effect would be larger in older adults (signalling increased detection of sensory level conflict).

A second limitation of the current thesis was the restricted use of tasks. Although this thesis used tasks that are under-represented (i.e. the cross-modal Stroop), it is likely that most cross-modal interference paradigms are under-represented relative to their unimodal counterparts, in the literature. Furthermore, it is possible that the findings derived from these studies do not generalise to all types of distraction, including many real world examples of distraction. Given that the current findings are in line with the conclusions by Guerreiro et al (2010), whose extensive review encompassed a wide range of tasks, it would be hypothesised that the currently reported findings should generalise to other tasks. Nevertheless, future research should assess this. Furthermore, future work should investigate in what circumstances older adults can focus on audition and ignore vision and examine whether the current effects generalise to real-world distraction.

The issue of limited tasks is also the case for studies of sensory dominance in this thesis. I focused on the Colavita task and the McGurk effect to measure changes in sensory dominance. However, there are many other approaches to studying this. For example, the ventriloquist effect (Thurlow & Jack, 1973), the sound-induced-flash-illusion (Shams et al., 2000) and the cross-bounce illusion (Sekuler et al., 1997). It is highly likely the current findings extend to these effects, however, this cannot be assumed as the neural mechanisms underlying these effects may differ. For example, Calvert (2001) highlighted that the cortical site of integration differs depending upon the nature of the task. As such, different tasks may be differentially effected by development and ageing. Future research should examine whether shifts in sensory dominance extend to other illusionary effects.

A third limitation of this thesis is that I only studied shifts in sensory dominance in children, yet the goal of this research was to make inferences regarding lifespan change. Although this work was not conducted in older adults, the current findings certainly provide preliminary support for future work investigating sensory dominance in older adults. Furthermore, as older adults are a heterogeneous population, it should be assessed whether individual differences in sensory dominance exist in ageing and the factors associated with these (for example, age-related sensory loss).

Finally, the main approach used to infer neural processes in this thesis was EEG. This was because specific hypotheses were developed regarding where in the processing stream interference occurred. Thus, an approach with temporal precision was selected. However, additional information would be gained from complementary research methods with spatial specificity. This may help in answering the question as to whether cross-modal interference can be achieved through non-frontal mechanisms and whether the ACC is responsible for triggering gating mechanisms in cross-modal control.

Conclusions

In sum, this thesis suggests different mechanisms underpin unimodal and cross-modal control. Unimodal interference control was lower in development and aging, whilst cross-modal control was susceptible to developmental maturation but spared in ageing. Unimodal interference occurred due to stimulus and response-interference, whilst cross-modal interference occurred due to stimulus- and not response-interference. Following this, a developmental shift in sensory dominance from auditory towards visual dominance was observed. However, individual differences in visual dominance were associated with unimodal and not cross-modal control. I hypothesise that cross-modal interference may be achieved via cortico-thalamic projections, by which the ACC detects sensory conflict and triggers thalamic gating mechanisms. However, future research is needed to assess this hypothesis.

References

Adleman, N. (2002). A developmental fMRI study of the Stroop color-word task. *NeuroImage*, *16*(1), 61–75. http://doi.org/10.1006/nimg.2001.1046

Algina, J., & Keselman, H. J. (2003). Approximate confidence intervals for effect sizes. *Educational and Psychological Measurement*, *63*(4), 537–553. http://doi.org/10.1177/0013164403256358

Allard, R., & Faubert, J. (2008). The noisy-bit method for digital displays: Converting a 256 luminance resolution into a continuous resolution. *Behavior Research Methods*, *40*(3), 735–743. http://doi.org/10.3758/BRM.40.3.735

Allen, H. A., & Payne, H. (2012). Similar behaviour, different brain patterns: Age-related changes in neural signatures of ignoring. *NeuroImage*, *59*, 4113–4125. http://doi.org/10.1016/j.neuroimage.2011.10.070

Anstey, K. J., Dain, S., Andrews, S., & Drobny, J. (2002). Visual abilities in older adults explain age-differences in Stroop and fluid intelligence but not face recognition: Implications for the vision-cognition connection. *Aging, Neuropsychology, and Cognition*, *9*(4), 253–265. http://doi.org/10.1076/anec.9.4.253.8770

Appelbaum, L. G., Donohue, S. E., Park, C. J., & Woldorff, M. G. (2013). Is one enough? The case for non-additive influences of visual features on crossmodal Stroop interference. *Frontiers in Psychology*, *4*, 799. http://doi.org/10.3389/fpsyg.2013.00799

Appelbaum, L. G., Meyerhoff, K. L., & Woldorff, M. G. (2009). Priming and backward influences in the human brain: Processing interactions during the stroop interference effect. *Cerebral Cortex*, *19*(11), 2508–2521. http://doi.org/10.1093/cercor/bhp036

Ashitaka, Y., & Shimada, H. (2010). Stroop effect with synonymous homographic heterophonemic words in Japanese Kanji. In

Proceedings of the 26th Annual Meetings of the International Society for Psychophysics.

Aspinall, P., Mavros, P., Coyne, R., & Roe, J. (2015). The urban brain: Analysing outdoor physical activity with mobile EEG. *British Journal of Sports Medicine*, *49*(4), 272–276. http://doi.org/10.1136/bjsports-2012-091877

Augustinova, M., Silvert, L., Ferrand, L., Llorca, P. M., & Flaudias, V. (2014). Behavioral and electrophysiological investigation of semantic and response conflict in the Stroop task. *Psychonomic Bulletin & Review*, *Online*(August), 1–7. http://doi.org/10.3758/s13423-014-0697-z

Badcock, N. A., Mousikou, P., Mahajan, Y., de Lissa, P., Thie, J., & McArthur, G. (2013). Validation of the Emotiv EPOC(®) EEG gaming system for measuring research quality auditory ERPs. *PeerJ*, *1*(1), e38. http://doi.org/10.7717/peerj.38

Badcock, N. A., Preece, K. A., de Wit, B., Glenn, K., Fieder, N., Thie, J., & McArthur, G. (2015). Validation of the Emotiv EPOC EEG system for research quality auditory event-related potentials in children. *PeerJ*, *3*, e907. http://doi.org/10.7717/peerj.907

Badcock, N. A., & Thei, J. (2015). matEPOC: Toolbox for handling Emotiv EPOC data and insertion of physical event markers in MATLAB.

Baier, B., Kleinschmidt, A., & Müller, N. G. (2006). Cross-modal processing in early visual and auditory cortices depends on expected statistical relationship of multisensory information. *Journal of Neuroscience*, *26*(47), 12260–12265. http://doi.org/10.1523/JNEUROSCI.1457-06.2006

Banich, M. T., Milham, M. P., Atchley, R. A., Cohen, N. J., Webb, A., Wszalek, T., … Brown, C. (2000). Prefrontal regions play a predominant role in imposing an attentional “set”: evidence from fMRI. *Cognitive Brain Research*, *10*(1–2), 1–9. http://doi.org/10.1016/S0926-6410(00)00015-X

Barham, M. P., Clark, G. M., Hayden, M. J., Enticott, P. G., Conduit, R., & Lum, J. A. G. (2017). Acquiring research-grade ERPs on a shoestring budget: A comparison of a modified Emotiv and commercial SynAmps EEG system. *Psychophysiology*, *54*(9), 1393–1404. http://doi.org/10.1111/psyp.12888

Barnea-Goraly, N., Menon, V., Eckert, M., Tamm, L., Bammer, R., Karchemskiy, A., … Reiss, A. L. (2005). White matter development during childhood and adolescence: A cross-sectional diffusion tensor imaging study. *Cerebral Cortex*, *15*(12), 1848–1854. http://doi.org/10.1093/cercor/bhi062

Barnhart, W. R., Rivera, S., & Robinson, C. W. (2018). Different patterns of modality dominance across development. *Acta Psychologica*, *182*, 154–165. http://doi.org/10.1016/j.actpsy.2017.11.017

Barutchu, A., Crewther, D. P., & Crewther, S. G. (2009). The race that precedes coactivation: development of multisensory facilitation in children. *Developmental Science*, *12*(3), 464–73. http://doi.org/10.1111/j.1467-7687.2008.00782.x

Barutchu, A., Danaher, J., Crewther, S. G., Innes-Brown, H., Shivdasani, M. N., & Paolini, A. G. (2010). Audiovisual integration in noise by children and adults. *Journal of Experimental Child Psychology*, *105*(1–2), 38–50. http://doi.org/10.1016/j.jecp.2009.08.005

Bedard, G., & Barnett-Cohen, M. (2016). Impaired timing of audiovisual events in the elderly. *Experimental Brain Research*, *234*(1), 331–340. http://doi.org/10.1007/s00221-015-4466-7

Bell, R., & Buchner, A. (2007). Equivalent irrelevant-sound effects for old and young adults. *Memory & Cognition*, *35*(2), 352–364. http://doi.org/10.3758/BF03193456

Belleville, S., Rouleau, N., Van Der Linden, M., & Collette, F. (2003). Effect of manipulation and irrelevant noise on working memory capacity of patients with Alzhiemers dementia. *Neuropsychology*, *17*, 69–81.

Ben-David, B. M., & Schneider, B. A. (2009). A sensory origin for color-word Stroop effects in aging: A meta-analysis. *Aging, Neuropsychology, and Cognition*, *16*(5), 505–534. http://doi.org/10.1080/13825580902855862

Ben-David, B. M., & Schneider, B. A. (2010). A sensory origin for color-word Stroop effects in aging: Simulating age-related changes in color-vision mimics age-related changes in Stroop. *Aging, Neuropsychology, and Cognition*, *17*(6), 730–746. http://doi.org/10.1080/13825585.2010.510553

Bishop, D. V. M., Hardiman, M., Uwer, R., & Von Suchodoletz, W. (2007). Maturation of the long-latency auditory ERP: Step function changes at start and end of adolescence: REPORT. *Developmental Science*, *10*(5), 565–575. http://doi.org/10.1111/j.1467-7687.2007.00619.x

Blasi, G., Goldberg, T. E., Weickert, T., Das, S., Kohn, P., Zoltick, B., … Mattay, V. S. (2006). Brain regions underlying response inhibition and interference monitoring and suppression. *The European Journal of Neuroscience*, *23*(6), 1658–64. http://doi.org/10.1111/j.1460-9568.2006.04680.x

Boersma, P., & Weenink, D. (2013). Praat: doing phonetics by computer. Retrieved from http://www.praat.org/

Borenstein, M., Hedges, L. V, Higgins, J. P. T., & Rothstein, H. R. (2009). Effect sizes based on means. In *Introduction to Meta-Analysis* (pp. 21–32). Chichester, UK: John Wiley & Sons, Ltd. http://doi.org/10.1002/9780470743386.ch6

Botvinick, M. ., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and the anterior cingulate cortex: an update. *Trends in Cognitive Sciences*, *8*(12), 539–546.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652. http://doi.org//10.1037/0033-

295X.108.3.624

Braga, R. M., Wilson, L. R., Sharp, D. J., Wise, R. J. S., & Leech, R. (2013). Separable networks for top-down attention to auditory non-spatial and visuospatial modalities. *NeuroImage*, *74*, 77–86. http://doi.org/10.1016/j.neuroimage.2013.02.023

Brooks, C. J., Chan, Y. M., Anderson, A. J., & Mckendrick, A. M. (2018). Audiovisual temporal perception in aging : The role of multisensory integration and age-related sensory loss. *Frontiers in Human Neuroscience*, *12*(192). http://doi.org/10.3389/fnhum.2018.00192

Burr, D., & Gori, M. (2012). Multisensory integration develops late in humans. In M. Murray & M. T. Wallace (Eds.), *The Neural Bases of Multisensory Processes* (pp. 345–363). Boca Raton (FL): CRC Press/Taylor & Francis. Retrieved from //www.ncbi.nlm.nih.gov/books/NBK92864/

Calvert, G. A. (2001). Crossmodal processing in the human brain: Insights from functional neuroimaging studies. *Cerebral Cortex*, *11*(12), 1110–1123. http://doi.org/10.1093/cercor/11.12.1110

Campbell, J., & Sharma, A. (2014). Cross-modal re-organization in adults with early stage hearing loss. *PLoS ONE*, *9*(2). http://doi.org/10.1371/journal.pone.0090594

Carriere, B. N., Royal, D. W., Perrault, T. J., Morrison, S. P., Vaughan, J. W., Stein, B. E., & Wallace, M. T. (2007). Visual deprivation alters the development of cortical multisensory integration. *Journal of Neurophysiology*, *98*(5), 2858–2867. http://doi.org/10.1152/jn.00587.2007.

Čeponiene, R., Westerfield, M., Torki, M., & Townsend, J. (2008). Modality-specificity of sensory aging in vision and audition: Evidence from event-related potentials. *Brain Research*, *1215*, 53–68. http://doi.org/10.1016/j.brainres.2008.02.010

Chan, Y. M., Pianta, M. J., & Mckendrick, A. M. (2014). Older age results in difficulties separating auditory and visual signals in time. *Journal of Vision*, *14*(11), 1–11. http://doi.org/10.1167/14.11.13

Chen, A., Bailey, K., Tiernan, B. N., & West, R. (2011). Neural correlates of stimulus and response interference in a 2-1 mapping stroop task. *International Journal of Psychophysiology*, *80*, 129–38. http://doi.org/10.1016/j.ijpsycho.2011.02.012

Christensen, T. A., Lockwood, J. L., Almryde, K. R., & Plante, E. (2011). Neural substrates of attentive listening assessed with a novel auditory Stroop task. *Frontiers in Human Neuroscience*, *4*, 236. http://doi.org/10.3389/fnhum.2010.00236

Clay Smith, H. (1943). Age differences in color discrimination. *Journal of General Psychology*, *29*, 191–226.

Coderre, E., Conklin, K., & Van Heuven, W. J. B. (2011). Electrophysiological measures of conflict detection and resolution in the Stroop task. *Brain Research*, *1413*, 51–59. http://doi.org/10.1016/j.brainres.2011.07.017

Cohn, N. B., Dustman, R. E., & Bradford, D. C. (1984). Age-related decrements in stroop color test performance. *Journal of Clinical Psychology*, *40*(5), 1244–1250. http://doi.org/10.1002/1097-4679(198409)40:5<1244::AID-JCLP2270400521>3.0.CO;2-D

Colavita, F. B. (1974). Human sensory dominance. *Perception & Psychophysics*, *16*(2), 409–412. http://doi.org/10.3758/BF03203962

Colavita, F. B. (1982). Visual dominance and attention in space. *Bulletin of the Psychonomic Society*, *19*(5), 261–262. http://doi.org/10.3758/BF03330251

Colavita, F. B., Tomko, R., & Weisberg, D. (1976). Visual prepotency and eye orientation. *Society*, *8*(l), 25–26. http://doi.org/10.3758/BF03337062

Colavita, F. B., & Weisberg, D. (1979). A further investigation of visual dominance. *Perception & Psychophysics*, *25*(4), 345–347. http://doi.org/10.3758/BF03198814

Coles, M. G. H. (1989). Modern mind-brain reading: psychophysiology, physiology, and cognition. *Psychophysiology*, *26*(3), 251–269.

Comalli, P. E., Wapner, S., & Werner, H. (1962). Interference effecs of Stroop colour-word test in childhood adulthood and aging. *The Journal of Genetic Psychology*, *100*, 47–53. http://doi.org//10.1080/00221325.1962.10533572

Cooper, B. A., Ward, M., Gowland, C. A., & McIntosh, J. M. (1991). The use of the Lanthony New Color Test in determining the effects of aging on color vision. *Journal of Gerontology: Psychological Sciences*, *46*(6), 320–324. http://doi.org//10.1093/geronj/46.6.P320

Cornsweet, T. M. (1962). The staircase-method in psychophysics. *The American Journal of Psychology*, *75*(3), 485–491.

Cosman, J. D., Lowe, K. A., Woodman, G. F., & Schall, J. D. (2018). Prefrontal control of visual distraction. *Current Biology*, *28*(3), 414–420.e3. http://doi.org/10.1016/j.cub.2017.12.023

Cowan, N. (1989a). A reply to Miles, Madden and Jones (1989): Mistakes and other flaws in the challenge to the cross-modal Stroop effect. *Perception & Psychophysics*, *45*(I), 82–84.

Cowan, N. (1989b). The reality of cross-modal Stroop effects. *Perception & Psychophysics*, *45*(1), 87–88. http://doi.org/10.3758/BF03208039

Cowan, N., & Barron, A. (1987). Cross-modal, auditory-visual Stroop interference and possible implications for speech memory. *Perception & Psychophysics*, *41*(5), 393–401. http://doi.org/10.3758/BF03203031

Cragg, L. (2016). The development of stimulus and response interference control in midchildhood. *Developmental Psychology*, *52*(2), 242–252. http://doi.org//10.1037/dev0000074

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta- analysis.* New York: Routedge.

Cumming, G., & Calin-Jagerman, R. (2017). *Introduction to the new statistics: Estimation, open science and beyond*. New York: Routedge.

Davidson, D. J., Zacks, R. T., & Williams, C. . (2003). Stroop interference, practice and aging. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *10*(2), 85–98. http://doi.org/10.1111/j.1467-9876.2009.00682.x.Disaggregated

De Houwer, J. (2003). On the role of stimulus-response and stimulus-stimulus compatibility in the Stroop effect. *Memory & Cognition*, *31*(3), 353–359. http://doi.org/10.3758/BF03194393

de Lissa, P., Sorensen, S., Badcock, N., Thie, J., & McArthur, G. (2015). Measuring the face-sensitive N170 with a gaming EEG system: A validation study. *Journal of Neuroscience Methods*, *253*(August), 47–54. http://doi.org/10.1016/j.jneumeth.2015.05.025

De Vos, M., Gandras, K., & Debener, S. (2014). Towards a truly mobile auditory brain-computer interface: Exploring the P300 to take away. *International Journal of Psychophysiology*, *91*(1), 46–53. http://doi.org/10.1016/j.ijpsycho.2013.08.010

Debener, S., Minow, F., Emkes, R., Gandras, K., & De Vos, M. (2012). How about taking a low-cost, small, and wireless EEG for a walk? *Psychophysiology*, *49*(11), 1617–1621. http://doi.org/10.1111/j.1469-8986.2012.01471.x

Delano, P. H., Elgueda, D., Hamame, C. M., & Robles, L. (2007). Selective attention to visual stimuli reduces cochlear sensitivity in chinchillas. *Journal of Neuroscience*, *27*(15), 4146–4153. http://doi.org/10.1523/JNEUROSCI.3702-06.2007

Dempster, F. N., & Vegas, L. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review*, *12*(1), 45–75. http://doi.org/10.1016/0273-2297(92)90003-K

Diaconescu, A. O., Alain, C., & McIntosh, A. R. (2011). The co-occurrence of multisensory facilitation and cross-modal conflict in the human brain. *Journal of Neurophysiology*, *106*(6), 2896–2909. http://doi.org/10.1152/jn.00303.2011

Diaconescu, A. O., Hasher, L., & McIntosh, A. R. (2013). Visual dominance and multisensory integration changes with age. *NeuroImage*, *65*, 152–66. http://doi.org/10.1016/j.neuroimage.2012.09.057

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168. http://doi.org/10.1146/annurev-psych-113011-143750

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. http://doi.org/10.3758/BF03195006

Diederich, A., Colonius, H., & Schomburg, A. (2008). Assessing age-related multisensory enhancement with the time-window-of-integration model. *Neuropsychologia*, *46*(10), 2556–2562. http://doi.org/10.1016/j.neuropsychologia.2008.03.026

Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., … Poeppel, D. (2017). Brain-to-Brain Synchrony Tracks Real-World Dynamic Group Interactions in the Classroom. *Current Biology*, *27*(9), 1375–1380. http://doi.org/10.1016/j.cub.2017.04.002

Donohue, S. E., Appelbaum, L. G., Mckay, C. C., & Woldorff, M. G. (2016). The neural dynamics of stimulus and response conflict processing as a function of response complexity and task demands. *Neuropsychologia*, *84*, 14–28. http://doi.org/10.1016/j.neuropsychologia.2016.01.035.The

Donohue, S. E., Appelbaum, L. G., Park, C. J., Roberts, K. C., & Woldorff, M. G. (2013). Cross-modal stimulus conflict: The behavioral effects of stimulus input timing in a visual-auditory Stroop task. *PLoS ONE*, *8*(4). http://doi.org/10.1371/journal.pone.0062802

Donohue, S. E., Todisco, A. E., & Woldorff, M. G. (2013). The rapid distraction of attentional resources toward the source of incongruent stimulus input during multisensory conflict. *Journal of Cognitive Neuroscience*, *25*(4), 623–635.

http://doi.org/10.1162/jocn_a_00336

Dryden, A., Allen, H. A., Henshaw, H., & Heinrich, A. (2017). The association between cognitive performance and speech-in-noise perception for adult listeners: A systematic literature review and meta-analysis. *Trends in Hearing*, *21*, 1–21. http://doi.org/10.1177/2331216517744675

Dunlap, W. P., Jose, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, *1*(2), 170–177. http://doi.org/10.1037/1082-989X.1.2.170

Duvinage, M., Castermans, T., Petieau, M., Hoellinger, T., Cheron, G., & Dutoit, T. (2013). Performance of the Emotiv Epoc headset for P300-based applications. *BioMedical Engineering Online*, *12*(1), 1–15. http://doi.org/10.1186/1475-925X-12-56

Dyer, F. N. (1973). The Stroop phenomenon and its use in the stlldy of perceptual, cognitive, and response processes. *Memory & Cognition*, *1*(2), 106–120. http://doi.org/10.3758/BF03198078

Egeth, H. E., & Sager, L. C. (1977). On the locus of visual dominance. *Perception & Psychophysics*, *22*(1), 77–86. http://doi.org/10.3758/BF03206083

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, *315*(7109), 629–634. http://doi.org/10.1136/bmj.316.7129.469

Elliott, E. M., Barrilleaux, K., & Cowan, N. (2006). Individual differences in the ability to avoid distracting sounds. *European Journal of Cognitive Psychology*, *18*(1), 90–108. http://doi.org/10.1080/09541440500216044

Elliott, E. M., & Cowan, N. (2001). Habituation to auditory distractors in a cross-modal, color-word interference task. *Journal of Experimental Psychology*, *27*(3), 654–667.

Elliott, E. M., Cowan, N., & Valle-Inclan, F. (1998). The nature of cross-modal color-word interference effects. *Perception & Psychophysics*, *60*(5), 761–767. http://doi.org/10.3758/BF03206061

Elliott, E. M., Morey, C. C., Morey, R. D., Eaves, S. D., Shelton, J. T., & Lutfi-Proctor, D. A. (2014). The role of modality: Auditory and visual distractors in Stroop interference. *Journal of Cognitive Psychology*, *26*(1), 15–26. http://doi.org/10.1080/20445911.2013.859133

Ernst, M. O. (2008). Multisensory integration: A late bloomer. *Current Biology*, *18*(12), 519–521. http://doi.org//10.1016/j.cub.2008.05.002

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, *8*(4), 162–169. http://doi.org/10.1016/j.tics.2004.02.002

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–91. http://doi.org/10.3758/BF03193146

Fetsch, C. R., Deangelis, G. C., & Angelaki, D. E. (2013). Bridging the gap between theories of sensory cue integration and the physiology of multisensory neurons. *Nature Reviews Neuroscience*, *14*(6), 429–442. http://doi.org/10.1038/nrn3503

Fixmer, E., & Hawkins, S. (1998). The influence of quality of information on the McGurk effect. *Proceedings of the International Conference on Auditory-Visual Speech Processing*, 27–32.

Floresco, S. B., & Grace, A. a. (2003). Gating of hippocampal-evoked activity in prefrontal cortical neurons by inputs from the mediodorsal thalamus and ventral tegmental area. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *23*(9), 3930–3943. http://doi.org/23/9/3930 [pii]

Flowers, J. H., & Stoup, C. M. (1977). Selective attention between words, shapes and colors in speeded classification and vocalization tasks. *Memory & Cognition*, *5*(3), 299–307. http://doi.org/10.3758/BF03197574

Forster, B., Cavina-Pratesi, C., Aglioti, S. M., & Berlucchi, G. (2002). Redundant target effect and intersensory facilitation from visual-tactile interactions in simple reaction time. *Experimental Brain Research*, *143*(4), 480–487. http://doi.org/10.1007/s00221-002-1017-9

Francis, W. S., MacLeod, C. M., & Taylor, R. S. (2017). Joint influence of visual and auditory words in the Stroop task. *Attention, Perception, and Psychophysics*, *79*(1), 200–211. http://doi.org/10.3758/s13414-016-1218-0

Frassinetti, F., Bolognini, N., & Làdavas, E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Experimental Brain Research*, *147*(3), 332–43. http://doi.org/10.1007/s00221-002-1262-y

Fritz, J. B., Elhilali, M., David, S. V., & Shamma, S. A. (2007). Auditory attention - focusing the searchlight on sound. *Current Opinion in Neurobiology*, *17*(4), 437–455. http://doi.org/10.1016/j.conb.2007.07.011

Fuster, J. M. (2002). Frontal lobe and cognitive development. *Journal of Neurocytology*, *31*, 373–385. http://doi.org//10.1023/A:1024190429920

Gagné, J. P., Besser, J., & Lemke, U. (2017). Behavioral assessment of listening effort using a dual-task paradigm: A review. *Trends in Hearing*, *21*, 1–25. http://doi.org/10.1177/2331216516687287

Getzmann, S., Wascher, E., & Falkenstein, M. (2015). What does successful speech-in-noise perception in aging depend on ? Electrophysiological correlates of high and low performance in older adults. *Neuropsychologia*, *70*, 43–57. http://doi.org/10.1016/j.neuropsychologia.2015.02.009

Ghatan, P. H., Hsieh, J. C., Petersson, K. M., Stone-Elander, S., & Ingvar, M. (1998). Coexistence of attention-based facilitation and inhibition in the human cortex. *NeuroImage*, *7*(1), 23–29. http://doi.org/10.1006/nimg.1997.0307

Giard, M., Collet, L., Bouchet, P., & Pernier, J. (1994). Auditory selective attention in the human cochlea, *633*, 353–356.

Gisiger, T., & Boukadoum, M. (2011). Mechanisms gating the flow of information in the cortex: What they might look like and what their uses may be. *Frontiers in Computational Neuroscience*, *5*, 1. http://doi.org/10.3389/fncom.2011.00001

Gmehlin, D., Kreisel, S. H., Bachmann, S., Weisbrod, M., & Thomas, C. (2011). Age effects on preattentive and early attentive auditory processing of redundant stimuli: Is sensory gating affected by physiological aging? *Journals of Gerontology - Series A Biological Sciences and Medical Sciences*, *66 A*(10), 1043–1053. http://doi.org/10.1093/gerona/glr067

Golden, C. J. (1974). Effect of differing number of colours on the Stroop color and word test. *Perceptual and Motor Skills*, *39*(1, Pt 2), 550.

Gondan, M., Niederhaus, B., Rösler, F., & Röder, B. (2005). Multisensory processing in the redundant-target effect: a behavioral and event-related potential study. *Perception & Psychophysics*, *67*(4), 713–726. http://doi.org//10.3758/BF03193527

Gratton, G., Coles, M. G. H., & Donchin, E. (1992). Optimizing the use of information: stategic control of activation of responses. *Journal of Experimental Psychology*, *121*(4), 480–506.

Graven, S. N., & Browne, J. V. (2008a). Auditory development in the fetus and infant. *Newborn and Infant Nursing Reviews*, *8*(4), 187–193. http://doi.org/10.1053/j.nainr.2008.10.010

Graven, S. N., & Browne, J. V. (2008b). Visual development in the human fetus, infant, and young child. *Newborn and Infant Nursing Reviews*, *8*(4), 194–201. http://doi.org//10.1053/j.nainr.2008.10.011

Green, E. J., & Barber, P. J. (1981). An auditory Stroop effect with judgements of speaker gender. *Perception & Psychophysics*, *30*(5), 459–466. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7329763

Greenwood, P. M. (2000). The frontal aging hypothesis evaluated. *Journal of the International Neuropsychological Society : JINS*, *6*(6), 705–726. http://doi.org/10.1017/S1355617700666092

Gregg, M. K., & Purdy, K. A. (2007). Graded auditory Stroop effects generated by gender words. *Percept Mot Skills*, *105*(2), 549–555. http://doi.org/10.2466/pms.105.2.549-555

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. http://doi.org/10.1111/j.1469-8986.2011.01273.x

Grummett, T. S., Leibbrandt, R. E., Lewis, T. W., DeLosAngeles, D., Powers, D. M. W., Willoughby, J. O., … Fitzgibbon, S. P. (2015). Measurement of neural signals from inexpensive, wireless and dry EEG systems. *Physiological Measurement*, *36*(7), 1469–1484. http://doi.org/10.1088/0967-3334/36/7/1469

Guerreiro, M. J. S., Adam, J. J., & Van Gerven, P. W. M. (2012). Automatic selective attention as a function of sensory modality in aging. *The Journals of Gerontology. Series B: Psychological Sciences and Social Sciences*, *67*(2), 194–202. http://doi.org/10.1093/geronb/gbr090

Guerreiro, M. J. S., Adam, J. J., & Van Gerven, P. W. M. (2014). Aging and response interference across sensory modalities. *Psychonomic Bulletin & Review*, *21*, 836–842. http://doi.org/10.3758/s13423-013-0554-5

Guerreiro, M. J. S., Anguera, J. A., Mishra, J., Van Gerven, P. W. M., & Gazzaley, A. (2014). Age-equivalent top-down modulation during cross-modal selective attention. *Journal of Cognitive Neuroscience*, *26*(12), 2827–2839. http://doi.org//10.1162/jocn_a_00685

Guerreiro, M. J. S., Eck, J., Moerel, M., Evers, E. A. T., & Van Gerven, P. W. M. (2015). Top-down modulation of visual and auditory cortical processing in aging. *Behavioural Brain Research*, *278*, 226–234. http://doi.org/10.1016/j.bbr.2014.09.049

Guerreiro, M. J. S., Murphy, D. R., & Van Gerven, P. W. M. (2010). The role of sensory modality in age-related distraction: a critical review and a renewed view. *Psychological Bulletin*, *136*(6), 975–1022. http://doi.org/10.1037/a0020731

Guerreiro, M. J. S., Murphy, D. R., & Van Gerven, P. W. M. (2013). Making sense of age-related distractibility: The critical role of sensory modality. *Acta Psychologica*, *142*(2), 184–194. http://doi.org/10.1016/j.actpsy.2012.11.007

Guerreiro, M. J. S., & Van Gerven, P. W. M. (2011). Now you see it, now you don't: evidence for age-dependent and age-independent cross-modal distraction. *Psychology and Aging*, *26*(2), 415–426. http://doi.org/10.1037/a0021507

Guerreiro, M. J. S., & Van Gerven, P. W. M. (2017). Disregarding hearing loss leads to overestimation of age-related cognitive decline. *Neurobiology of Aging*, *56*, 180–189. http://doi.org/10.1016/j.neurobiolaging.2017.05.001

Haegerstrom-Portnoy, G., Hewlett, S. E., & Barr, S. A. N. (1989). S cone loss with aging. In B. Drum & G. Verriest (Eds.), *Colour Vision Deficiencies IX: Proceedings of the ninth symposium of the International Research Group on Colour Vision Deficiencies, held at St. John's College, Annapolis, Maryland, U.S.A., 1--3 July 1987* (pp. 345–352). Dordrecht: Springer Netherlands. http://doi.org/10.1007/978-94-009-2695-0_40

Hairston, W. D. (2012). Accounting for timing drift and variability in contemporary electroencephologography (EEG) systems. *DTIC Document*, 1–22. Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA561715

Hanauer, J. B., & Brooks, P. J. (2003). Developmental change in the cross-modal Stroop effect. *Perception & Psychophysics*, *65*(3), 359–366. http://doi.org/10.3758/BF03194567

Hanauer, J. B., & Brooks, P. J. (2005). Contributions of response set and semantic relatedness to cross-modal Stroop-like picture-word interference in children and adults. *Journal of Experimental Child Psychology*, *90*(1), 21–47. http://doi.org/10.1016/j.jecp.2004.08.002

Hartcher-O'Brien, J., Levitan, C., & Spence, C. (2010). Extending visual dominance over touch for input off the body. *Brain Research*, *1362*, 48–55. http://doi.org/10.1016/j.brainres.2010.09.036

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, *108*(3), 356–388. http://doi.org//10.1037/0096-3445.108.3.356

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view BT - The psychology of learning and motivation. *The Psychology of Learning and Motivation*. http://doi.org/10.1016/s0079-7421(08)60041-9

Haxby, J. V., Horwitz, B., Ungerleider, L. G., Maisog, J. M., Pietrini, P., & Grady, C. L. (1994). The functional organization of human extrastriate cortex: a PET-rCBF study of selective attention to faces and locations. *The Journal of Neuroscience*, *14*(11), 6336–6353. http://doi.org/7965040

Hazan, V., & Li, E. (2008). The effect of auditory and visual degradation on audiovisual perception of native and non-native speakers. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, (January 2008), 1191–1194.

Hecht, D., & Reiner, M. (2009). Sensory dominance in combinations of audio, visual and haptic stimuli. *Experimental Brain Research*, *193*(2), 307–314. http://doi.org/10.1007/s00221-008-1626-z

Hennelly, M. L., Barbur, J. L., Edgar, D. F., & Woodward, E. G. (1998). The effect of age on the light scattering characteristics of the eye. *Ophthalmic and Physiological Optics*, *18*(2), 197–203. http://doi.org/10.1016/S0275-5408(97)00077-X

Hillyard, S. A., Hink, R. F., Schwent, V. L., Picton, T. W., Series, N., & Oct, N. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*(4108), 177–180.

Hirst, R. J., & Allen, H. A. (2018). An adapted noisy bit method for creating continuous luminance resolution for digital displays through altering image opacity. GitHub repository. Retrieved from https://github.com/RebeccaHirst/NoisyBit_RJHirst

Huang, S., Li, Y., Zhang, W., Zhang, B., Liu, X., Mo, L., & Chen, Q. (2015). Multisensory competition is modulated by sensory pathway interactions with fronto-sensorimotor and default-mode network regions. *Journal of Neuroscience*, *35*(24), 9064–9077. http://doi.org/10.1523/JNEUROSCI.3760-14.2015

Hubel, D. H., Henson, C. O., Rupert, A., & Galambos, R. (1959). “Attention” units in the auditory cortex. *Psychobiology*, *129*, 1279–180.

Hugenschmidt, C. E., Mozolic, J. L., Tan, H., Kraft, R. A., & Laurienti, P. J. (2009). Age-related increase in cross-sensory noise in resting and steady-state cerebral perfusion. *Brain Topography*, *21*(3–4), 241–251. http://doi.org/10.1007/s10548-009-0098-1

Hullett, C. R., & Levine, T. R. (2003). The overestimation of effect sizes from F values in meta-analysis: The cause and a solution. *Communication Monographs*, *70*(1), 52–67. http://doi.org/10.1080/715114664

Hunter, S. K., Gillow, S. J., & Ross, R. G. (2015). Stability of P50 auditory sensory gating during sleep from infancy to 4 years of age. *Brain and Cognition*, 4–9. http://doi.org/10.1016/j.bande.2014.12.004

Innes-Brown, H., Barutchu, A., Shivdasani, M. N., Crewther, D. P., Grayden, D. B., & Paolini, A. (2011). Susceptibility to the flash-beep illusion is increased in children compared to adults. *Developmental Science*, *14*(5), 1089–1099. http://doi.org/10.1111/j.1467-7687.2011.01059.x

Iwasaki, M., Kellinghaus, C., Alexopoulos, A. V., Burgess, R. C., Kumar, A. N., Han, Y. H., … Leigh, R. J. (2005). Effects of eyelid closure, blinks, and eye movements on the electroencephalogram. *Clinical Neurophysiology*, *116*(4), 878–885. http://doi.org/10.1016/j.clinph.2004.11.001

Jensen, A. R., & Rohwer, W. D. J. (1966). The Stroop colour-word test: A review. *Acta Psychologica*, *25*, 36–93.

Johnson, J. A., Strafella, A. P., & Zatorre, R. J. (2007). The role of the dorsolateral prefrontal cortex in bimodal divided attention: Two transcranial magnetic stimulation studies. *Journal of Cognitive Neuroscience*, *19*(6), 907–920. http://doi.org/10.1162/jocn.2007.19.6.907

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

Johnson, J. A., & Zatorre, R. J. (2006). Neural substrates for dividing and focusing attention between simultaneous auditory and visual events. *NeuroImage*, *31*(4), 1673–1681. http://doi.org/10.1016/j.neuroimage.2006.02.026

Johnson, T. L., & Shapiro, K. L. (1989). Attention to auditory and peripheral visual stimuli: Effects of arousal and predictability. *Acta Psychologica*, *72*(3), 233–245. http://doi.org/10.1016/0001-6918(89)90031-0

Jones, L. A., Hills, P. J., Dick, K. M., Jones, S. P., & Bright, P. (2016). Cognitive mechanisms associated with auditory sensory gating. *Brain and Cognition*, *102*, 33–45. http://doi.org/10.1016/j.bandc.2015.12.005

Jongen, E. M. M., & Jonkman, L. M. (2008). The developmental pattern of stimulus and response interference in a color-object Stroop task: an ERP study. *BMC Neuroscience*, *9*, 82. http://doi.org/10.1186/1471-2202-9-82

Kahneman, D., & Chajczyk, D. (1983). Tests of the automaticity of reading: Dilution of Stroop effects by color-irrelevant stimuli.

Journal of Experimental Psychology: Human Perception and Performance, *9*(4), 497–509. http://doi.org/10.1037/0096-1523.9.4.497

Kawashima, R., O'Sullivan, B. T., & Roland, P. E. (1995). Positron-emission tomography studies of cross-modality inhibition in selective attentional tasks: closing the "mind's eye". *Proceedings of the National Academy of Sciences*, *92*(13), 5969–5972. http://doi.org/10.1073/pnas.92.13.5969

Keil, J., Müller, N., Ihssen, N., & Weisz, N. (2012). On the variability of the McGurk effect: Audiovisual integration depends on prestimulus brain states. *Cerebral Cortex*, *22*(1), 221–231. http://doi.org/10.1093/cercor/bhr125

Killikelly, C., & Szűcs, D. (2010). The development of interference control: A pilot study using the manual colour word stroop paradigm. *Procedia - Social and Behavioral Sciences*, *2*, 4842–4847. http://doi.org/10.1016/j.sbspro.2010.03.781

Killikelly, C., & Szűcs, D. (2013). Asymmetry in stimulus and response conflict processing across the adult lifespan: ERP and EMG evidence. *Cortex*, *49*(10), 2888–903. http://doi.org/10.1016/j.cortex.2013.08.017

Kisley, M. A., Polk, S. D., Ross, R. G., Levisohn, P. M., & Freedman, R. (2003). Early postnatal development of sensory gating. *Neuroreport*, *14*(5), 693–697. http://doi.org/10.1097/01.wnr.0000064988.96259.b8 [doi]

Koppen, C., Alsius, A., & Spence, C. (2008). Semantic congruency and the Colavita visual dominance effect. *Experimental Brain Research*, *184*(4), 533–546. http://doi.org/10.1007/s00221-007-1120-z

Koppen, C., Levitan, C. A., & Spence, C. (2009). A signal detection study of the Colavita visual dominance effect. *Experimental Brain Research*, *196*(3), 353–360. http://doi.org/10.1007/s00221-009-1853-y

Koppen, C., & Spence, C. (2007a). Assessing the role of stimulus probability on the Colavita visual dominance effect. *Brain Research*, *1186*(1), 224–232. http://doi.org/10.1016/j.brainres.2007.09.076

Koppen, C., & Spence, C. (2007b). Audiovisual asynchrony modulates the Colavita visual dominance effect. *Brain Research*, *1186*(1), 224–232. http://doi.org/10.1016/j.brainres.2007.09.076

Koppen, C., & Spence, C. (2007c). Seeing the light: exploring the Colavita visual dominance effect. *Brain Research*, *1186*(1), 224–232. http://doi.org/10.1016/j.brainres.2007.09.076

Koppen, C., & Spence, C. (2007d). Spatial coincidence modulates the Colavita visual dominance effect. *Neuroscience Letters*, *417*, 107–111. http://doi.org/10.1016/j.brainres.2007.09.076

Kotowski, K., Stapor, K., Leski, J., & Kotas, M. (2018). Validation of Emotiv EPOC+ for extracting ERP correlates of emotional face processing. *Biocybernetics and Biomedical Engineering*, *38*(4), 773–781. http://doi.org/10.1016/j.bbe.2018.06.006

Kraft, J. M., & Werner, J. S. (1999). Aging and the saturation of colors. 1. Colorimetic purity discrimination. *Journal of the Optical Society of America. A*, *16*(2), 231–235. http://doi.org/10.1364/JOSAA.16.000231

Kray, J., Eppinger, B., & Mecklinger, A. (2005). Age differences in attentional control: An event-related potential approach. *Psychophysiology*, *42*(4), 407–416. http://doi.org/10.1111/j.1469-8986.2005.00298.x

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

Larson, M. J., Kaufman, D. A. S., & Perlstein, W. M. (2009). Neural time course of conflict adaptation effects on the Stroop task. *Neuropsychologia*, *47*(3), 663–670. http://doi.org/10.1016/j.neuropsychologia.2008.11.013

Laurienti, P. J., Burdette, J. H., Maldjian, J. A., & Wallace, M. T. (2006). Enhanced multisensory integration in older adults. *Neurobiology of Aging*, *27*(8), 1155–1163. http://doi.org/10.1016/j.neurobiolaging.2005.05.024

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

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(3), 451–468. http://doi.org/10.1037/0096-1523.21.3.451

Lewis, J. W., Beauchamp, M. S., & DeYoe, E. A. (2000). A comparison of visual and auditory motion processing in human cerebral cortex. *Cerebral Cortex*, *10*(9), 873–888. http://doi.org/10.1093/cercor/10.9.873

Lewkowicz, D. J., & Flom, R. (2014). The audio-visual temporal binding window narrows in early childhood. *Child Development*, *85*(2), 685–694. http://doi.org/10.1111/cdev.12142

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, *49*(4), 764–766. http://doi.org/10.1016/j.jesp.2013.03.013

Li, Y., Liu, M., Zhang, W., Huang, S., Zhang, B., Liu, X., & Chen, Q. (2017). Neurophysiological correlates of visual dominance: A lateralized readiness potential investigation. *Frontiers in Psychology*, *8*, 303. http://doi.org/10.3389/fpsyg.2017.00303

Lorist, M. M., & Jolij, J. (2012). Trial history effects in Stroop task performance are independent of top-down control. *PLoS ONE*, *7*(6). http://doi.org/10.1371/journal.pone.0039802

Lovelace, C. T., Stein, B. E., & Wallace, M. T. (2003). An irrelevant light enhances auditory detection in humans: A psychophysical analysis of multisensory integration in stimulus detection. *Cognitive Brain Research*, *17*(2), 447–453. http://doi.org/10.1016/S0926-6410(03)00160-5

Lutfi-proctor, D. A., Elliott, E. M., & Cowan, N. (2014). The role of visual stimuli in cross-modal Stroop interference. *Psych J*, *3*(1), 17–29. http://doi.org/10.1002/pchj.51.The

Lüttke, C. S., Ekman, M., Van Gerven, M. A. J., & De Lange, F. P. (2016). McGurk illusion recalibrates subsequent auditory perception. *Scientific Reports*, *6*(September), 1–7. http://doi.org/10.1038/srep32891

Macaskill, P., Walter, S. D., & Irwig, L. (2001). A comparison of methods to detect publication bias in meta-analysis. *Statistics in Medicine*, *20*(4), 641–654. http://doi.org/10.1002/sim.698

Macdonald, A. W. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, *288*, 1835–1838. http://doi.org/10.1126/science.288.5472.1835

Macleod, C. M. (1998). Training on integrated versus separated Stroop tasks: The progression of interference and facilitation. *Memory & Cognition*, *26*(2), 201–211. http://doi.org/10.3758/BF03201133

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*(2), 163–203. http://doi.org/10.1037/0033-2909.109.2.163

MacLeod, C. M. (1992). The Stroop task: The “gold standard” of attentional measures. *Journal of Experimental Psychology: General*, *121*(1), 12–14. http://doi.org/10.1037/0096-3445.121.1.12

Macleod, C. M., & Hodder, S. L. (1998). Presenting two incongruent color words on a single trial does not alter Stroop interference. *Memory and Cognition*, *26*(2), 212–219. http://doi.org/10.3758/BF03201134

Magnotti, J. F., & Beauchamp, M. S. (2015). The noisy encoding of disparity model of the McGurk effect. *Psychonomic Bulletin & Review*, *22*(3), 701–709. http://doi.org/10.1002/nbm.3369.Three

Mallick, D. B., Magnotti, J. F., & Beuchamp, M. S. (2015). Variability and stability in the McGurk effect: contributions of participants, stimuli, time and response type. *Psychonomic Bulletin & Review*, *22*(5), 1299–1307. http://doi.org/10.1016/j.coviro.2015.09.001.Human

Markela-Lerenc, J., Ille, N., Kaiser, S., Fiedler, P., Mundt, C., & Weisbrod, M. (2004). Prefrontal-cingulate activation during executive control: which comes first? *Cognitive Brain Research*, *18*(3), 278–287. http://doi.org/10.1016/j.cogbrainres.2003.10.013

Masaki, H., Wild-Wall, N., Sangals, J., & Sommer, W. (2004). The functional locus of the lateralized readiness potential. *Psychophysiology*, *41*(2), 220–230. http://doi.org/10.1111/j.1469-8986.2004.00150.x

Massaro, D. W., Thompson, L. A., Barron, B., & Laren, E. (1986). Developmental changes in visual and auditory contributions to speech perception. *Journal of Experimental Child Psychology*, *41*(1), 93–113. http://doi.org/10.1016/0022-0965(86)90053-6

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. http://doi.org/10.1016/j.cognition.2014.11.031

Matusz, P. J., Dikker, S., Huth, A. G., & Perrodin, C. (2018). Are we ready for real-world neuroscience ? *Journal of Cognitive Neuroscience*. http://doi.org/10.1162/jocn_e_01276

Mayer, A. R., Franco, A. R., Canive, J., & Harrington, D. L. (2009). The effects of stimulus modality and frequency of stimulus presentation on cross-modal distraction. *Cerebral Cortex*, *19*(5), 993–1007. http://doi.org/10.1093/cercor/bhn148

Mayer, A. R., & Kosson, D. S. (2004). The effects of auditory and visual linguistic distractors on target localization. *Neuropsychology*, *18*(2), 248–257. http://doi.org/10.1037/0894-4105.18.2.248

Mayer, A. R., Ryman, S. G., Hanlon, F. M., Dodd, A. B., & Ling, J. M. (2017). Look hear! The prefrontal cortex is stratified by modality of sensory input during multisensory cognitive control. *Cerebral Cortex*, *27*(5), 2831–2840. http://doi.org/10.1093/cercor/bhw131

McArthur, G. M., Atkinson, C., & Ellis, D. (2009). Atypical brain responses to sounds in children with specific language and reading impairments. *Developmental Science*, *12*(5), 768–783. http://doi.org/10.1111/j.1467-7687.2008.00804.x

McArthur, G. M., & Bishop, D. V. M. (2004). Which people with specific language impairment have auditory processing deficits? *Cognitive Neuropsychology*, *21*(1), 79–94. http://doi.org/10.1080/02643290342000087

McClain, L. (1983). Effects of response type and set size on Stroop color–word performance. *Perceptual and Motor Skills*, *56*(3), 735–743. http://doi.org/10.2466/pms.1983.56.3.735

McCormick, D. A., & Bal, T. (1994). Sensory gating mechanisms of the thalamus. *Current Opinion in Neurobiology*, *4*(4), 550–556. http://doi.org/10.1016/0959-4388(94)90056-6

McGraw, K. O., & Wong, S. . (1992). A common language effect size statistic. *Psychological Bulletin*, *111*(2), 361–365.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746–748. http://doi.org/10.1038/264746a0

Melara, R. D., & Algom, D. (2003). Driven by information: a tectonic theory of Stroop effects. *Psychological Review*, *110*(3), 422–471. http://doi.org/10.1037/0033-295X.110.3.422

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.

Miles, C., & Jones, D. M. (1989). The fallacy of the cross-modal Stroop effect : A rejoinder to Cowan (1989). *Perception & Psychophysics*, *45*(I), 85–86.

Miles, C., Madden, C., & Jones, D. M. (1989). Cross-modal , auditory-visual Stroop interference : A reply to Cowan and Barron (1987). *Perception & Psychophysics*, *45*(I), 77–81.

Milham, M. P., Banich, M. T., Webb, A., Barad, V., Cohen, N. J., Wszalek, T., & Kramer, A. F. (2001). The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on nature of conflict. *Cognitive Brain Research*, *12*, 467–473. http://doi.org/10.1016/S0926-6410(01)00076-3

Milham, M. P., Erickson, K. I., Banich, M. T., Kramer, A. F., Webb, A., Wszalek, T., & Cohen, N. J. (2002). Attentional control in the aging brain: Insights from an fMRI study of the Stroop task. *Brain and Cognition*, *49*(3), 277–296. http://doi.org/10.1006/brcg.2001.1501

Miller, G. A. (1947). The masking of speech. *Psychological Bulletin*, *44*(2), 105–129.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, a H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “Frontal Lobe” tasks: a latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. http://doi.org/10.1006/cogp.1999.0734

Miyakoshi, M., & Kothe, C. (2014). clean_rawdata. Retrieved from http://sccn.ucsd.edu/wiki/Plugin_list_process

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine*, *6*(7). http://doi.org/https://doi.org/10.1371/journal.pmed.1000097

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. http://doi.org/10.1016/S0926-6410(02)00066-6

Monem, R. G., & Fillmore, M. T. (2016). Alcohol-related visual cues impede the ability to process auditory information : Seeing but not hearing. *Psychology of Addictive Behaviour*, *30*(1), 12–17. http://doi.org/10.1037/adb0000140.Alcohol-Related

Moore, T., & Armstrong, K. M. (2003). Selective gating of visual signals by microstimulation of frontal cortex. *Nature*, *421*, 370–373. http://doi.org/10.1038/nature01285.

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, *229*(4715), 782–784. http://doi.org/10.1126/science.4023713

Morey, C. C., Elliott, E. M., Wiggers, J., Eaves, S. D., Shelton, J. T., & Mall, J. T. (2012). Goal-neglect links Stroop interference with working memory capacity. *Acta Psychologica*, *141*(2), 250–260. http://doi.org/10.1016/j.actpsy.2012.05.013

Morgan, A. L. R., & Brandt, J. F. (1989). An auditory Stroop effect for pitch, loudness, and time. *Brain and Language*, *36*(4), 592–603. http://doi.org/10.1016/0093-934X(89)90088-6

Moro, S. S., & Steeves, J. K. E. (2012). No Colavita effect: Equal auditory and visual processing in people with one eye. *Experimental Brain Research*, *216*(3), 367–373. http://doi.org/10.1007/s00221-011-2940-4

Moro, S. S., & Steeves, J. K. E. (2013). No Colavita effect: Increasing temporal load maintains equal auditory and visual processing in people with one eye. *Neuroscience Letters*, *556*, 186–190. http://doi.org/10.1016/j.neulet.2013.09.064

Motter, B. C. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *Journal of Neurophysiology*, *70*(3), 909–919. http://doi.org/0022-3077/93

Mozolic, J. L., Joyner, D., Hugenschmidt, C. E., Peiffer, A. M., Kraft, R. A., Maldjian, J. A., & Laurienti, P. J. (2008). Cross-modal deactivations during modality-specific selective attention. *BMC Neurology*, *8*(35), 1–11. http://doi.org/10.1186/1471-2377-8-35

Müller, N., Leske, S., Hartmann, T., Szebényi, S., & Weisz, N. (2015). Listen to yourself: The medial prefrontal cortex modulates auditory alpha power during speech preparation. *Cerebral Cortex*, *25*(11), 4029–4037. http://doi.org/10.1093/cercor/bhu117

Müller, V. I., Cieslik, E. C., Kellermann, T. S., & Eickhoff, S. B. (2013). Crossmodal emotional integration in major depression. *Social Cognitive and Affective Neuroscience*, *9*(6), 839–848. http://doi.org/10.1093/scan/nst057

Napolitano, A. C., & Sloutsky, V. M. (2004). Is a picture worth a thousand words? The flexible nature of modality dominance in young children. *Child Development*, *75*(6), 1850–1870. http://doi.org/10.1111/j.1467-8624.2004.00821.x

Narinesingh, C., Goltz, H. C., Raashid, R. A., & Wong, A. M. (2015). Developmental trajectory of McGurk effect susceptibility in children and adults with amblyopia. *Investigative Ophthalmology & Visual Science*, *56*(3), 2107–2113. http://doi.org/10.1167/iovs.14-15898

Narinesingh, C., Wan, M., Goltz, H. C., Chandrakumar, M., & Wong, A. M. F. (2014). Audiovisual perception in adults with amblyopia: A study using the McGurk effect. *Investigative Ophthalmology and Visual Science*, *55*(5), 3158–3164. http://doi.org/10.1167/iovs.14-14140

Nava, E., & Pavani, F. (2013). Changes in sensory dominance during childhood: Converging evidence from the Colavita effect and the sound-induced flash illusion. *Child Development*, *84*(2), 604–616. http://doi.org/10.1111/j.1467-8624.2012.01856.x

Naylor, L. J., Stanley, E. M., & Wicha, N. Y. Y. (2012). Cognitive and electrophysiological correlates of the bilingual stroop effect. *Frontiers in Psychology*, *3*, 81. http://doi.org/10.3389/fpsyg.2012.00081

Ngo, M. K., Cadieux, M. L., Sinnett, S., Soto-Faraco, S., & Spence, C. (2011). Reversing the Colavita visual dominance effect. *Experimental Brain Research*, *214*(4), 607–618.

http://doi.org/10.1007/s00221-011-2859-9

Ngo, M. K., Sinnett, S., Soto-Faraco, S., & Spence, C. (2010). Repetition blindness and the Colavita effect. *Neuroscience Letters*, *480*(3), 186–190. http://doi.org/10.1016/j.neulet.2010.06.028

Nguyen-Tri, D., Overbury, O., & Faubert, J. (2003). The role of lenticular senescence in age-related color vision changes. *Investigative Ophthalmology and Visual Science*, *44*(8), 3698–3704. http://doi.org/10.1167/iovs.02-1191

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin*, *126*(2), 220.

Noel, J., Niear, M. De, Burg, E. Van Der, & Wallace, M. T. (2016). Audiovisual simultaneity judgment and rapid recalibration throughout the lifespan. *PloS One*, *11*(8), 1–14. http://doi.org/10.1371/journal.pone.0161698

Noesselt, T., Bergmann, D., Hake, M., Heinze, H. J., & Fendrich, R. (2008). Sound increases the saliency of visual events. *Brain Research*, *1220*, 157–163. http://doi.org/10.1016/j.brainres.2007.12.060

Nolan, H., Whelan, R., & Reilly, R. B. (2010). FASTER: Fully Automated Statistical Thresholding for EEG artifact Rejection. *Journal of Neuroscience Methods*, *192*(1), 152–162. http://doi.org/10.1016/j.jneumeth.2010.07.015

Occelli, V., Hartcher-O’Brien, J., Spence, C., & Zampini, M. (2010). Assessing the audiotactile Colavita effect in near and rear space. *Experimental Brain Research*, *203*(3), 517–532. http://doi.org/10.1007/s00221-010-2255-x

Odgaard, E. C., Arieh, Y., & Marks, L. E. (2003). Cross-modal enhancement of perceived brightness : Sensory interaction versus response bias. *Perception & Psychophysics*, *65*(1), 123–132.

Odgaard, E. C., Arieh, Y., & Marks, L. E. (2004). Brighter noise:

Sensory enhancement of perceived loudness by concurrent visual stimulation. *Cognitive, Affective and Behavioral Neuroscience*, *4*(2), 127–132. http://doi.org/10.3758/CABN.4.2.127

Okajima, K., Tsuchiya, N., & Yamashita, K. (2002). Age-related changes in color appearance depend on unique-hue components. In *Proc. SPIE 4421, 9th Congress of the International Colour Association* (p. 259). http://doi.org/10.1117/12.464589

Osborn, W. C., Sheldon, R. W., & Baker, R. A. (1963). Vigilance performance under conditions of redundant and nonredundant signal presentation. *Journal of Applied Psychology*, *47*(2), 130–134. http://doi.org/10.1037/h0040874

Ouhaz, Z., Fleming, H., & Mitchell, A. S. (2018). Cognitive functions and neurodevelopmental disorders involving the prefrontal cortex and mediodorsal thalamus. *Frontiers in Neuroscience*, *12*, 33. http://doi.org/10.3389/fnins.2018.00033

Ousterhout, T. (2016). BCI effectiveness test through N400 replication study. In *Proceedings of the 4th European and 7th Nordic Symposium on Multimodal Communication*.

Paneri, S., & Gregoriou, G. G. (2017). Top-down control of visual attention by the prefrontal cortex. Functional specialization and long-range interactions. *Frontiers in Neuroscience*, *11*, 545. http://doi.org/10.3389/fnins.2017.00545

Parker, J. L., & Robinson, C. W. (2018). Changes in multisensory integration across the lifespan. *Psychology and Aging*, *33*(3), 545–558. http://doi.org//10.1037/pag0000244

Peiffer, A. M., Hugenschmidt, C. E., Haldjian, J. A., Casanova, R., Srlkanth, R., Hayasaka, S., … Laurienti, P. J. (2009). Aging and the interaction of sensory cortical function and structure. *Human Brain Mapping*, *30*(1), 228–240. http://doi.org/10.1002/hbm.20497

Peirce, J. W. (2007). Psychopy - Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13. http://doi.org//10.1016/j.jneumeth.2006.11.017

Peirce, J. W. (2009). Generating stimuli for neuroscience using Psychopy. *Frontiers in Neuroinformatics*, *2*, 10. http://doi.org/10.3389/neuro.11.010.2008

Pelli, D. G., & Zhang, L. (1991). Accurate control of contrast on microcomputer displays. *Vision Research*, *31*(7), 1337–1350. http://doi.org/10.1016/0042-6989(91)90055-A

Petit, S., Badcock, N. A., Grootswagers, T., Dermody, N., Yau, S., Schmidt, E., … Woolgar, A. (n.d.). *Detecting language comprehension in individual children using the N400 and Emotiv EPOC+ EEG*. Sydner.

Pfefferbaum, A., Mathalon, D. H., Sullivan, E. V, Rawles, J. M., Zipursky, R. B., & Lim, K. O. (1994). A quantitative magnetic resonance imaging study of changes in brain morphology from infancy to late adulthood. *Arch Neurol*, *51*(9), 874–887. http://doi.org/8080387

Pham, T. D., & Tran, D. (2012). Emotion recognition using the Emotiv EPOC device. In T. Huang, Z. Zeng, C. Li, & C. S. Leung (Eds.), *International Conference on Neural Information Processing* (pp. 394–399). Berlin: Springer. http://doi.org/DOI https://doi.org/10.1007/978-3-642-34500-5_47

Podlesek, A., & Komidar, L. (2006). Comparison of three psychophysical methods for measuring displacement in frontal plane motion. *Review of Pyschology*, *13*(1), 51–60.

Poldrack, R. A., Baker, C. I., Durnez, J., Gorgolewski, K. J., Matthews, P. M., Munafò, M. R., … Yarkoni, T. (2017). Scanning the horizon: Towards transparent and reproducible neuroimaging research. *Nature Reviews Neuroscience*, *18*(2), 115–126. http://doi.org/10.1038/nrn.2016.167

Polly, D. B., Hillock, A. R., Spankovich, C., Popescu, M. V., Royal, D. W., & Wallace, M. T. (2008). Development and plasticity of intra- and intersensory information processing. *Journal of American Academy of Audiology*, *19*(10), 780–798.

Prakash, R. S., Erickson, K. I., Colcombe, S. J., Kim, J. S., Voss, M. W., & Kramer, A. F. (2009). Age-related differences in the involvement of the prefrontal cortex in attentional control. *Brain and Cognition*, *71*(3), 328–335. http://doi.org/10.1016/j.bandc.2009.07.005.Age-related

Ratti, E., Waninger, S., Berka, C., Ruffini, G., & Verma, A. (2017). Comparison of medical and consumer wireless EEG systems for use in clinical trials. *Frontiers in Human Neuroscience*, *11*(398). http://doi.org/10.3389/fnhum.2017.00398

Raz, N. (1997). Selective aging of the human cerebral cortex observed in vivo: differential vulnerability of the prefrontal gray matter. *Cerebral Cortex*, *7*(3), 268–282. http://doi.org/10.1093/cercor/7.3.268

Raz, N. (2000). Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings. In F. I. M. Craik & T. A. Salthouse (Eds.), *Handbook of aging and cognition* (Vol. 2). Mahwah, NJ: Erlbaum.

Razumiejczyk, E., Girardi, C. P., Crivello, M. del C., Fioramonti, M., Macbeth, G., & Marmolejo-Ramos, F. (2016). Crossmodal interference between language and flavour. *Revista Latinoamericana de Psicologia*, *49*(2), 91–101. http://doi.org/10.1016/j.rlp.2016.01.002

Razumiejczyk, E., Macbeth, G., Marmolejo-Ramos, F., & Noguchi, K. (2015). Crossmodal integration between visual linguistic information and flavour perception. *Appetite*, *91*, 76–82. http://doi.org/10.1016/j.appet.2015.03.035

Ries, A. J., Touryan, J., Vettel, J., McDowell, K., & Hairston, W. D. (2014). A comparison of electroencephalography signals acquired from conventional and mobile systems. *Journal of Neuroscience and Neuroengineering*, *3*(1), 10–20. http://doi.org/10.1166/jnsne.2014.1092

Robinson, C. W., Chandra, M., & Sinnett, S. (2016a). Existence of competing modality dominances. *Attention, Perception, & Psychophysics*, *78*(4), 1104–1114. http://doi.org/10.3758/s13414-016-1061-3

Robinson, C. W., Chandra, M., & Sinnett, S. (2016b). Existence of competing modality dominances. *Attention, Perception & Psychophysics*, *78*(4), 1104–14. http://doi.org/10.3758/s13414-016-1061-3

Robinson, C. W., & Sloutsky, V. M. (2004). Auditory dominance and its change in the course of development. *Child Development*, *75*(5), 1387–401. http://doi.org/10.1111/j.1467-8624.2004.00747.x

Robinson, C. W., & Sloutsky, V. M. (2010). Development of cross-modal processing. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(1), 135–141. http://doi.org/10.1002/wcs.12

Roelofs, A. (2005). The visual-auditory color-word Stroop asymmetry and its time course. *Memory & Cognition*, *33*(8), 1325–1336. http://doi.org/10.3758/BF03193365

Ruff, C. C., Blankenburg, F., Bjoertomt, O., Bestmann, S., Freeman, E., Haynes, J. D., … Driver, J. (2006). Concurrent TMS-fMRI and psychophysics reveal frontal influences on human retinotopic visual cortex. *Current Biology*, *16*(15), 1479–1488. http://doi.org/10.1016/j.cub.2006.06.057

Savoy, R. L. (1986). Making quantized images appear smooth: Tricks of the trade in vision research. *Behavior Research Methods, Instruments, & Computers*, *18*(6), 507–517. http://doi.org/10.3758/BF03201422

Schall, J. D. (2009). Frontal Eye Fields. In *Encyclopedia of Neuroscience* (pp. 367–374). http://doi.org/10.1016/B978-008045046-9.01111-6

Schorr, E. A., Fox, N. A., van Wassenhove, V., & Knudsen, E. I. (2005). Auditory-visual fusion in speech perception in children with cochlear implants. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(51), 18748–18750.

http://doi.org/10.1073/pnas.0508862102

Sekiyama, K., & Burnham, D. (2008). Impact of language on development of auditory-visual speech perception. *Developmental Science*, *11*(2), 306–320. http://doi.org/10.1111/j.1467-7687.2008.00677.x

Sekiyama, K., Soshi, T., & Sakamoto, S. (2014). Enhanced audiovisual integration with aging in speech perception : a heightened McGurk effect in older adults. *Frontiers in Psychology*, *5*, 323. http://doi.org/10.3389/fpsyg.2014.00323

Sekiyama, K., & Tohkura, Y. (1991). McGurk effect in non-English listeners: Few visual effects for Japanese subjects hearing Japanese syllables of high auditory intelligibility. *The Journal of the Acoustical Society of America*, *90*(4 Pt 1), 1797–1805. http://doi.org/https://doi.org/10.1121/1.401660

Sekuler, R., Sekuler, A. B., & Lau, R. (1997). Sound alters visual motion perception. *Nature*. http://doi.org/10.1038/385308a0

Setti, A., & Chan, J. S. (2011). Familiarity of objects affects susceptibility to the sound-induced flash illusion. *Neuroscience Letters*, *492*(1), 19–22. http://doi.org/10.1016/j.neulet.2011.01.042

Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature*, *408*, 788. http://doi.org/10.1038/35048669

Shapiro, K. L., Egerman, B., & Klein, R. M. (1984). Effects of arousal on human visual dominance. *Perception & Psychophysics*, *35*(6), 547–552. http://doi.org/10.3758/BF03205951

Shapiro, K. L., & Johnson, T. L. (1987). Effects of arousal on attention to central and peripheral visual stimuli. *Acta Psychologica*, *66*(2), 157–172. http://doi.org/10.1016/0001-6918(87)90031-X

Shimada, H. (1990). Effect of auditory presentation of words on color naming: the intermodal Stroop effect. *Perceptual and Motor Skills*, *30*(3), 1155–1161.

Shimada, H., & Ashitaka, Y. (2010). Are Stroop facilitation and interference effects enhanced by an accompanying auditory word?

In *Proceedings of the 26th Annual Meetings of the International Society for Psychophysics*.

Shor, R. E. (1975). An auditory analog of the Stroop test. *The Journal of Genetic Psychology*, *93*, 281–288.

Shulman, G., Corbetta, M., Buckner, R. L., Raichlet, M. E., Fiez, J. A., Miezin, F. M., & Petersen, S. E. (1997). Top-down modulation of early sensory cortex. *Cerebral Cortex*, *7*(3), 193–206. http://doi.org/10.1093/cercor/7.3.193

Silton, R. L., Heller, W., Towers, D. N., Engels, A. S., Spielberg, J. M., Edgar, J. C., … Miller, G. A. (2010). The time course of activity in dorsolateral prefrontal cortex and anterior cingulate cortex during top-down attentional control. *NeuroImage*, *50*(3), 1292–302. http://doi.org/10.1016/j.neuroimage.2009.12.061

Sinnett, S., Soto-Faraco, S., & Spence, C. (2008). The co-occurrence of multisensory competition and facilitation. *Acta Psychologica*, *128*(1), 153–161. http://doi.org/10.1016/j.actpsy.2007.12.002

Sinnett, S., Spence, C., & Soto-Faraco, S. (2007). Visual dominance and attention: The Colavita effect revisited. *Perception & Psychophysics*, *69*(5), 673–686. http://doi.org/10.3758/BF03193770

Sloutsky, V. M., & Napolitano, A. C. (2003). Is a picture worth a thousand words? Preference for auditory modality in young children. *Child Development*, *74*(3), 822–833. http://doi.org/10.1111/1467-8624.00570

Smulders, F. T. Y., Kok, A., Kenemans, J. L., & Bashore, T. R. (1995). The temporal selectivity of additive factor effects on the reaction process revealed in ERP component latencies. *Acta Psychologica*, *90*(1–3), 97–109. http://doi.org/10.1016/0001-6918(95)00032-P

Solé-Padullés, C., Castro-Fornieles, J., De La Serna, E., Calvo, R., Baeza, I., Moya, J., … Sugranyes, G. (2016). Intrinsic connectivity networks from childhood to late adolescence: Effects of age and sex. *Developmental Cognitive Neuroscience*, *17*, 35–44.

http://doi.org/10.1016/j.dcn.2015.11.004

Sowell, E. R., Delis, D., Stiles, J., & Jernigan, T. L. (2001). Improved memory functioning and frontal lobe maturation between childhood and adolescence: A structural MRI study. *Journal of the International Neuropsychological Society*, *7*(3), 312–322. http://doi.org/10.1017/S135561770173305X

Spence, C. (2009). Explaining the Colavita visual dominance effect. *Progress in Brain Research*, *176*, 245–258. http://doi.org/10.1016/S0079-6123(09)17615-X

Spence, C., Parise, C., & Chen, Y. C. (2012). The Colavita visual dominance effect. In M. M. Murray & M. T. Wallace (Eds.), *The Neural Bases of Multisensory Processes* (pp. 529–557). Boca Raton (FL): Taylor & Francis.

Sreenivasan, K. K., Curtis, C. E., & D'Esposito, M. (2014). Revisiting the role of persistent neural activity during working memory. *Trends in Cognitive Sciences*, *18*(2), 82–89. http://doi.org/10.1016/j.tics.2013.12.001

Stacey, J. E., Howard, C., Mitra, S., & Stacey, P. (2017). The relationship between eye movements and the McGurk effect when stimuli are presented in noise. In *Poster presented at: International Multisensory Research Forum*. Nashville, TN.

Stanford, T. R., Wallace, M. T., Vaughan, W. J., & Jiang, W. (2004). Crossmodal spatial interactions in subcortical and cortical circuits. In C. Spence & J. Driver (Eds.), *Crossmodal space and crossmodal attention* (pp. 25–50). Oxford: Oxford University Press.

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.

Stekelenburg, J. J., & Keetels, M. (2016). The effect of synesthetic associations between the visual and auditory modalities on the Colavita effect. *Experimental Brain Research*, *234*(5), 1209–1219. http://doi.org/10.1007/s00221-015-4363-0

Stern, R. M., Brown, G. J., & Wang, D. (2006). Binaural sound localisation. In D. Wang & G. J. Brown (Eds.), *Computational Auditory Scene Analysis: Principles, Algorithms and Applications* (pp. 147–185). Wiley-IEEE Press.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. http://doi.org/10.1037/h0054651

Stuart, D. M., & Carrasco, M. (1993). Semantic component of a cross-modal stroop-like task. *The American Journal of Psychology*, *106*(3), 383–405.

Stubblefield, A., Jacobs, L. A., Kim, Y., & Goolkasian, P. (2013). Colavita dominance effect revisited: The effect of semantic congruity. *Attention, Perception & Psychophysics*, *75*(8), 1827–39. http://doi.org/10.3758/s13414-013-0530-1

Sufani, C., De Blasio, F. M., McDonald, S., & Rushby, J. A. (2015). Validating the use of Emotiv EPOC in resting EEG coherence research. In *25th Annual Conference of the Australasian Society for Pscyhophysiology*. Front. Hum. Neuroscie. http://doi.org/doi: 10.3389/conf.fnhum.2015.219.00013

Swick, D., & Jovanovic, J. (2002). Anterior cingulate cortex and the Stroop task: neuropsychological evidence for topographic specificity. *Neuropsychologia*, *40*(8), 1240–1253. http://doi.org/10.1016/S0028-3932(01)00226-3

Szűcs, D., & Soltész, F. (2007). Event-related potentials dissociate facilitation and interference effects in the numerical Stroop paradigm. *Neuropsychologia*, *45*(14), 3190–3202. http://doi.org/10.1016/j.neuropsychologia.2007.06.013

Szűcs, D., & Soltész, F. (2010). Stimulus and response conflict in the color-word Stroop task: A combined electro-myography and event-related potential study. *Brain Research*, *1325*, 63–76. http://doi.org/10.1016/j.brainres.2010.02.011

Szűcs, D., & Soltész, F. (2012). Functional definition of the N450 event-related brain potential marker of conflict processing: a numerical Stroop study. *BMC Neuroscience*, *13*(35), 1–14. http://doi.org/10.1186/1471-2202-13-35

Szűcs, D., Soltész, F., Bryce, D., & Whitebread, D. (2009). Real-time tracking of motor response activation and response competition in a Stroop task in young children: a lateralized readiness potential study. *Journal of Cognitive Neuroscience*, *21*(11), 2195–2206. http://doi.org/10.1162/jocn.2009.21220

Thackray, R. I., & Jones, K. N. (1971). Level of arousal during Stroop performance: Effects of speed stress and “distraction.” *Psychonomic Science*, *23*(2), 133–135. http://doi.org/10.3758/BF03336518

Thie, J. (2013). A wireless marker system to enable evoked potential recordings using a wireless EEG system (EPOC) and a portable computer. *PeerJ PrePrints*, 1–24. http://doi.org/10.7287/peerj.preprints.32v1

Thomas, R. L., Nardini, M., & Mareschal, D. (2017). The impact of semantically congruent and incongruent visual information on auditory object recognition across development. *Journal of Experimental Child Psychology*, *162*, 72–88. http://doi.org/10.1016/j.jecp.2017.04.020

Thompson, S. G., & Higgins, J. P. T. (2002). How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine*, *21*(11), 1559–1573. http://doi.org/10.1002/sim.1187

Thurlow, W. R., & Jack, C. E. (1973). Certain determinants of the “ventriloquism effect”. *Perceptual and Motor Skills*, *36*, 1171–1184. http://doi.org/10.2466/pms.1973.36.3c.1171

Townsend, J., Adamo, M., & Haist, F. (2006). Changing channels: An fMRI study of aging and cross-modal attention shifts. *NeuroImage*, *31*(4), 1682–1692. http://doi.org/10.1016/j.neuroimage.2006.01.045

Tremblay, C., Champoux, F., Voss, P., Bacon, B. A., Lepore, F., & Théoret, H. (2007). Speech and non-speech audio-visual illusions:

A developmental study. *PLoS ONE*, *2*(8). http://doi.org/10.1371/journal.pone.0000742

Treue, S. (2003). Visual attention: The where, what, how and why of saliency. *Current Opinion in Neurobiology*, *13*(4), 428–432. http://doi.org/10.1016/S0959-4388(03)00105-3

Tsujimoto, S. (2008). The prefrontal cortex: functional neural development during early childhood. *The Neuroscientist*, *14*(4), 345–358. http://doi.org/10.1177/1073858408316002

Tsukiko, M., & Desmarais, G. (2014). The impact of sensory dominance and congruency effects on multisensory integration. *Canadian Journal of Experimental Psychology*, *68*(4), 303.

Tun, P. A., Mccoy, S., & Wingfield, A. (2009). Aging, hearing acuity, and the attentional costs of effortful listening. *Psychology and Aging*, *24*(3), 761–766. http://doi.org/10.1037/a0014802

Tyler, C. W. (1997). Colour bit-stealing to enhance the luminance resolution of digital displays on a single pixel basis. *Spatial Vision*, *10*(4), 369–377. http://doi.org/10.1163/156856897X00294

Van Damme, S., Crombez, G., & Spence, C. (2009). Is visual dominance modulated by the threat value of visual and auditory stimuli? *Experimental Brain Research*, *193*(2), 197–204. http://doi.org/10.1007/s00221-008-1608-1

Van Gerven, P. W. M., & Guerreiro, M. J. S. (2016). Selective attention and sensory modality in aging: Curses and blessings. *Frontiers in Human Neuroscience*, *10*(147). http://doi.org/10.3389/fnhum.2016.00147

Verhaeghen, P., & De Meersman, L. (1998). Aging and the Stroop effect: A meta-analysis. *Psychology and Aging*, *13*(1), 120–126. http://doi.org/10.1037/0882-7974.13.1.120

Verhaeghen, P., Steitz, D. W., Sliwinski, M. J., & Cerella, J. (2003). Aging and dual-task performance: A meta-analysis. *Psychology and Aging*, *18*(3), 443–460. http://doi.org/10.1037/0882-7974.18.3.443

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. http://doi.org/10.1103/PhysRevB.91.121108

Wang, K. (2011). An electrophysiological investigation of the role of orthography in accessing meaning of Chinese single-character words. *Neuroscience Letters*, *487*(3), 297–301. http://doi.org/10.1016/j.neulet.2010.10.041

Wang, W., Viswanathan, S., Lee, T., & Grafton, S. T. (2016). Coupling between theta oscillations and cognitive control network during cross-modal visual and auditory attention: Supramodal vs modality-specific mechanisms. *PLoS ONE*, *11*(7), 1–16. http://doi.org/10.1371/journal.pone.0158465

Weijkamp, J., & Sadakata, M. (2017). Attention to affective audio-visual information: Comparison between musicians and non-musicians. *Psychology of Music*, *45*(2), 204–215. http://doi.org/10.1177/0305735616654216

Weissman, D. H., Warner, L. M., & Woldorff, M. G. (2004). The neural mechanisms for minimizing cross-modal distraction. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *24*(48), 10941–9. http://doi.org/10.1523/JNEUROSCI.3669-04.2004

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, *88*(3), 638–667. http://doi.org/http://dx.doi.org/10.1037/0033-2909.88.3.638

Welch, R. B., & Warren, D. H. (1986). Intersensory Interactions. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of Perception and Human Performance* (p. 25.1-25.36). New York: Wiley.

Werner, J. S., & Steele, V. G. (1988). Sensitivity of human foveal color mechanisms throughout the life span. *Journal of the Optical Society of America A*, *5*(12), 2122–2130. http://doi.org/10.1364/JOSAA.5.002122

West, R. (1996a). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin*, *120*(2), 272–292.

http://doi.org/10.1037/0033-2909.120.2.272

West, R. (1996b). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin*, *120*(2), 272–292. http://doi.org/10.1037/0033-2909.120.2.272

West, R., & Alain, C. (2000). Age-related decline in inhibitory control contributes to the increased Stroop effect observed in older adults. *Psychophysiology*, *37*(2), 179–189. http://doi.org/10.1017/S0048577200981460

West, R., & Bell, M. A. (1997). Stroop color-word interference and electroencephalogram activation: Evidence for age-related decline of the anterior attention system. *Neuropsychology*, *11*(3), 421–427. http://doi.org/10.1037/0894-4105.11.3.421

West, R., Bowry, R., & McConville, C. (2004). Sensitivity of medial frontal cortex to response and nonresponse conflict. *Psychophysiology*, *41*(5), 739–748. http://doi.org/10.1111/j.1469-8986.2004.00205.x

West, R., Jakubek, K., Wymbs, N., Perry, M., & Moore, K. (2005). Neural correlates of conflict processing. *Experimental Brain Research*, *167*(1), 38–48. http://doi.org/10.1007/s00221-005-2366-y

White, T. L., & Prescott, J. (2007). Chemosensory cross-modal stroop effects: Congruent odors facilitate taste identification. *Chemical Senses*, *32*(4), 337–341. http://doi.org/10.1093/chemse/bjm001

Wille, C., & Ebersbach, M. (2016). Semantic congruency and the (reversed) Colavita effect in children and adults. *Journal of Experimental Child Psychology*, *141*, 23–33. http://doi.org/10.1016/j.jecp.2015.07.015

Winkler, I., Haufe, S., & Tangermann, M. (2011). Automatic classification of artifactual ICA- components for artifact removal in EEG signals. *Behavioral and Brain Functions*, *7*(1), 30. http://doi.org/10.1186/1744-9081-7-30

Witten, I. B., & Knudsen, E. I. (2005). Why seeing is believing: Merging auditory and visual worlds. *Neuron*, *48*(3), 489–496. http://doi.org/10.1016/j.neuron.2005.10.020

Woldorff, M. G., & Hillyard, S. A. (1991). Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalography and Clinical Neurophysiology*, *79*, 170–191.

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

Xiao, X., Dupuis-Roy, N., Yang, X. L., Qiu, J. F., & Zhang, Q. L. (2014). The taste-visual cross-modal Stroop effect: An event-related brain potential study. *Neuroscience*, *263*, 250–256. http://doi.org/10.1016/j.neuroscience.2014.01.002

Yakovlev, P. I., & Lecours, A.-R. (1967). The myelogenetic cycles of regional maturation of the brain. In *Regional Development of Brain in Early Life* (pp. 3–70). http://doi.org/10.1002/ana.410220408/full

Yue, Z., Jiang, Y., Li, Y., Wang, P., & Chen, Q. (2015). Enhanced visual dominance in far space. *Experimental Brain Research*, *233*(10), 2833–2843. http://doi.org/10.1007/s00221-015-4353-2

Yuval-Greenberg, S., & Deouell, L. Y. (2009). The dog’s meow: Asymmetrical interaction in cross-modal object recognition. *Experimental Brain Research*, *193*(4), 603–614. http://doi.org/10.1007/s00221-008-1664-6

Zhang, H., & Kornblum, S. (1998). The effects of stimulus-response mapping and irrelevant stimulus-response and stimulus-stimulus overlap in four-choice Stroop tasks with single-carrier stimuli. *Journal of Experimental Psychology. Human Perception and Performance*, *24*(1), 3–19. http://doi.org/10.1037/0096-1523.24.1.3

Zinchenko, A., Obermeier, C., Kanske, P., Schröger, E., Villringer, A., & Kotz, S. A. (2017). The influence of negative emotion on cognitive and emotional control remains intact in aging. *Frontiers in Aging*

Neuroscience, *9*, 349. http://doi.org/10.3389/fnagi.2017.00349

Appendix

Supplementary analyses for Chapter 3

Analysis of raw response times and accuracy (Experiment 1)

In the main body of this thesis, ratio scores are used throughout and, although neutral conditions were employed in the Stroop task these results are not reported (due to the novelty effect seen with neutral stimuli). However, a 2 (sensory condition) x 4 (congruency) x 3 (age group) mixed ANOVA was performed to assess effects in raw response time and accuracy data. These are shown below.

Response times

There was a main effect of congruency ($F(2.65, 270.1) = 23.14$, $p < 0.001$, $\eta_p^2 = .18$) and a main effect of age group ($F(2, 102) = 40.315$, $p < .001$, $\eta_p^2 = .44$). The effect of congruency was not modulated by age group ($F(5.3, 270.1) = 1.93$, $p = .09$, $\eta_p^2 = .03$), or sensory condition ($F(2.49, 270.1) = 2.81$, $p = .05$, $\eta_p^2 = .03$) and there was no three way interaction between sensory condition, age group and congruency ($F(4.98, 254.18) = 2.079$, $p = .069$, $\eta_p^2 = .04$).

Response times were faster under congruent ($M = 703$ms, $SE = 17$ms) vs. response-incongruent ($M = 766$ms, $SE = 19$ms) conditions ($p < .001$), congruent vs. stimulus-incongruent ($M = 748$ms, $SE = 20$ms) conditions ($p < .001$) and congruent vs. neutral ($M = 748$ms, $SE = 19$ms) conditions ($p < .001$). Response times did not significantly differ between stimulus- and response-incongruent conditions ($p = .237$). Notably response times did not differ between neutral and stimulus-incongruent conditions ($p = 1$) or neutral and response-incongruent conditions ($p = .227$), suggesting interference occurring on neutral trials. This may have resulted from stimulus novelty (as the word “Brown” was not included during the thresholding procedure) or stimulus-interference itself. Indeed, previous literature has utilised stimuli that are not in the response set to probe stimulus-interference (Milham et al., 2001). For this reason neutral conditions were not used in the calculations of interference ratios noted within the main body of this thesis.

There was no main effect of sensory condition ($F(1, 102) = .004$, $p = .95$, $\eta_p^2 < .001$), suggesting similar response times between unimodal and cross-modal tasks generally. However the effect of age group was modulated by sensory condition ($F(2, 102) = 4.589$, $p = .012$, $\eta_p^2 = .08$). Children ($M = 929$ms, $SE = 29$ms) were slower than young adults ($M = 525$ms, $SE = 34$ms $p < .001$) and older adults ($M = 770$ms, $SE = 34$ms $p = .001$), and older adults were slower than young adults ($p < .001$). The interaction between age group and sensory condition occurred because, under unimodal conditions, children ($M = 904$ms, $SE = 32$ms) and older adults ($M = 801$ms, $SE = 37$ms) were significantly slower than young adults ($M = 517$ms, $SE = 38$ms) ($p < .001$) whilst children and older adults did not differ ($p = .12$). Conversely, under cross-modal conditions, children ($M = 953$ms, $SE = 31$ms) were significantly slower than young adults ($M = 532$ms, $SE = 36$ms) and older adults ($M = 740$ms, $SE = 36$ms) ($p < .001$) and older adults were significantly slower than young adults ($p < .001$). This finding coincides with the conclusion reported in the main body of this thesis, that children experience more cross-modal distraction compared with adults.

As shown in Table S 1 children and older adults manifested larger Stroop effects than younger adults (i.e. they were slower under incongruent vs. congruent conditions compared with young adults). However the interaction between congruency and age group did not reach significance. This is not in line with the age related Stroop effects reported in past unimodal literature (Comalli et al., 1962). Notably, the interaction between congruency, sensory condition and age group also did not reach significance, suggesting this was not due to the inclusion of cross-modal conditions.

Accuracy

Accuracy scores for each age group under each condition are shown in Table S 1. A 2 (sensory condition) x 4 (congruency) x 3 (age group) mixed ANOVA on the accuracy scores across groups revealed a main effect of congruency (*F*(2.63, 268.14) = 7.48, *p* < .001, η_p^2=.08) and a main effect of age group (*F*(2, 102) = 19.48, *p* < .001, η_p^2=.28). The main effect of congruency was not modulated by age group (*F*(5.26, 268.14) = 1.026, *p* = .4, η_p^2 = .02) suggesting similar congruency effects between groups. There was however a significant interaction between congruency and sensory condition (*F*(3, 306) = 5.16, *p* = .002, η_p^2 = .05) suggesting different congruency effects between unimodal and cross-modal conditions. The three way interaction between sensory condition, congruency and age group failed to meet significance (*F*(6, 306) = 2.25, *p* = .04, η_p^2 = .04).

Accuracy was significantly lower in response-incongruent (*M* = 90% *SE* = .75) vs. congruent (*M* = 92%, *SE* = .61) (*p* = .001) and response-incongruent vs. stimulus-incongruent (*M* = 92%, *SE* = .66) (*p* = .002) but not response-incongruent vs. neutral (*M* =91%, *SE* = .7) (*p* = .53) conditions. Notably, stimulus-incongruent conditions should not prime errors as irrelevant stimuli are also mapped to the correct responses (thus differences between response-incongruent and stimulus-incongruent conditions can also result from facilitation on stimulus-incongruent trials). Older adults (*M* = 96%, *SE* = 1.03) were significantly more accurate than younger adults (*M* = 90%, *SE* = 1.04, *p* < .001), and children (*M* = 88%, *SE* = .9, *p* < .001). However accuracy between children and adults did not significantly differ (*p* = .298).

The interaction between sensory condition and congruency occurred because the main effect of congruency only reached significance under unimodal (*p* < .001) and not cross-modal (*p* = .52) conditions. In line with the analyses reported in the main body of this thesis, accuracy effects were seen under unimodal but not cross-modal conditions. Unimodal response-incongruent conditions reduced accuracy (*M* = 88.63%, *SE* = .9) compared with congruent (*M* = 92%, *SE* = .73 *p* < .001), stimulus-incongruent (*M* = 93.37%, *SE* = .72 *p* < .001) and neutral conditions (*M* = 91.16%, *SE* = .76, p = .023), whilst accuracy did not significantly differ between unimodal stimulus-incongruent and congruent conditions *p* = .924. These differences did not occur under cross-modal conditions (respective *p* values under cross-modal conditions were all *p* = 1).

However, the main effect of sensory condition did not reach significance (F(1, 102) = .073, p = .79, η_p^2 < .001) and this was not modulated by age group (*F*(2, 102) = .474, *p* = .62, η_p^2 = .009), suggesting similar accuracy between unimodal and cross-modal conditions overall.

		Response times						Accuracy					
			95% CI			95%CI			95% CI			95%CI	
		Uni-modal	Lower	Upper	Cross-modal	Lower	Upper	Uni-modal	Lower	Upper	Cross-modal	Lower	Upper
Children	C	851	793	909	905	846	964	87.6	85.34	89.86	89.29	86.98	91.6
	SI	913	839	987	965	902	1028	89.19	86.96	91.41	88.49	86.01	90.97
	RI	944	872	1016	968	901	1035	85.62	82.84	88.4	86.01	83.27	88.76
	N	907	839	975	975	908	1042	87.6	85.24	89.96	88.19	85.35	91.04
Young adults	C	499	432	567	513	445	582	92.88	90.25	95.5	91.13	88.44	93.82
	SI	511	425	597	542	468	615	92.88	90.29	95.47	89.38	86.49	92.27
	RI	524	440	608	544	466	622	85.62	82.38	88.85	91.26	88.07	94.46
	N	535	456	615	530	452	607	89.79	87.03	92.54	88.04	84.73	91.34
Older adults	C	728	662	794	722	654	789	96.88	94.29	99.46	95.7	93.06	98.35
	SI	808	724	893	749	676	821	98.05	95.5	100.6	95.44	92.6	98.29
	RI	861	778	944	754	677	831	94.66	91.48	97.85	95.7	92.56	98.85
	N	805	727	883	734	658	811	96.09	93.39	98.8	96.62	93.36	99.87

Table S 1. RT and accuracy for each age group under congruent (C) stimulus-incongruent (SI) response-incongruent (RI) and neutral (N) conditions under unimodal and cross-modal sensory conditions for Chapter 3, Experiment 1. Confidence intervals (CI) represent the 95% bootstrapped confidence intervals.

Analysis of raw response times and accuracy (Experiment 2)

Response times

A 2 (distractor type; auditory vs. visual) x 4 (congruency; congruent, stimulus-incongruent, response-incongruent and neutral) x 3 (age group) ANOVA showed a main effect of distractor type ($F(1, 109) = 15.54$, $p < .001$, $\eta_p^2 = .11$) which significantly interacted with age group ($F(2, 109) = 5.65$, $p = .005$, $\eta_p^2 = .08$). Response times were significantly slower with visual ($M = 874$ms, $SE = 26$ms) vs. auditory ($M = 781$ms, $SE = 34$ms) distractors. However this difference was only seen in young adults ($p < .001$) and not children ($p = .72$) or older adults ($p = .12$). This is consistent with previous literature showing visual distractors have a stronger effect in young adults versus children (Donohue, Appelbaum, et al., 2013; Yuval-Greenberg & Deouell, 2009).

There was also a significant main effect of congruency ($F(2.58, 280.98) = 13.57$, $p < .001$, $\eta_p^2 = .11$), which interacted with age group ($F(5.16, 280.98) = 3.133$, $p = .008$, $\eta_p^2 = .05$). Response times were slower under response-incongruent ($M = 858$ms, $SE = 31$ms) vs. congruent ($M = 783$ms, $SE = 25$ms, $p < .001$) and stimulus-incongruent ($M = 858$ms, $SE = 31$ms, $p = .009$) conditions but response times did not significantly differ between response-incongruent and neutral ($M = 841$ms, $SE = 29$ms, $p = 1$). Response times were also slower under stimulus-incongruent vs. congruent conditions ($p = .001$) but did not significantly differ between stimulus-incongruent and neutral conditions ($p = 1$). As in Experiment 1, and despite presenting no stimuli other than the colour rectangle/word to be identified on neutral trials (compared with the word “Brown” in Experiment 1), response times were significantly slower on neutral vs. congruent ($p < .001$). As such the calculation of interference ratios within the main thesis were based on comparisons between stimulus and response-incongruent conditions with congruent conditions.

Interestingly the effect of congruency interacted with age group because the effect of congruency reached significance in children ($p < .001$)

but not young adults (*p* = .06) or older adults (*p* = .09). This is consistent with the results reported in the main body of this thesis, suggesting stronger cross-modal interference in children. The effect of congruency was not modulated by distractor type (*F*(2.66, 290.37) = .006, *p* = .35, η_p^2 =.003) and there was no three-way interaction between congruency, distractor type and age (*F*(5.33, 290.37) = .009, *p* = .72, η_p^2 = .01).

There was a main effect of age group overall (*F*(2, 109) = 20.28, *p* < .001, η_p^2 =.28). Children were significantly slower than younger and older adults (*p* < .001 for both comparisons) but response time did not significantly differ between young and older adult groups (*p* = 1). This, again, is consistent with the findings reported in the main thesis, suggesting children are more susceptible to cross-modal distraction, whilst older adults perform similarly to young adults.

Accuracy

A 2 x 4 x 3 ANOVA showed no main effect of distractor type on accuracy (*F*(1, 109) = 2.98, *p* = .09, η_p^2 = .03) and this was not influenced by age group (*F*(2, 109) = 1.09, *p* = .34, η_p^2 =.02). There was a significant main effect of congruency (*F*(2.69, 293.51) = 16.44, *p* < .001, η_p^2 = .12) and this interacted with age group (*F*(5.39, 293.51) = 3.51, *p* = .002, η_p^2 =.05). Accuracy was significantly lower under response-incongruent (*M* = 81.81%, *SE* = 1.15) vs. congruent (*M* = 86.28%, *SE* = 0.97, p < .001), stimulus-incongruent (*M* = 86.82, *SE* = 0.97, *p* < .001) and neutral (*M* = 86.49, *SE* = 0.98, *p* < .001) conditions. Accuracy under stimulus-incongruent, congruent and neutral conditions did not differ from one another (*p* = 1). Simple main effects showed the effect of congruency only reached significance in children (p < .001) and not young adults (*p* = .017) or older adults (*p* = .45). This also supports the conclusion drawn in Chapter 3 of this thesis, that children experienced more cross-modal interference than adults and this may be because cross-modal interference also induced response-interference in this age group.

The effect of congruency was not modulated by distractor type (*F*(2.49, 271.24) = .67, *p* = .57, η_p^2 =.006), suggesting similar congruency effects on accuracy with visual and auditory cross-modal distractors. There was no three-way interaction between congruency, distractor type and age group (*F*(4.98, 271.24) = 2.185, *p* = .044, η_p^2 = .04).

		Response times						Accuracy					
			95% CI			95%CI			95% CI			95%CI	
		Unimodal	Lower	Upper	Cross-modal	Lower	Upper	Unimodal	Lower	Upper	Cross-modal	Lower	Upper
Children	C	961	875	1048	978	908	1047	80.79	77.58	84	81.31	78.09	84.54
	SI	1060	959	1161	1082	1001	1163	84.28	81.05	87.51	84.22	81.13	87.32
	RI	1114	1005	1223	1097	1018	1177	73.06	68.94	77.17	76.84	73.15	80.52
	N	1045	940	1149	1072	996	1148	80.07	76.72	83.42	84.31	81.4	87.23
Young adults	C	560	447	674	737	645	829	86.25	82.03	90.48	88.51	84.26	92.76
	SI	570	437	703	784	678	891	84.32	80.06	88.57	89.45	85.37	93.52
	RI	610	466	753	793	688	897	83.01	77.59	88.43	80.89	76.04	85.74
	N	587	449	724	809	709	909	84.15	79.74	88.57	88.56	84.72	92.4
Older adults	C	698	584	812	762	670	853	89.51	85.29	93.74	91.28	87.03	95.53
	SI	693	559	826	775	668	882	89.51	85.26	93.76	89.17	85.09	93.25
	RI	723	580	866	813	709	917	89.13	83.71	94.55	87.92	83.07	92.76
	N	747	609	884	788	688	888	91.75	87.33	96.16	90.11	86.27	93.95

Table S 2. RT and accuracy for each age group under congruent (C) stimulus-incongruent (SI) response-incongruent (RI) and neutral (N) conditions under unimodal and cross-modal sensory conditions for Chapter 3, Experiment 2. Confidence intervals (CI) represent the 95% bootstrapped confidence intervals.

Supplementary analyses for Chapter 4

Additional tasks (Chapter 4)

Threshold task

Four thresholds were derived using four blocks of staircases in a counterbalanced order. Thresholds were derived for:

1. Reading colour-words on each colour background *without* a desaturation mask.
2. Reading colour-words on each colour background *with* a desaturation mask.
3. Hearing spoken colour-words with a randomly selected colour rectangle.
4. Identifying the colour of rectangles presented with a desaturation mask of varying opacity.

The staircase protocol used for reading and hearing words was identical to the staircase described in Chapter 3 Experiment 1 (adult participants). However timings were adjusted to reflect the timings of the Stroop task used in Chapter 4. Thresholds for colour identification were derived using four interleaved staircases (one for each colour) in which the opacity of a desaturation mask overlaying the colours adjusted in a 1-up 3-down protocol (step size (% opacity) = [30, 20, 10, 10, 5,1], start value = 0% opacity) until 8 reversals had been reached. Each threshold was taken as the average of the final 6 reversals. The written word remained the same throughout the task to prevent Stroop interference. Notably, thresholds for identifying colours were measured only to assess how visible colours were in the Stroop task relative to threshold. Unlike thresholds for reading words, colour thresholds were not used to control the opacity of the mask in the Stroop task (in which the mask was 60% opacity for all participants).

Response time task

Table S 3 illustrates the four types of choice response time tasks used to assess participants' speed for reading/hearing colour-words and naming colours under each saturation condition. Each response time task contained 40 trials preceded by 3 practice trials[34]. The trial structure of the response time task is illustrated in Figure S 1. This structure (fixation cross > blank period > stimulus presentation > response period > recovery period) was used to mimic the structure used in the main Stroop task whilst emphasising speed and minimising overall testing time.

RT assessed for	Rectangle colour	Written information	Auditory information	Desaturation mask
Reading colour-words	White	Colour-word	Auditory babble	No
Colour naming (high saturation)	Coloured	Written babble	Auditory babble	No
Colour naming (low saturation)	Coloured	Written babble	Auditory babble	Yes
Spoken word repetition	White	Written babble	Spoken colour-word	No

Table S 3. Design of the four choice response time tasks used to assess participants colour reading and colour naming speed.

---

[34] Pilot testing indicated the task was very simple and three trials were sufficient to accustom participants to the trial structure/stimuli. All participants confirmed prior to completing the main task that they understood the task well enough to proceed.

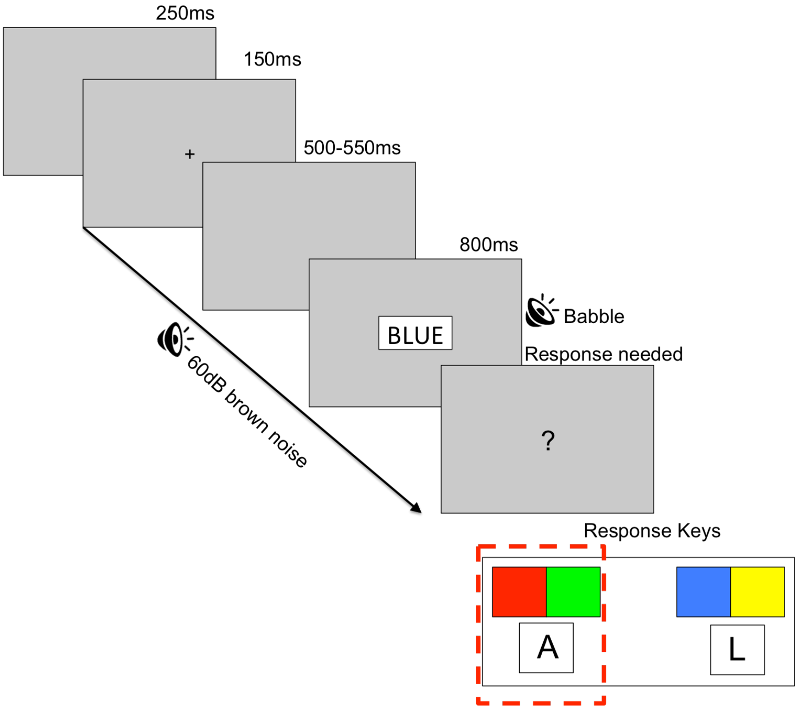


Figure S 1. Trial structure used for response time tasks in Chapter 4. Participants identified the relevant dimension (in this case the written word) using either the A or the L key.

Additional analyses (Chapter 4)

Assessment of dimensional imbalance

Defining dimensional imbalance

In Chapter 4 it was assumed that lowering saturation would decrease the accessibility of relevant colour information relative to irrelevant word information (i.e. *dimensional imbalance*). To assess whether this was the case, two sets of additional analyses were performed. First, I compared the response times for word reading versus colour naming under high versus low saturation conditions. This measure of dimensional imbalance is comparable to that used by Ben-David and Schneider (2010), except I implemented a ratio rather than a difference score:

Equation 15

$$DT^{RT} = \frac{RT\ to\ colour}{RT\ to\ word}$$

Thus, a score of 1 would indicate equal response times for colour naming and word reading, >1 would indicate faster responses for word versus colour naming and <1 faster responses for colour versus word naming.

Second, I compared the accessibility of colour and word information relative to threshold under high versus low saturation conditions. This was based on the detectability of the word relative to the colour based on thresholds:

Equation 16

$$DT^{Thresh} = \frac{Colour\ visibility}{Word\ visibility}$$

Here a score of 1 would indicate colour and word were equally detectible relative to threshold, >1 would indicate the colour was more detectible than the word and <1 the word more detectible than the colour.

In this equation $Word\ visibility$ was calculated as:

Equation 17

$$Word\ visibility = \frac{Mean\ presented\ word\ opacity}{Mean\ threshold\ word\ opacity}$$

Where$Mean\ presented\ word\ opacity$ $Mean\ presented\ word\ opacity$ is the mean opacity at which words were presented in the Stroop task (across the four colour backgrounds) and $Mean\ threshold\ word\ opacity$ is the mean opacity level at which participants could read words 79% of the time in the threshold task (across the four colour backgrounds). Similarly, $Colour\ visibility$ $Colour\ visibility$ was defined as:

Equation 18

$$Colour\ visibility = \frac{(100 - \mathrm{Pr}\,esented\ mask\ opacity)}{(100 - Mean\ threshold\ mask\ opacity)}$$

Where $\mathrm{Pr}\,esented\ mask\ opacity$ is the opacity at which the masker was presented in the Stroop task (60% in low saturation, 0% In high saturation conditions) and $Mean\ threshold\ mask\ opacity$ is the average opacity at which participants could identify colours correctly 79% of the time (averaged across the 4 colours). DI^{Thresh} therefore reflects whether the colour or word was more detectable, relative to threshold, in the Stroop and response time tasks.

Results

Response time ratios (DI^{RT})

Raw response times to written words, spoken words and colours are shown in Table S 4.

		N	**Minimum**	**Maximum**	***M (ms)***	***SE***
Colour	High saturation	32	400	1050	595	26
	Low saturation	27	460	990	598	24
Words	Written	32	470	960	592	20
	Spoken	32	490	970	676	18

Table S 4. Response times (ms) to colours without and with a desaturation mask, written words and spoken words. A software crash with 5 subjects meant a response time for colours with a desaturation mask was not obtained.

Figure S 2 shows DI^{RT} under high and low saturation conditions (and for spoken words relative to high saturation colours). DI^{RT} did not significantly differ between high (*M* = .98, *SE* = .02) and low (*M* = 1.01, *SE* = .02) saturation conditions (*F*(1, 26) = 2.805, *p* = .106, η_p^2 = .10). Furthermore, neither ratio significantly differed from 1 (*t*(31) = .04, *p* = .97 and *t*(26) = .82, *p*

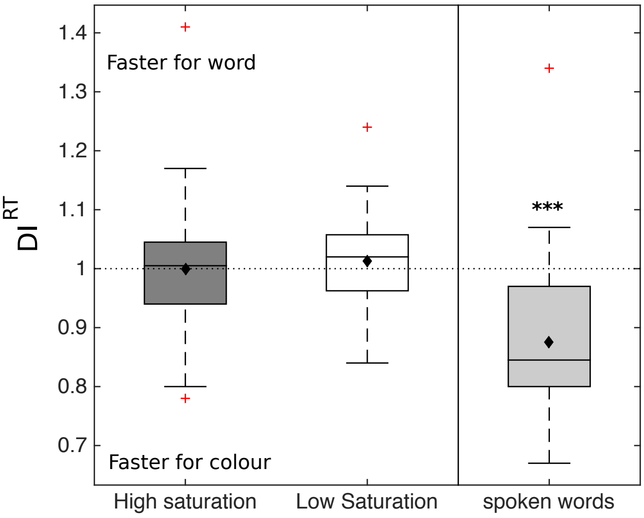


Figure S 2. Response time defined dimensional imbalance measures (D^{RT}) under high (grey) and low (white) saturation conditions. A ratio score greater than 1 indicates response time was faster to the colour compared with the word and score of 1 indicates no difference. Black diamonds indicate means used in analyses. Asterisks indicate t-tests comparing ratio to 1 (* = .05, ** = .01, ***= .001) Bonferroni corrected for 6 comparisons.

= .417 for high and low saturation respectively) suggesting response times were similar between words and colours under both conditions.

As shown in the right hand column of Figure S 2, an exploratory comparison was performed to assess accessibility of spoken word information relative to the colour information. Although this was not the primary aim of Chapter 4 this was done to explore potential differences in dimensional imbalance between unimodal and cross-modal conditions. In short, the result suggests stronger dimensional imbalance, in terms of response time, under cross-modal conditions. There was a significant difference in DI^{RT} between written words (*M* = 1, *SE* = .02) and spoken words (*M* = .88, *SE* = .02) presented with high saturation colours (*F*(1, 31) = 41.708, *p* <.001, η_p^2 = .57). DI^{RT} also significantly differed from 1 for spoken words (*t*(31) = -5.206, *p* <.001). This suggested response times were slower to spoken words relative to colours, whilst response times were similar between written words and colours. Thus, greater dimensional imbalance, in favour of colour versus word information, might provide one explanation of smaller cross-modal Stroop effects in the main body of this thesis.

Threshold ratios (DI^{Thresh})

Figure S 3 shows DI^{Thresh} under high and low saturation conditions. DI^{Thresh}

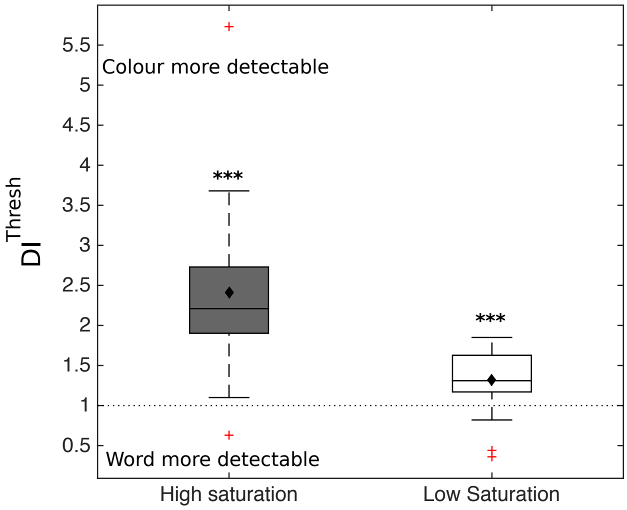


Figure S 3. Threshold defined dimensional imbalance measures (D^{Thresh}) under high (grey) and low (white) saturation conditions. A ratio score greater than 1 indicates the colour was more detectable than the word, a score of 1 indicates no difference. Black diamonds indicate means used in analyses. Asterisks indicate t-tests comparing ratio to 1 (* = .05, ** = .01, ***= .001) Bonferroni corrected for 6 comparisons.

was significantly higher (in favour of colour information) under high saturation (*M* = 2.4, *SE* = .16) conditions compared with low saturation (*M* =1.32, *SE* = .07) conditions (*t*(30) = 7.25, *p* < .001). However, both ratios were significantly higher than 1 (*t*(30) = 8.84, *p* <.001 and *t*(30) = 4.93, *p* <.001, for high and low saturation respectively) suggesting that the colour was presented higher above threshold compared with the word, and this was more so in high saturation conditions.

Analysis of Threshold data

The mean threshold for detecting spoken words in Brown noise was 43.45dB (*SE* = .78). Thresholds derived for visual stimuli are presented in Table S 5.

A 2 (saturation) x4 (colour background) ANOVA indicated a main effect of saturation (*F*(1, 31) = 66.301, *p* < .001, η_p^2 = .68) , colour background, (*F*(2.07, 64.09) = 35.06, *p* < .001, η_p^2 = .53) and an interaction between saturation and colour background (*F*(1.98, 61.24) = 5363.16, *p* < .001, η_p^2 = .66) . Thresholds for reading words were significantly higher with the desaturation mask (*M* = 15.9%, *SE* = 0.83%) compared to without the desaturation mask (*M* = 6.81%, *SE* = 1.29%). Thresholds also significantly differed between all colour backgrounds (red vs. green, *p* = .002; all other significant comparisons *p* <.001) apart from green and yellow (*p* = 1).

	Red		**Green**		**Blue**		**Yellow**	
	M	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***
Colours	95.66	0.1	94.83	.03	94.37	.03	94.18	.08
Words (no mask)	8.16	8.74	6.56	6.34	6.10	8.7	6.51	8.7
Words (+ mask)	15.04	5.68	12.06	3.69	25.7	7.23	9.94	6.2

Table S 5. Mean (*M*) thresholds for identifying colours presented behind a desaturation mask (in which the threshold indicates the % mask opacity at which participants could still identify colours) and thresholds for reading words presented with different coloured backgrounds with and without a desaturation mask set to 60% opacity (here the threshold is the % opacity at which the word needed to be set in order to be identified).

However, the interaction between saturation and colour background showed that the effect of background only occurred under low saturation conditions. Thresholds were all comparably low without the mask.

Analysis of raw Stroop data

Response times

Response times for each condition in Chapter 4 are shown in Table S 6. A 2 (saturation) x 4 (congruency) ANOVA on response times showed no main effect of saturation ($F(1, 30) = .129$, $p = .722$, $\eta_p^2 = .004$) a main effect of congruency ($F(1, 30) = 30.509$, $p < .001$, $\eta_p^2 = .5$) but no interaction between saturation and congruency ($F(1, 30) = .553$, $p = .648$, $\eta_p^2 = .02$). The main effect of congruency occurred because response time was significantly faster for congruent (*M* = 564ms *SE* = 14) vs. all other conditions ($p < .001$ for all comparisons). Response times were also slower for response-incongruent (*M* = 616ms *SE* = 18) vs. stimulus-incongruent (*M* = 585ms, *SE* = 17; $p < .001$) and neutral (*M* = 591ms, *SE* = 16; $p = .005$). However, response times did not significantly differ between stimulus-incongruent and neutral ($p = 1$). For this reason, it was concluded that neutral conditions still produce stimulus interference, and therefore focus on ratios within the main thesis.

Accuracy

Accuracy for each condition in Chapter 4 are shown in Table S 6. A 2 (saturation) x 4 (congruency) ANOVA on accuracy showed a main effect of saturation ($F(1, 30) = 4.917$, $p = .034$, $\eta_p^2 = .14$), a main effect of congruency ($F(1, 30) = 23.336$, $p < .001$, $\eta_p^2 = .44$) and no interaction between saturation and congruency ($F(1, 30) = .406$, $p = .749$, $\eta_p^2 = .01$). The main effect of congruency occurred because accuracy was significantly lower for response-incongruent (*M* = 91.71%, *SE* = .89) vs. congruent (*M* = 95.71%, *SE* = .52; $p < .001$) and stimulus-incongruent (*M* = 96.25% *SE* = .55; $p < .001$) conditions. Accuracy was also lower for response-incongruent vs. neutral (*M* = 93.93, *SE* = .57, $p = .014$). Accuracy did not significantly differ between stimulus-incongruent and congruent conditions (*p* = .05) but was significantly lower for neutral compared with stimulus-incongruent conditions (*p* <.001). There was no significant difference in accuracy between neutral and congruent conditions (*p* = .43).

		Response times						Accuracy					
		High saturation	95% CI		Low saturation	95%CI		High saturation	95% CI		Low saturation	95%CI	
			Lower	Upper		Lower	Upper		Lower	Upper		Lower	Upper
All trials	C	569	531	606	559	529	589	95.46	94.22	96.70	94.94	93.58	96.31
	SI	590	546	634	580	546	614	96.51	95.31	97.70	96.00	94.81	97.19
	RI	619	571	666	614	578	650	92.28	90.19	94.37	91.14	89.49	92.79
	N	592	547	636	591	561	622	94.50	92.97	96.04	93.35	92.20	94.50

Table S 6. RT and accuracy under each saturation conditions (C) stimulus-incongruent (SI) response-incongruent (RI) and neutral (N) conditions under unimodal and cross-modal sensory conditions for Chapter 4. Confidence intervals (CI) represent the 95% bootstrapped confidence intervals.

Analysis of first 96 Stroop trials

In Chapter 4, it was predicted that lowering saturation would increase Stroop interference. This effect was not found. Here I analysed the first 96 trials to assess whether the overall lack of differences could be explained by practice effects, and, therefore, if an effect would be seen in the first 96 trials. In summary, the results below did not show any effect of saturation in the first 96 trials.

Response times

A 2(saturation) x2 (interference type) ANOVA showed no main effect of saturation ($F(1, 30) = .275$, $p = .604$, $\eta_p^2 = .01$), no main effect of interference type ($F(1, 30) = .914$, $p = .347$, $\eta_p^2 = .03$) and no interaction between interference type and sensory condition ($F(1, 30) = .749$, $p = .394$, $\eta_p^2 = .02$). Thus the effect of saturation did not modulate interference type even when only considering the first block of trials.

Bonferroni corrected t-tests comparing response time ratios to 1 showed both stimulus and response-interference ratios under high and low saturation conditions all significantly differed from 1 (high saturation: $p = .013$, $p = .029$; low saturation: $p < .001$ and $p = .025$ respectively).

Accuracy

A 2x2 ANOVA showed no main effect of saturation ($F(1, 30) = 1.304$, $p = .263$, $\eta_p^2 = .04$) but a main effect of interference type ($F(1, 30) = 8.676$, $p = .006$, $\eta_p^2 = .22$) which did not interact with sensory condition ($F(1, 30) = .844$, $p = .366$, $\eta_p^2 = .03$). This finding parallels the analysis across all trials, suggesting accuracy was lowered by response interference but not stimulus interference, and this did not differ between saturation conditions.

Bonferroni corrected t-tests comparing accuracy ratios to 1 showed response-interference but not stimulus-interference ratios under high saturation conditions significantly differed from 1 (p =.028 and $p = .592$ respectively). Conversely under low saturation conditions stimulus interference ratios did not significantly differ from 1 (*p* = .2) whilst response interference ratios did significantly differ from 1 (*p* <.001).

Supplementary analyses for Chapter 5

Analysis of threshold data (Chapter 5)

The average threshold for hearing words embedded in brown noise was 45dB (*SE* = .47).

Average thresholds for reading words on each colour background are shown in Table S 7. A four-way ANOVA showed a significant main effect of colour background (*F*(2.03, 56.8) = 5.402, *p* =.007, η_p^2=.16). This occurred because thresholds were significantly higher for reading words on red vs. blue backgrounds (*p* = .028) and red vs. yellow backgrounds (*p* = .017), however no other differences reached significance.

			95% Confidence Interval	
	M (% opacity)	**SE**	**Lower Bound**	**Upper Bound**
Red	3.95	0.24	3.46	4.44
Green	3.95	0.15	3.64	4.25
Blue	3.54	0.22	3.09	4.00
Yellow	3.98	0.16	3.65	4.30

Table S 7. Average thresholds for reading words on each colour background in Chapter 5 (% opacity). M = mean; SE = standard error of the mean.

Correlations between detectability imbalance and unimodal vs. cross-modal Stroop performance

Within Chapter 5 I attempted to match the detectability of visual and auditory distractors by presenting all stimuli 10x (20db) above threshold. However, due to a calibration error, all auditory stimuli were presented 5dB below expected. This meant that whilst visual stimuli were always presented 10x (20dB) above threshold, auditory stimuli were capped at 60dB. Thus auditory stimuli had to be presented at the capped value for all participants and auditory stimuli were therefore presented 4-14dB above threshold. Given this, it might be expected that visual distractors were more detectable (and therefore more distracting) than auditory distractors. As such it would be expected that unimodal Stroop effects should be higher than cross-modal effects.

To explore whether the imbalance in detectability between visual and auditory distractors influenced unimodal vs. cross-modal Stroop interference supplementary correlational analyses were performed. First, I calculated detectability imbalance between visual and auditory information:

Equation 19

$$Detectability\ imbalance = D^{V} / D^{A}$$

Where D^{V} indicates the detectability of the visual stimulus in the Stroop task relative to threshold (all visual stimuli were successfully presented 10x (20dB) above threshold) and D^{A} indicates the detectability of the auditory stimulus relative to threshold. Given this, a detectability imbalance score indicates how much higher above threshold the visual information was presented relative to the auditory information (i.e. if visual information were presented 10x (20dB) above threshold and the auditory stimulus was presented 10dB above threshold the detectability imbalance score would be 2, showing the increase in visual intensity relative to threshold was 2x the increase in auditory intensity relative to threshold).

Following this, I calculated Stroop imbalance with regards to response time and accuracy through comparing ratios under unimodal vs. cross-modal conditions:

Equation 20

$$Stroop\ Imbalance^{RT} = Unimodal\ General\ Interference^{RT} / Cross\text{-}modal\ General\ Interference^{RT}$$

Equation 21

$$Stroop\ Imbalance^{Acc} = Unimodal\ General\ Interference^{Acc} / Cross\text{-}modal\ General\ Interference^{Acc}$$

Thus, a $Stroop\ Imbalance^{RT}$ score higher than 1 indicates the participant was more distracted by (i.e. was slowed further by) unimodal compared with cross-modal Stroop interference. A $Stroop\ Imbalance^{Acc}$ score lower than 1 indicates the participant was more distracted (i.e. had more accuracy decrements) under unimodal compared with cross-modal Stroop conditions.

The results of this exploratory correlational analysis are shown in Figure S 4. Both relationships were negative and non-significant, suggesting detectability imbalance was not associated with increased unimodal relative to cross-modal interference.

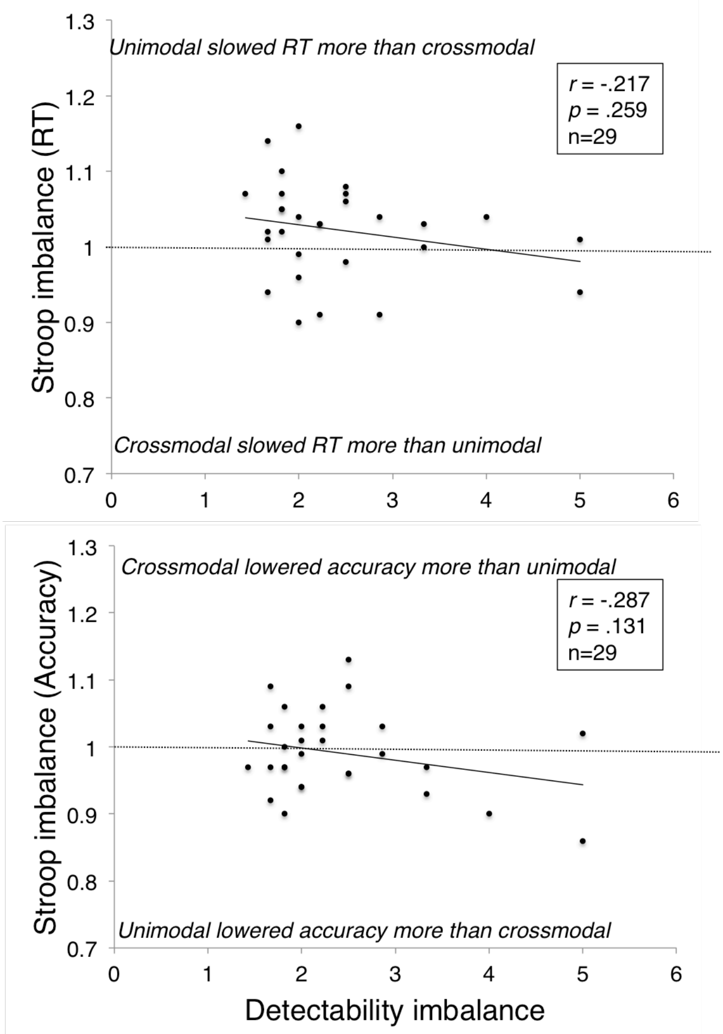


Figure S 4. Correlations between detectability imbalance and Stroop imbalance for response time (a) and accuracy (b) ratios. Stroop imbalance scores higher than 1 for response times show unimodal incongruence slowed response times more than cross-modal incongruence. Stroop imbalance scores for accuracy lower than 1 indicate accuracy was lowered more by unimodal incongruence vs. cross-modal incongruence. Detectability scores indicate how many more times above threshold visual distractors were presented relative to auditory distractors. *r* = Pearsons correlations.

Analysis of raw Stroop data

Response times

Response times for each condition in Chapter 5 are shown in Table S 8. A 2 (saturation) x 4 (congruency) ANOVA on response times showed no main effect of sensory condition (*F*(1, 28) = .026, *p* = .872, η_p^2=.001) a main effect of congruency (*F*(2.157, 60.388) = 8.538, *p* < .001, η_p^2 = .23) and an interaction between saturation and congruency (*F*(2.63, 73.73) = .553, *p* = .042, η_p^2 = .09). The main effect of congruency occurred because response time was significantly faster for congruent (*M* = 541ms, *SE* = 19) vs. response-incongruent (*M* = 568ms, *SE* = 24; *p* = .005). However response times did not differ between congruent and stimulus-incongruent (*M* = 551ms, *SE* = 21, *p* = .171) or neutral conditions (*M* = 553ms, *SE* = 20, *p* = .17). Response times were also significantly faster for stimulus-incongruent vs. response-incongruent (*p* = .004). However, response times did not significantly differ between neutral conditions and stimulus-incongruent (*p* = 1) or response-incongruent (*p* = .073) conditions. Because response times to neutral stimuli did not significantly differ from stimulus-incongruent response times, it was again assumed that neutral trials elicited stimulus interference, and neutral trails were therefore not considered in the main analysis.

The interaction between congruency and sensory condition occurred because congruency effects only reached significance under unimodal and not cross-modal conditions. Under unimodal conditions response times were significantly slower to response-incongruent vs. congruent and stimulus-incongruent (*p* <.001 for both) as well as neutral (*p* = .035) however no comparisons reached significance under cross-modal conditions.

Accuracy

Accuracy scores for each condition are shown in Table S 8. A 2 (sensory condition) x 2 (congruency) ANOVA showed a main effect of sensory condition (*F*(1, 28) = 10.64, *p* = .003, η_p^2=.26), a main effect of congruency (*F*(3, 84) = 11.027, *p* <.001, η_p^2=.28) but no interaction between sensory condition and congruency (*F*(3, 84) = 1.714, *p* = .17, η_p^2 =.06). The main effect of sensory condition occurred because accuracy was higher for cross-modal (*M* = 96.15%, *SE* = .51) vs. unimodal (*M* = 94.91%, *SE* = .6) conditions. The main effect of congruency occurred because accuracy was significantly lower for response-incongruent (*M* = 93.7%, *SE* = .8) vs. congruent (*M* = 96.31%, *SE* = .51, *p* = .002), stimulus-incongruent (*M* = 96.38%, *SE* = .54, *p* = .001) and neutral (*M* = 95.71%, *SE* = .58, *p* = .001) conditions. However there were no significant differences between accuracy for congruent, stimulus-incongruent or neutral conditions.

	Response times						Accuracy					
		95% CI			95%CI			95% CI			95%CI	
	Unimodal	Lower	Upper	Cross-modal	Lower	Upper	Unimodal	Lower	Upper	Cross-modal	Lower	Upper
C	543	496	591	538	497	580	95.79	94.48	97.09	96.84	95.59	98.09
SI	549	499	599	553	504	603	96.31	94.93	97.69	96.46	95.43	97.48
RI	576	520	632	561	507	615	92.67	90.85	94.49	94.73	92.88	96.58
N	552	506	598	553	505	601	94.88	93.47	96.28	96.55	95.27	97.84

Table S 8. RT and accuracy under each saturation conditions (C) stimulus-incongruent (SI) response-incongruent (RI) and neutral (N) conditions under unimodal and cross-modal sensory conditions for chapter(s) 5/6. Confidence intervals (CI) represent the 95% bootstrapped confidence intervals.

Supplementary analyses for Chapter 6

Number of retained trials for stimulus interference and neutral conditions.

To simplify comparisons in Chapter 6 congruent and response-incongruent trials are used for comparison. However, this experiment also included stimulus-incongruent and neutral conditions. As shown in Table S 9, the number of trials retained in these conditions was comparable to those reported in Chapter 6.

		EEG system			
Condition	Congruency	Neuroscan	EPOC	*Z*	Shared epochs
Unimodal	C	63.47 (1.08)	61.63 (1.07)	-3.533***	61.7 (1.05)
	SI	64.03 (.84)	62.1 (1.07)	-3.551***	62.33 (.99)
	RI	58.33 (2.08)	56.73 (2.07)	-3.448***	57.03 (2.05)
	N	62.93 (1.14)	60.93 (1.29)	-3.307***	61.03 (1.27)
Cross-modal	C	66.27 (.8)	64.33 (.94)	-3.535***	64.37 (.93)
	SI	65.8 (.7)	63.77 (.76)	-3.748***	63.8 (.75)
	RI	64 (1.07)	62.53 (1.06)	-3.201***	62.53 (1.06)
	N	65.27 (.87)	63.33 (.99)	-3.305***	63.33 (.99)

p* <= .05, ** *p* <= .01, * *p* <= .001

Table S 9. Mean (standard error) number of accepted and shared epochs for Neuroscan and EPOC in each condition. Seventy-two trials were presented per condition. Wilcoxen Signed Rank Tests (*Z*) were used to test the difference between systems.

Supplementary analyses for Chapter 8

Analysis of congruent trials

Within Chapter 8 of this thesis I focus on the McGurk effect. This entailed an analysis focused on responses given to incongruent McGurk stimuli – in which auditory information (“Ba”) and visual information (“Da”) conflicted. However, stimuli also included 25 congruent “Ba”, 25 congruent “Da” and 25 congruent “Ga” stimuli (each presented in 5 levels of auditory noise and 5 levels of visual noise). Here I provide a supplementary overview of:

a) Responses made to congruent stimuli (regardless of noise).

b) The effect of auditory noise on accuracy in each of these conditions (regardless of visual noise)

c) The effect of visual noise on accuracy in each of these conditions (regardless of auditory noise).

It was expected that increasing auditory and visual noise would reduce accuracy on congruent trials in adults and children.

Overall accuracy for congruent stimuli in adults and children.

A 4 (age group: 3-6-year-olds, 7-9-year-olds, 10-12-year-olds and adults) x 3 (stimulus type: congruent “Ba”, “Ga” and “Da”) x 3 (response type: “Ba”, “Ga”, “Da”/”Tha”) ANOVA showed no main effect of stimulus type ($F(1.26, 148.42) = 1.6$, $p =.204$, $\eta^2 = .01$) and no main effect of age group ($F(3, 118) = 1.93$, $p = .128$, $\eta^2 = .05$) but a main effect of response type ($F(1.67, 197.45) = 270.55$, $p < .001$, $\eta^2 = .68$). This main effect occurred because participants made significantly more “Da”/”Tha” ($M = 47.3$, $SE = .88$) compared with “Ga” ($M = 17.38$, $SE = .74$; $p <.001$) and “Ba” ($M = 35.23$, $SE = .59$; $p < .001$). The proportion of “Ba” responses was also higher than “Ga” ($p <.001$).

There was a significant interaction between response type and age group ($F(5.02, 197.45) = 2.4$, $p = .029$, $\eta^2 = .02$). This occurred because whilst there was a main effect of age for “Da”/”Tha” responses ($p = .027$) and “Ba” responses ($p = .020$) there was no effect of age on “Ga” responses ($p = .565$). However, post-hoc pairwise comparisons did not reveal ant differences between age groups in any type of response.

There was also a significant interaction between stimulus type and response type ($F(2.66, 313.47) = 1110.607$, $p < .001$, $\eta^2 = .89$). This was because on “Da” trials, the proportion of “Da”/”Tha” responses made were significantly higher than other responses ($p < .001$ for both comparisons) and the proportion of “Ga” and “Ba” responses did not significantly differ from one another ($p =.609$). Similarly, on congruent “Ba” trials the proportion of “Ba” responses was significantly higher than other responses ($p <.001$ for both comparisons) however the proportion of “Da”/”Tha” errors was significantly higher than “Ga” errors ($p <.001$). In contrast to this, participants appeared to confuse congruent “Ga” stimuli with “Da”, the proportion of “Da”/”Tha” and “Ga” responses did not significantly differ from one another (*p* = .148) however both of these responses were made significantly higher than “Ba” responses (*p* <.001) for both comparisons) .

Finally, there was a significant three-way interaction between stimulus type, response type and age group (*F(*7.97, 313.47) = 7.01, *p* <= .001 η^2 = .02). As can be seen in Table 14, page 210 in Chapter 8, when presented with congruent “Ga” stimuli there was no effect of age group on the type of response made (*p* = .333), and this was likely due to lower accuracy across groups. However, when presented with congruent “Ba” stimuli 3-6-year-olds were significantly more likely to make “Da”/”Tha” (*p* = .01) or “Ga” (*p* =.029) errors compared with adults (although it should be emphasised that accuracy was high despite errors). Errors in 3-6-year-olds did not significantly differ from other child groups however, and there was no difference between 7-9-year-olds, 10-12-year-olds and adults. When presented with congruent “Da” stimuli 3-6-year-olds also presented more errors, making significantly fewer correct “Da”/”Tha” responses compared with the 10-12-year-olds (*p* = .001) and adults (*p* <.001). 7-9-year-olds also made significantly fewer correct “Da”/”Tha” responses compared with adults (*p* = .002), 3-6-year-olds and 7-9-year-olds did not significantly differ from one another in this aspect. Compared with adults, 3-6-year-olds made significantly more “Ga” (*p* = .01) and “Ba” (*p* < .001) errors, and made more “Ba” errors compared with the 10-12-year-olds (*p* =.002).

The effect of auditory noise on accuracy for congruent stimuli in adults and children

A 4 (age group: 3-6-year-olds, 7-9-year-olds, 10-12-year-olds and adults) x 3 (stimulus type: congruent “Ba”, “Ga” and “Da”) x 5 (auditory noise level: no noise, -2 SNR, -8 SNR, -14 SNR, -20 SNR) ANOVA showed a main effect of stimulus type (*F*(1.70, 200.92) = 371.57, *p* <.001, η^2 =.76), a main effect of auditory noise level (*F*(4, 472) = 221.42, *p* <.001, η^2 =.64) and a significant effect of age group (*F*(3, 118) = 10.54, *p* <.001, η^2 =.21). In line with the above analyses, the main effect of stimulus occurred because accuracy was significantly higher for congruent “Ba” and “Da” stimuli compared with “Ga” ($p < .001$ for both comparisons) and accuracy did not significantly differ between “Ba” and “Da” ($p = 1$). The main effect of age group occurred because accuracy gradually increased with age (3-6-year-olds: $M = 62.07\%$, $SE = 1.6$, 7-9-year-olds: $M = 69.75\%$, $SE = 1.53$, 10-12-year-olds: $M = 70.39\%$, $SE = 1.6$, adults: $M = 74.29\%$, $SE = 1.53$). Accuracy was significantly lower in 3-6-year-olds compared with 7-9-year-olds ($p = .004$), 10-12-year-olds ($p = .002$) and adults ($p < .001$). Comparisons between other groups did not reach significance. There was no significant interaction between stimulus type and age group ($F(5.11, 200.92) = .943$, $p =.465$, $\eta^2 = .01$).

The main effect of auditory noise occurred because accuracy progressively decreased as auditory noise increased (no noise: $M = 91.28\%$, $SE = 1.02$; -2 SNR: $M = 70.9\%$, $SE =1.15$; -8 SNR: $M = 66.68\%$, $SE = 1.26$; -14 SNR: $M = 64.66\%$, $SE = 1.31$; -20 SNR: $M = 52.11\%$, $SE = 1.03$). Accuracy was significantly higher with no noise compared to all noise levels ($p < .001$ for all comparisons). Accuracy was also higher at -2 SNR compared with -8 SNR ($p = .015$) and higher levels ($p < .001$ for both). Accuracy did not significantly differ between -8 and -14 SNR ($p = 1$), but at -14 SNR accuracy was higher than – 20 SNR ($p < .001$).

The effect of auditory noise level significantly interacted with both age group ($F(12, 472) = 2.49$, $p = .004$, $\eta^2 = .02$) and stimulus type ($F(6.6, 778.21) = 70.66$, $p < .001$, $\eta^2 = .36$). The interaction between auditory noise level and age group occurred because age differences were significant in no noise and at the highest two noise levels ($p = .001$ and $p < .001$ respectively), however, age differences did not reach significance at -14 SNR ($p = .117$) and reached marginal significance at -8 SNR ($p = .045$). The interaction between auditory noise level and stimulus type arose due to different effects of noise on the pattern of accuracy for each syllable. As shown in Figure S 5, accuracy for “Ba” progressively decreased with each stage of auditory noise but then accuracy increased at the highest noise level. Accuracy for “Ga”

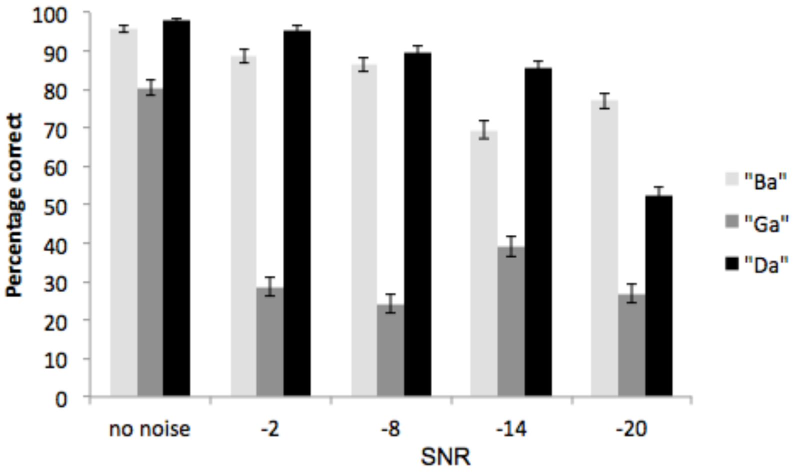
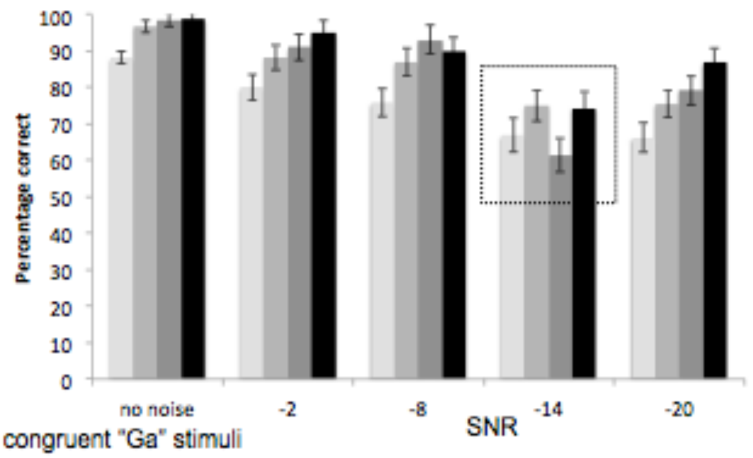


Figure S 5. Percentage correct (across all age groups) for congruent stimuli at each noise level.

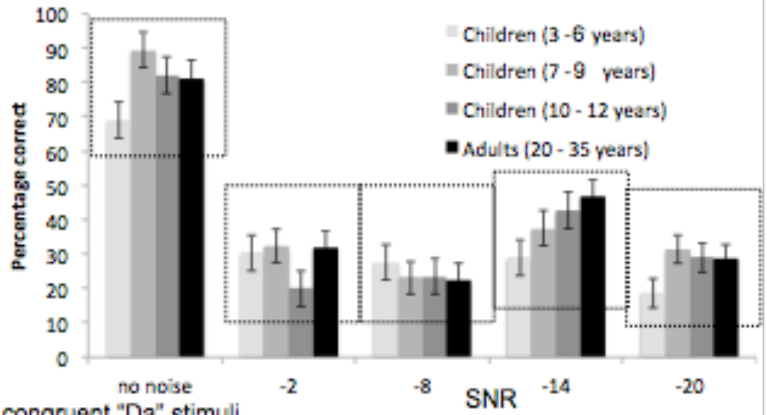
significantly decreased from no noise to -2 and -8 SNR, then increased at -14 SNR, then, again, decreased at -20 SNR. For “Da” accuracy progressively decreased at each increment in noise.

Finally, there was a three-way interaction between auditory noise level, stimulus type and age group (*F*(19.79, 778.21) = 2.5, $p < .001$, $\eta^2 = .04$). As shown in Figure S 6 this occurred because the main effect of age did not reach significance at all levels of noise in each syllable (non significant effects are indicated with dashed boxes). At levels in which the effect of age group reached significance there was a consistent pattern in which accuracy was lower in younger child groups compared with older children and adults.

a) congruent "Ba" stimuli



b) congruent "Ga" stimuli



c) congruent "Da" stimuli

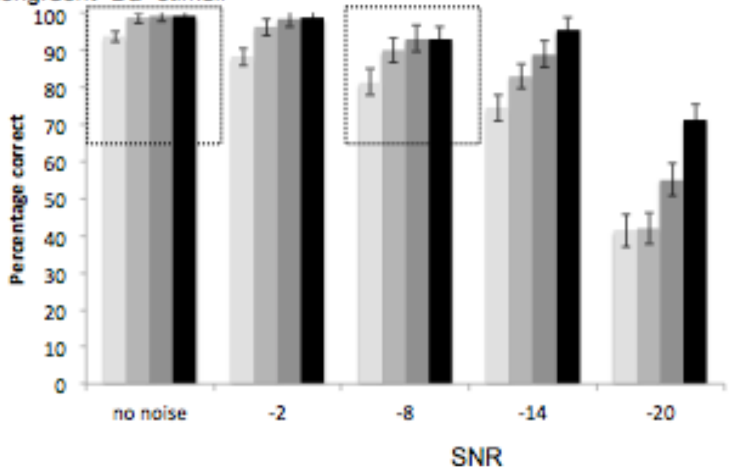


Figure S 6. Accuracy for congruent stimuli at each auditory noise level in each age group. Dashed boxes indicate levels of noise in each stimulus at which the effect of age was non significant.

The effect of visual noise on accuracy for congruent stimuli in adults and children

A 4 (age group: 3-6-year-olds, 7-9-year-olds, 10-12-year-olds and adults) x 3 (stimulus type: congruent "Ba", "Ga" and "Da") x 5 (visual noise level: no noise, 30% blur, -40% blur, 50% blur, 60% blur) ANOVA showed a main effect of stimulus type ($F(1.701, 200.92) = 371.57$, $p < .001$, $\eta^2 = .75$), a main effect of visual noise ($F(4, 472) = 6.88$, $p < .001$,η^2 = $\eta^2 = .05$) and a main effect of age group ($F(3, 118) = 10.54$, $p < .001$, $\eta^2 = .21$). The main effect of age and stimulus type were identical to those reported above.

The main effect of visual noise occurred because accuracy was significantly higher with no noise compared with 50% ($p < .001$) and 60% blur ($p = .003$). Accuracy was also higher at 30% blur compared with 50% blur ($p = .014$) – all other comparisons did not reach significance. Notably, there was no interaction between age group and visual noise ($F(12, 472) = 1.76$, $p = .052$, $\eta^2 =.04$) suggesting similar effects across age groups.

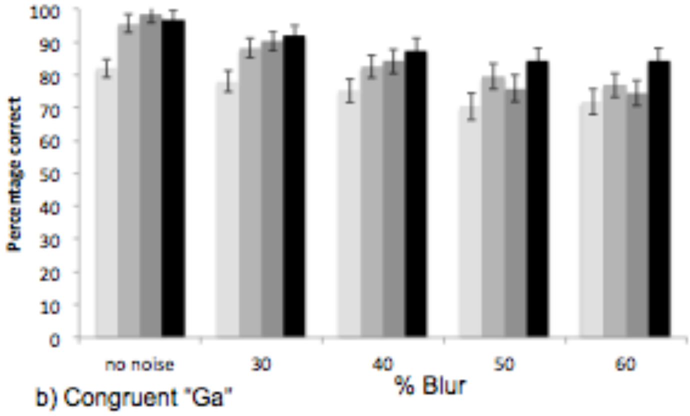
The effect of visual noise significantly interacted with stimulus type ($F(8, 944) = 9.672$, $p < .001$, $\eta^2 =.07$). This interaction occurred because the main effect of visual noise reached significance for “Ba” stimuli ($p < .001$) but not “Ga” ($p = .264$) or “Da” ($p = .539$) stimuli. As can be seen in Figure S 7 visual noise gradually reduced accuracy for congruent “Ba” stimuli. There was no significant three-way interaction between age group, stimulus type and visual noise ($F(20.52, 807.24) = 1.32$, $p = .137$, $\eta^2 =.03$)

Key points from analysis of congruent trials (Chapter 8)

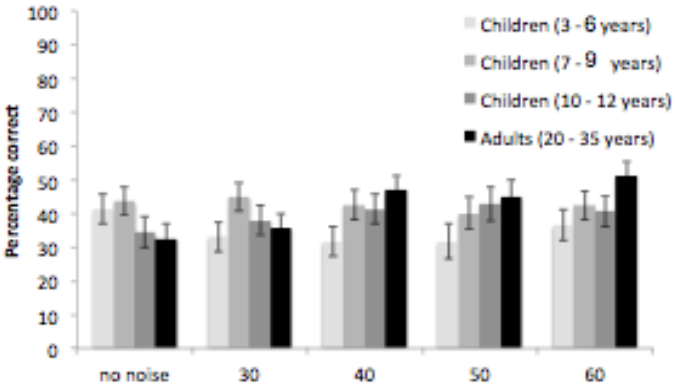
A few points of interest are notable from these supplementary analyses of responses to congruent trials.

1. Accuracy on congruent trials was always lower for “Ga” compared with “Ba” and “Da” trials.
2. Age differences were greatest in higher auditory noise levels.

a) Congruent "Ba"



b) Congruent "Ga"



c) Congruent "Da"

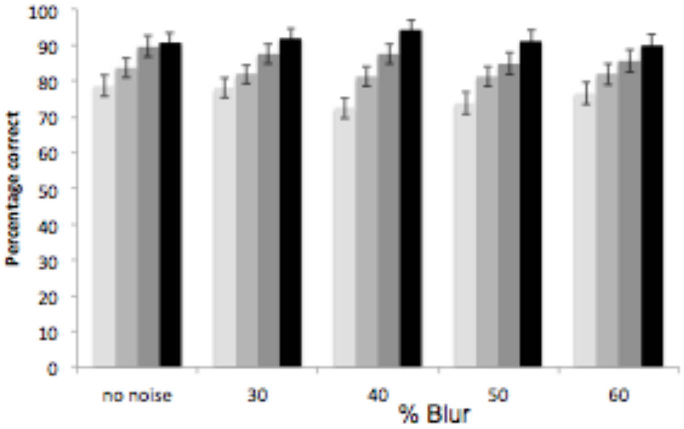


Figure S 7. Accuracy for congruent stimuli at each visual noise level in each age group.

With regards to point one, accuracy was lower for congruent “Ga” stimuli compared with the other two syllables and participants were equally likely to respond “Da” or “Ga” to these stimuli. This may have meant that for the main analyses of incongruent McGurk stimuli (auditory “Ba” visual “Ga”) participants would be biased towards correct auditory responses. Given this the interpretation of “fusion” vs. “visual” responses must be approached with caution, as if participants were likely to confuse these stimuli fusion responses may reflect participants responding to what they believe the visual stimulus represents, making it similar to a “visual” response.

With regards to point two, this suggest that children were more influenced by the effect of noise and found it more difficult to identify the correct, congruent, signal in high auditory noise compared with adults. This finding appears contrary to the analysis of McGurk responses, where children’s veridical perception of sound was less influenced by noise compared with adults (who were more easily swayed towards the influence of vision). One explanation of this may be that children struggle to integrate auditory and visual noise in levels of high auditory noise (Barutchu et al., 2010) and this results in reduced benefit for congruent speech signals. Following this, when auditory and visual signals are in conflict, but also in high levels of auditory noise, reduced integration may result in reduced McGurk perception in children.

Details regarding participants excluded from threshold analyses of Chapter 8

In Chapter 8, participants were only included in threshold analyses if their estimated threshold occurred within the range of noise presented. This resulted in 75 (of 90) child and 21 (of 32) adult data sets available for the analysis of auditory noise upon McGurk responses and 73 child and 24 adult data sets available for the analysis of visual noise upon the McGurk effect. Due to the large number of exclusions in this analysis, it is important to illustrate that these exclusions did not bias my main analyses.

Figure S 8 shows the distribution of estimated thresholds for child (left) and adult (right) participants excluded from analysis of auditory (top) and visual (bottom) analyses. These estimated thresholds must be treated with caution, however, the pattern of results is in line with the effects reported in Chapter 8. Inclusion of these participants would have inflated the effects reported in Chapter 8.

In terms of excluded adult participants, these subjects visual thresholds were shifted towards increased noise levels, indicating that these participants would have shown McGurk perception even in high levels of visual noise. In terms of auditory noise, excluded adults participants all had thresholds requiring a lower SNR (favouring signal over noise). Thus, these participants would have shown McGurk responses even in lower levels of auditory noise. Critically these data points are in line with the reported main analyses and removal would not have altered the direction of results.

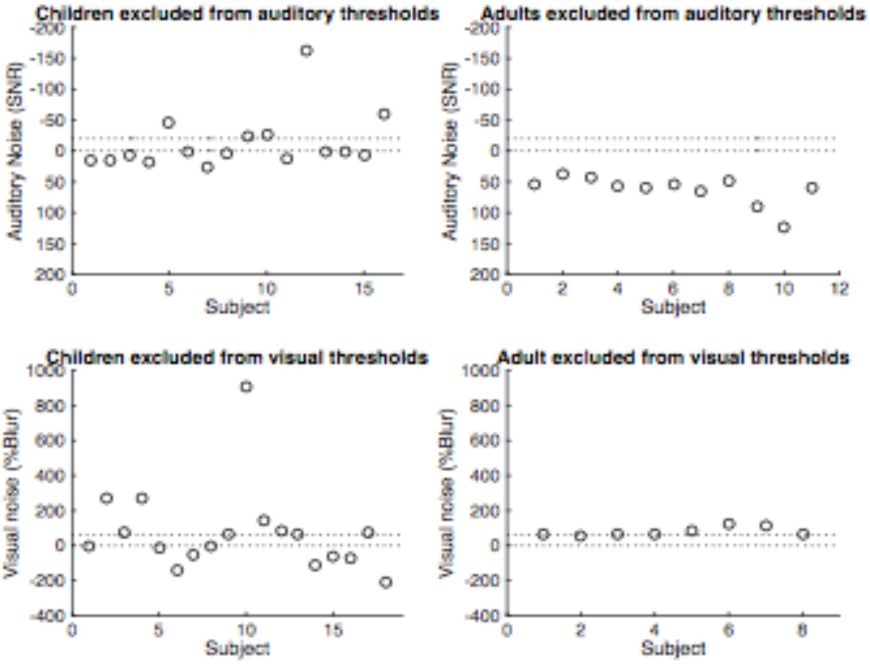


Figure S 8. Thresholds of participants who were excluded from analysis of auditory thresholds (upper) and visual thresholds (lower). Participants were excluded as their thresholds fell outside the range of presented intensities (shown with dashed lines).

Supplementary analyses for Chapter 9

Analysis of raw Stroop response times and accuracy

Response times

There was no main effect of sensory condition (*F* (1, 24) = .028, *p* = .87, $\eta_p^2 <$ .01) and no interaction between sensory condition and congruency (*F* (2.16, 51.89) = .94, *p* = .425, $\eta_p^2 = .04$) but there was a main effect of congruency (*F* (1, 24) = 11.39, *p* <.001, $\eta_p^2 = .32$). The main effect of congruency occurred because response times were significantly faster to congruent stimuli (*M* = 565ms, *SE* = 21) compared with stimulus-incongruent (*M* = 609ms, *SE* = 29; *p* = .005) and response-incongruent (*M* = 611ms, *SE* = 26; *p* <.001) but not neutral conditions (*M* = 578ms, *SE* = 25; *p* = .65). Response times were not significantly different between stimulus- and response-incongruent stimuli (*p* = 1). Response times to neutral stimuli were significantly faster than both stimulus-incongruent (*p* = .032) and response-incongruent (*p* = .004). However, to maintain consistency with previous analyses these trials were not included in the main analysis.

Accuracy

There was no main effect of sensory condition (*F* (1, 24) = .7, *p* = .412, η_p^2 = .03) and no interaction between sensory condition and congruency (*F* (3, 72) = 1.49, *p* = .225, η_p^2 = .06) but there was a main effect of congruency (*F* (1, 24) = 4.88, *p* = .004, η_p^2 = .17). This effect occurred because accuracy was significantly lower for response-incongruent vs. stimulus-incongruent conditions (*p* = .012). However all other comparisons did not reach significance.

	Response times						**Accuracy**					
		95% CI			**95%CI**		**Unimodal**	**95% CI**		**Cross-modal**	**95%CI**	
	Unimodal	**Lower**	**Upper**	**Cross-modal**	**Lower**	**Upper**		**lower**	**upper**		**lower**	**upper**
C	564	513	615	567	525	609	94.33	92.36	96.31	93.16	90.95	95.38
SI	604	549	660	613	536	690	95.17	93.20	97.13	93.67	90.77	96.57
RI	620	556	684	602	545	659	91.50	89.37	93.63	89.83	86.18	93.49
N	582	534	630	575	515	634	91.33	88.81	93.86	93.33	90.90	95.77

Table S 10. RT and accuracy under each saturation conditions (C) stimulus-incongruent (SI) response-incongruent (RI) and neutral (N) conditions under unimodal and cross-modal sensory conditions for Chapter 9. Confidence intervals (CI) represent the 95% bootstrapped confidence intervals.

Exploratory analysis of the effect of first language

Sixteen participants in Chapter 9 reported English as a second language. To ensure this did not influence Stroop effects a 2 (Sensory condition: unimodal vs. cross-modal) x 2 (Interference type: stimulus vs. response interference) x 2 (English first language vs. English not first language) mixed ANOVA was performed. As shown in Table S 11, the results of this ANOVA did not show first language to interact with Stroop effects. It was therefore concluded that English as a first language did not influence Stroop effects in this Experiment.

	F	***p***	η_p^2
Response times			
Sensory condition	3.70	0.07	0.14
Sensory condition x First language	1.18	0.29	0.05
Interference type	4.55	0.04	0.17
Interference type x First language	1.82	0.19	0.07
Sensory condition x Interference type	0.11	0.75	0.01
Sensory Condition x Interference type x First language	2.09	0.16	0.08
Accuracy			
Sensory condition	0.02	0.9	<.01
Sensory condition x First language	0.17	0.68	0.01
Interference type	6.23	0.02	0.21
Interference type x First language	1.07	0.31	0.05
Sensory condition x Interference type	0.02	0.90	<.01
Sensory Condition x Interference type x First language	4.05	0.06	0.15

Table S 11. Statistics resulting from 2 (Sensory condition: unimodal vs. cross-modal) x 2 (Interference type: stimulus vs. response interference) x 2 (English first language vs. English not first language) ANOVA. English first language did not interact with any effects.