Review article

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

Rebecca J. Hirst*, Lucy Cragg, Harriet A. Allen

University of Nottingham, United Kingdom

A B S T R A C T

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. We 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.

1. Introduction

Our world is perceived through multiple senses, but it is unclear whether information from all senses is treated equally. Whilst reading this paper, 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), Colavita et al. (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 and Weisberg, 1979; Johnson and Shapiro, 1989; Shapiro et al., 1984). These studies found Colavita "errors" to occur on a relatively large number of bimodal trials ranging from 38 to 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 and Spence, 2007a,b,c,d) 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 et al., 2012) have concluded the Colavita effect to be relatively insensitive to manipulations of stimulus intensity (Colavita, 1974; Shapiro and Johnson, 1987), attention bias to one or other modality created by the experiment (Egeth and Sager, 1977; Koppen and Spence, 2007a,c; Sinnett et al., 2007), response demands (Egeth and Sager, 1977; Hecht and Reiner, 2009; Koppen and Spence, 2007c; Sinnett et al., 2007) and stimulus complexity (Koppen et al., 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 et al. (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 and Browne, 2008a) whereas the visual cortex undergoes lengthy, protracted development throughout childhood (Graven and 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 and Brooks, 2003) and children manifest smaller, sometimes reverse, Colavita effects (Nava and Pavani, 2013; Wille and 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 et al., 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 and Colonius, 2004; Forster et al., 2002; Gondan et al., 2005; Sinnett et al., 2008). Furthermore, detection thresholds for luminance appear lower (Frassinetti et al., 2002), and the saliency (Noesselt et al., 2008) and perceived brightness (Odgaard et al., 2003) of visual events increases with simultaneous sound. Similarly, irrelevant visual stimuli can enhance auditory detection (Lovelace et al., 2003) and increase the perceived loudness of simultaneously presented sounds (Odgaard et al., 2004). However, Odgaard et al. (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 and 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 versus asymmetrical inhibition and facilitation between vision and audition, we aimed to test this within the existing Colavita literature. The hypothesis of Sinnett et al. (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 et al. note that with multiple response keys slowing can be observed. As such, we assessed whether asymmetrical response time effects are observed within the wider Colavita literature, in which multiple response keys were sometimes used.

The current paper provides the first quantitative synthesis of literature exploring the Colavita effect. The primary objectives of this analysis 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 versus asymmetrical facilitation between audition and vision. Given the specific predictions provided by Sinnett et al. with regards to auditory versus visual modalities, and the audio-visual nature of the Colavita effect in original reports (Colavita, 1974), we focus on studies comparing auditory versus visual modalities. Nevertheless it should be noted that the Colavita effect has since been extended to the visual-tactile domain (Hartcher-O'Brien et al., 2010; Hecht and Reiner, 2009; Occelli et al., 2010). By including data from multiple studies we can overcome 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) we included variables highlighted by Spence (2009) as potential moderator variables. Specifically, we 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[1]).
- 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.

Furthermore, we extend the comparisons to include:

---

[1] Koppen et al. (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.
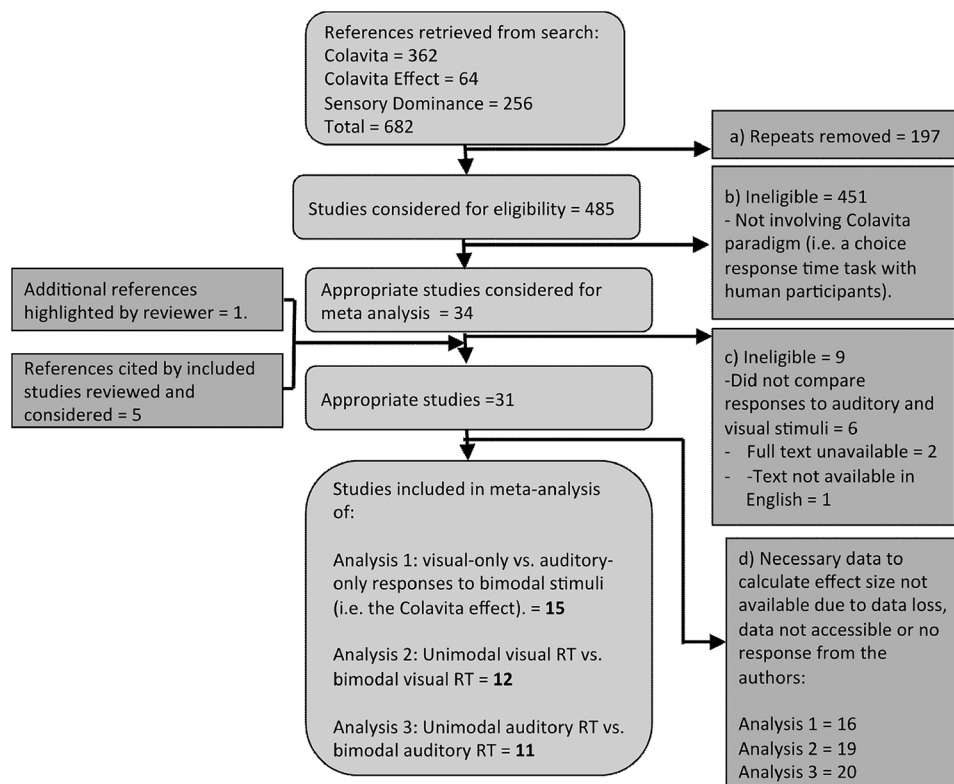
Fig. 1. Flow diagram illustrating search stratery and exclusion criteria used to isolate studies to be in cluded in the meta-analysis.

- Age group: child vs. adult. We predicted a reduced Colavita effect in children.
- Asymmetric facilitation and inhibition. We included studies using Colavita tasks that 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.

2. Method

2.1. Search and inclusion criteria

Studies were retrieved and selected using the guidelines outlined in PRISMA (Moher et al., 2009). Fig. 1 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 (Fig. 1; box b).
- Studies comparing responses to auditory, visual and audio-visual targets (Fig. 1; box c).
- Studies available to the author in English (Fig. 1; box c).
- Sources in which full text could be sourced (i.e. meeting abstracts and posters excluded - Fig. 1; 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 - Fig. 1; 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 and visual targets in the presence of auditory stimuli and/or unimodal auditory

targets, and auditory targets in the presence of visual stimuli.
- 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 and Steeves, 2012, 2013).

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 our 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 1.

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 Eq. (1), 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 we use ratio scores 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" versus 2% "auditory only").

$$Colavita\ effect = \frac{Vb}{Ab} \tag{1}$$

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).

2.2. Statistical analyses

Effect sizes were calculated for the percentage visual-only vs.

Table 1

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.

Author	Exp.	n	Age Group	Age M	Stimuli Type	Stimuli Matched	Stimuli Ratio	Keys (n)	Instructions on bimodal trials	Attentional manipulation	Congruency	Colavita	Unimodal vs. Bimodal RT Visual	Unimodal vs. Bimodal RT Auditory
Osborn et al. (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	43:43:14		Press the tone key	A (respond with tone key)	–	×	✓	✓
Colavita et al. (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 and 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 and 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 tone	A (respond with tone key)	–	×	×	×
	3.1	10					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				N (light half subjective intensity)			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)	–	×	×	✓
	6.1	12					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 et al. (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	–	–	×	×	×

Table 1 (continued)

Author	Exp.	n	Age Group	Age M	Stimuli Type	Stimuli Matched	Ratio	Keys (n)	Instructions on bimodal trials	Attentional manipulation	Congruency	Colavita	Unimodal vs. Bimodal RT Visual	Auditory
	2.1	16					45:45:10		Respond to the stimulus first perceived-informed that shock would occur if too fast or slow.	–	–	x	x	x
	2.2	16							Respond to the stimulus first perceived	–	–	x	x	x
	3.1	16							Respond to the stimulus first perceived-informed that shock would occur if too fast or slow. Participants also received additional tactile stimulus with shock.	–	–	x	x	x
Johnson and Shapiro (1989)	3.2	16							Respond to the stimulus first perceived	–	–	x	x	x
	1.1	12	Adult	–	Simple	N	40:40:20	2	Respond to the stimulus you detect first- Visual stimulus appeared in random location.	–	–	x	x	x
	1.2	10							Respond to the stimulus you detect first- Visual stimulus appeared in same location on all trials.	–	–	x	x	x
	1.3	12							Respond to the stimulus you detect first- Subjects received random, infrequent shocks following trials. Visual stimulus appeared in random location.	–	–	x	x	x
	1.4	10							Respond to the stimulus you detect first- subjects received random, infrequent shocks following trials Visual stimulus appeared in same location.	–	–	x	x	x
Koppen and 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			40:40:20	2	Press both keys	Total - visual cues and auditory cues	–	x	x	x
	4.2									V (visual cue)	–	✓	✓	✓
	4.3									A (auditory cue)	–	✓	✓	✓
Sinnett et al. (2007)	1	24	Adult	–	Complex	N	40:40:20	3	Press a separate key	–	–	x	x	x
	2	54								–	–	x	x	x
	4	24					33:33:33			–	–	x	x	x
	5.1	18					60:20:20			V (more visual targets)	–	x	x	x
	5.2	18					20:60:20			A (more auditory targets)	–	x	x	x
	6.1	18					40:40:20			A (low auditory load)	–	x	x	x
	6.2	18								V (low visual load)	–	x	x	x
Koppen and 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 and 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 position (26°))	✓	✓	✓

Table 1 (*continued*)

Author	Exp.	n	Age Group	Age M	Stimuli Type	Stimuli Matched	Ratio	Keys (n)	Instructions on bimodal trials	Attentional manipulation	Congruency	Colavita	Unimodal vs. Bimodal RT Visual	Auditory
Koppen and 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 et al. (2008)	1.1	12	Adult	21	Complex	N	40:40:20	2	Press both keys	–	C	✓	✓	✓
	1.2	30		21						–	I	✓	✓	×
	2.1	30		21				3	Press separate key	–	C	✓	×	×
	2.2	15		23	Complex					–	I	✓	×	×
	3.1									–	C	✓	×	×
	3.2			–						–	I	✓	×	×
Sinnett et al. (2008)	1.1	22	Adult	–	Complex	N	40:40:20	3	Press separate key	–	–	✓	×	×
	1.2	20						1	Press single key to all targets	–	–	✓	×	×
	2	22		20					Press single key to all targets	–	–	✓	×	×
Koppen et al. (2009)	1	22	Adult	20	Simple	Y	25:25:25:25(no target)	2	Press both keys	–	–	×	×	×
Hecht and Reiner (2009)	1	12	Adult	24.6	Simple	N	40:40:20	3	Press separate key	–	–	×	×	×
Van Damme et al. (2009)	1.1	20	Adult	19.5	Simple	N	40:40:20	3	Press separate key	Visual threat baseline	–	✓	×	×
	1.2									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 et al. (2010)	1	24	Adult	28	Complex	N	40:40:20	3	Press separate key- detect immediate repetition	–	–	✓	×	×
Ngo et al. (2011)	1	20	Adult	28	Complex	N	40:40:20	3	Press separate key- respond to repeated (n-1) targets	–	–	✓	×	×
	2	18		27						–	–	✓	×	×
	3	16		24						–	–	✓	×	×
	4	24		25					Press separate key- detect immediate repetition	–	–	✓	×	×
Moro and 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 and 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	–	–	✓	×	×

Table 1 (*continued*)

Author	Exp.	n	Age Group	Age M	Stimuli Type	Stimuli Matched	Ratio	Keys (n)	Instructions on bimodal trials	Attentional manipulation	Congruency	Colavita	Unimodal vs. Bimodal RT Visual	Unimodal vs. Bimodal RT Auditory
Stubblefield et al. (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	–	✗	✗	✓
	4.1	31		20					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 and 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 and Desmarais (2014)	1	31	Adult	19	Complex	N	33:33:33	3	Press separate Key	–	–	✗	✗	✗
Yue et al. (2015)	1.1	28	Adult	22.7	Simple	N	40:40:20	2	Press both keys	–	C (both near)	✓	✓	✓
	1.2	20		22.6						–	C (both far)	✓	✓	✓
	2.1									V	I (auditory far - visual near)	✓	✓	✓
	2.2									A	I (Visual far - auditory near)	✓	✓	✓

Table 1 (*continued*)

Author	Exp.	n	Age Group	Age M	Stimuli Type	Stimuli Matched	Ratio	Keys (n)	Instructions on bimodal trials	Attentional manipulation	Congruency	Colavita	Unimodal vs. Bimodal RT Visual	Unimodal vs. Bimodal RT Auditory
Wille and 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	✓	.	.
	2.2.1	17		9.06						–	Total	✓	.	.
	2.2.2									–	I	✓	.	.
	2.2.3									–	C	✓	.	.
	2.3.1	23	Adult	27.61	Complex					–	Total	✓	.	.
	2.3.2									–	I	✓	.	.
	2.3.3									–	C	✓	.	.
	3.1.1	18	Child	6.39	Complex					–	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	✓	.	.
	3.3.3									–	C	✓	.	.
Monem and 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 and 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	–	–	✓	✓	✓

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[2] 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 and 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.

2.2.1. Outliers

In line with the guidelines outlined by Viechtbauer (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) Cooks distance exceeded $X^2_{p, 0.5}$, indicating the mahalanobis distance between studies to be decreased following the deletion of ith study.
c) The study was shown to have considerable leverage upon the fit of the model based upon a hat value larger than $3(p/k)$.

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

2.2.2. 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 and Calin-Jagerman, 2017; Lakens, 2013)[3].

$$Cohen's\ d_{av} = \frac{M_{diff}}{\sqrt{\left(\frac{SD_1^2 + SD_2^2}{2}\right)}}$$

(2)

We acknowledge that this is not the optimal measure of effect size for studying within-subject phenomenon. 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 et al., 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 and 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) approximate method for the calculation of confidence intervals (Eq. (3)), and subsequently variance (Eq. (4)), for Cohen's d_{av}. This method still requires knowledge of r.[4]

$$Cohen's\ d_{av} CI = d_{av} \pm t_{\left(1-\frac{\alpha}{2}, n-1\right)}\sqrt{\frac{(2(SD_1^2 + SD_2^2 - 2r)}{n(SD_1^2 + SD_2^2)}}$$

(3)

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

(4)

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 we utilised a method adapted from the calculation of variance for Cohen's d for independent samples (Eq. (5)) where n_1 and n_2 signify the number of observations contributing towards M_{diff}.

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

(5)

Note this is a conservative method yielding marginally wider confidence intervals, relative to Algina and Keselman's (2003) approximate method (Eqs. (3) and (4)), and thus assuming slightly greater variance. Where possible, we also calculated Vd_{av} using Eq. (4) 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 our method proved to be more conservative; the mean variance was 0.114 (SD = 0.05) when calculated using Eq. (5) vs. 0.073 (SD = 0.03) when calculated using the approximate method outlined in Eq. (4) with knowledge of r.

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 Eq. (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).

$$Hedges\ g_{av} = d_{av} * \left(1 - \left(\frac{3}{4 * n - 1} - 1\right)\right)$$

(6)

To summarise, Hedges g_{av} (Eq. (6)) was used as the effect size measure within our analysis. The variance of g_{av} was calculated using Eq. (5), in which d_{av} was substituted with g_{av}.[5]

2.2.3. Moderator variables

Given the range of contexts in which the Colavita effect has been explored the studies included in our meta-analyses were heterogeneous in terms of the methods used. As such we explored the following 8 factors 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,

---

[2] Script available at https://osf.io/d7b3d/.
[3] 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 and Wong, 1992).

[4] Notation used by Algina and Keselman (2003) changed to be in line with current notation.
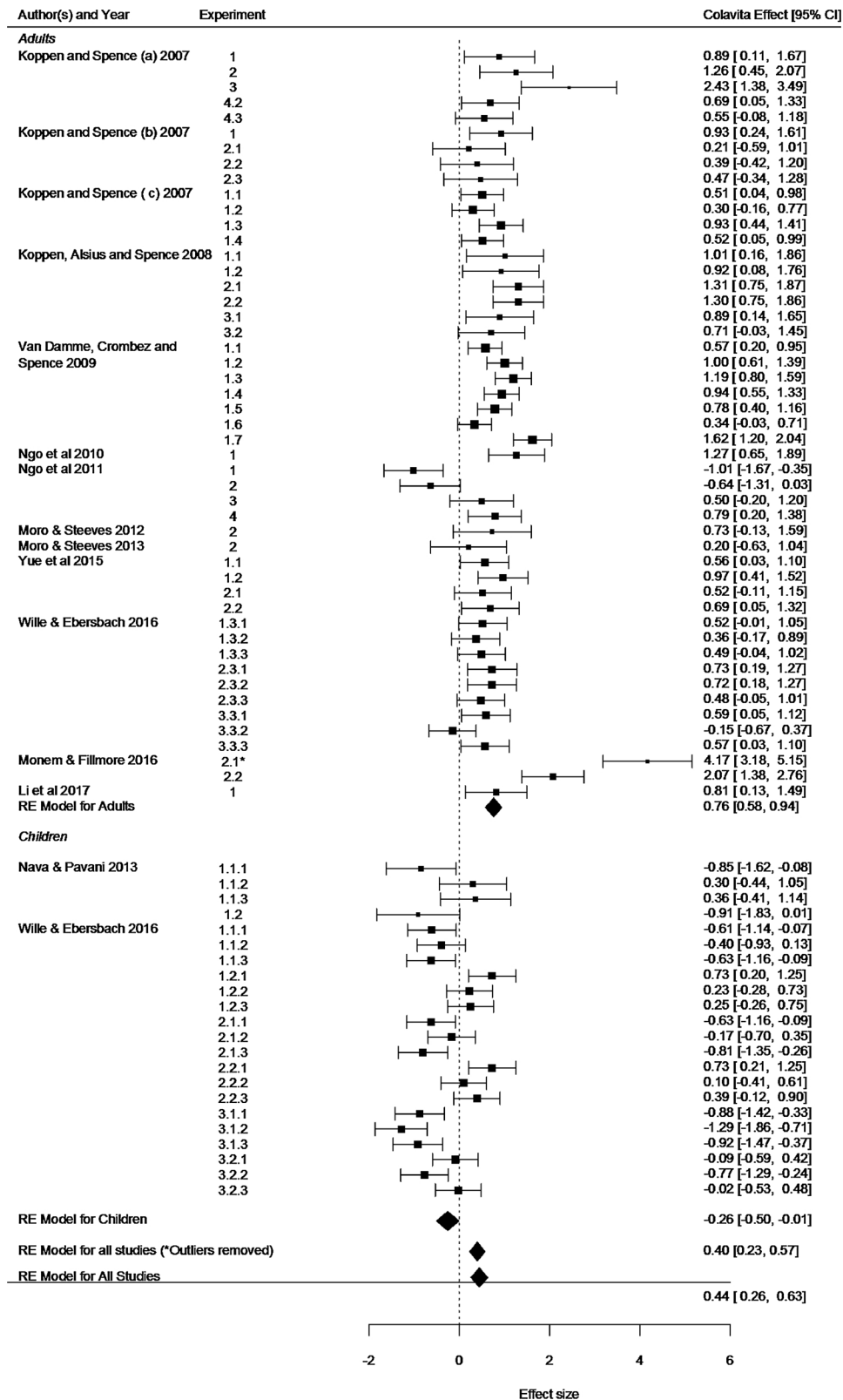[5] Spreadsheet allowing replication of effect size calculation available at https://osf.io/d7b3d/.

Fig. 2. 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.

e.g. 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. 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).

3. Results

3.1. Error data analyses: the Colavita effect

Fig. 2 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 and 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 2). 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.

3.1.1. Effect of age group

A further model was fitted to directly compare the effect sizes of

Table 2

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.

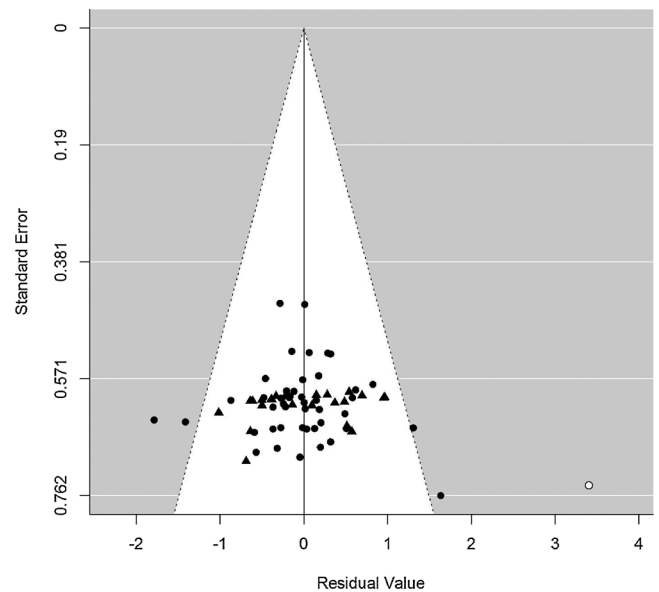| Factor	df	QM	p
Ratio	5	2.8568	0.7220
Response keys	1	.0078	.9297
Stimulus category	1	.2429	.6176
Congruency	2	.2868	.8664
Attentional manipulation	2	.0853	.9583
Age group	1	23.6089	< .0001

Fig. 3. 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.

studies using adult and child participants (regardless of other factors). For details of studies included in this comparison see Table 1, column 4 labelled Age group. Unlike the model described above, here we included studies using all types of ratio and stimuli (rather than only "typical" parameters). This model indicated that the effect size significantly differed from zero in adults (0.76, SE = 0.09; $p < .001$) but not children (-0.26, SE = 0.13; ns). Thus, although children appeared to show a small reverse Colavita effect, this did not reach significance. The effect size seen in experiments with children was significantly smaller than the effect size found in adults ($p < .001$).

3.1.2. 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 (Fig. 3). The amount of scatter around the true effect should decrease with decreased sampling variance/increased sample size, thus producing a classic "funnel" shape (Macaskill et al., 2001). Publication bias is associated with funnel plot asymmetry (Egger et al., 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 ($z = 1.04$, $p = .3$, Fig. 3), suggesting the reported findings were not influenced by publication bias.

3.2. Asymmetrical facilitation: response time analyses

We used studies that had reported response times to auditory and visual stimuli under unimodal and bimodal conditions to investigate whether the Colavita effect occurs due to asymmetrical facilitation and inhibition (Sinnett et al., 2008). Our 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. Our 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
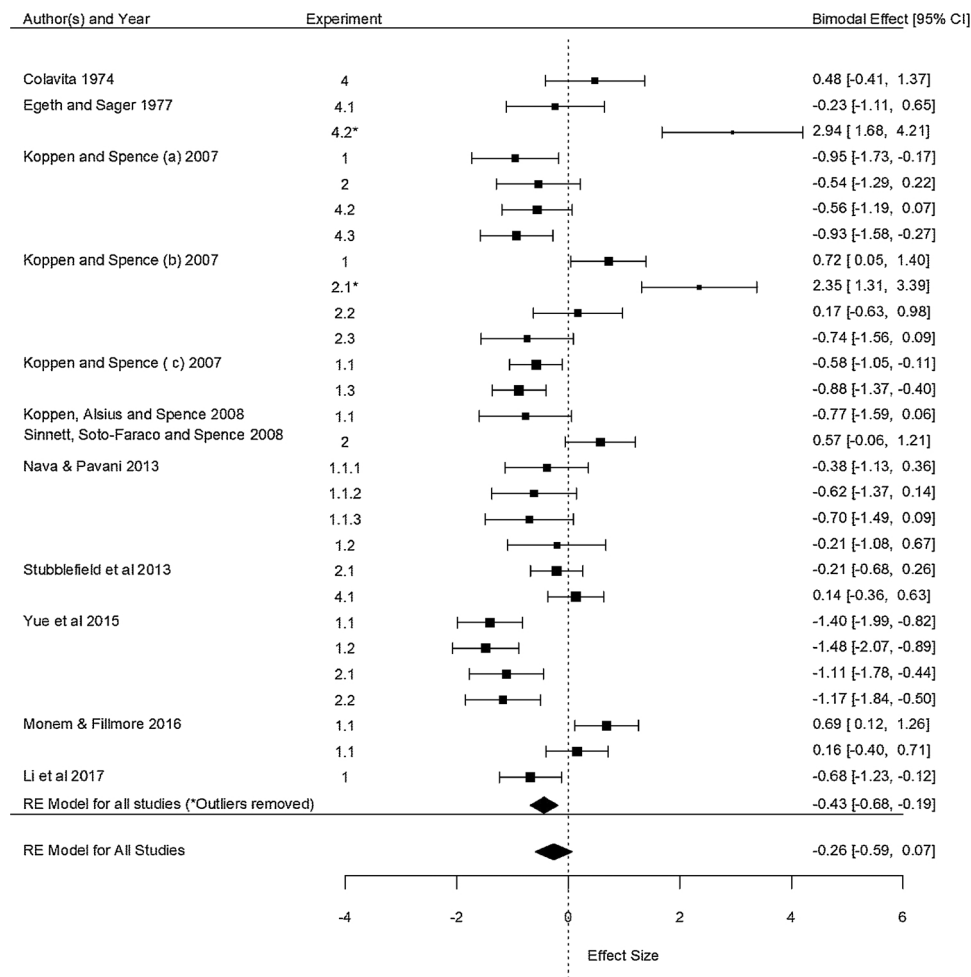
Author(s) and Year	Experiment		Bimodal Effect [95% CI]
Colavita 1974	4		0.48 [-0.41, 1.37]
Egeth and Sager 1977	4.1		-0.23 [-1.11, 0.65]
	4.2*		2.94 [1.68, 4.21]
Koppen and Spence (a) 2007	1		-0.95 [-1.73, -0.17]
	2		-0.54 [-1.29, 0.22]
	4.2		-0.56 [-1.19, 0.07]
	4.3		-0.93 [-1.58, -0.27]
Koppen and Spence (b) 2007	1		0.72 [0.05, 1.40]
	2.1*		2.35 [1.31, 3.39]
	2.2		0.17 [-0.63, 0.98]
	2.3		-0.74 [-1.56, 0.09]
Koppen and Spence (c) 2007	1.1		-0.58 [-1.05, -0.11]
	1.3		-0.88 [-1.37, -0.40]
Koppen, Alsius and Spence 2008	1.1		-0.77 [-1.59, 0.06]
Sinnett, Soto-Faraco and Spence 2008	2		0.57 [-0.06, 1.21]
Nava & Pavani 2013	1.1.1		-0.38 [-1.13, 0.36]
	1.1.2		-0.62 [-1.37, 0.14]
	1.1.3		-0.70 [-1.49, 0.09]
	1.2		-0.21 [-1.08, 0.67]
Stubblefield et al 2013	2.1		-0.21 [-0.68, 0.26]
	4.1		0.14 [-0.36, 0.63]
Yue et al 2015	1.1		-1.40 [-1.99, -0.82]
	1.2		-1.48 [-2.07, -0.89]
	2.1		-1.11 [-1.78, -0.44]
	2.2		-1.17 [-1.84, -0.50]
Monem & Fillmore 2016	1.1		0.69 [0.12, 1.26]
	1.1		0.16 [-0.40, 0.71]
Li et al 2017	1		-0.68 [-1.23, -0.12]
RE Model for all studies (*Outliers removed)			-0.43 [-0.68, -0.19]
RE Model for All Studies			-0.26 [-0.59, 0.07]

-4   -2   0   2   4   6

Effect Size

Fig. 4. 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.

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 our analysis of Colavita errors, we also test the effect of moderators 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.

3.2.1. 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 (Fig. 4). Two experiments (Egeth and Sager, 1977, experiment 4.2; Koppen and 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 et al.'s (2008) predictions of asymmetrical facilitation, response times were *slower* for visual stimuli accompanied by auditory stimuli compared to when they were presented alone.

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 3). 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 versus complex stimuli such as images and naturalistic sounds) found above, two further

Table 3

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.

Factor	df	QM	p
Ratio	5	28.4598	< .0001
Response keys	1	1.9309	.1647
Stimulus category	1	18.3067	< .0001
Congruency	2	7.0490	.0295
Attentional manipulation	2	1.1756	.5556
Age group	1	2.6828	.1014

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.

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.

3.2.2. 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 4). From this we conclude that participants were slower for auditory targets paired with visual stimuli

Table 4

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.

Factor	df	QM	p
Ratio	5	1.0200	0.6005
Response keys	1	2.2846	0.1307
Stimulus category	1	1.6394	.2004
Congruency	2	1.1317	0.2874
Attentional manipulation	2	3.2638	0.1956
Age group	1	3.1034	0.0781
Matched	1	2.5771	0.1084

compared with unimodal targets, as can be seen in Fig. 5, and this was not modulated by experimental parameters.

3.2.3. Is the bimodal slowing effect between vision and audition symmetrical?

Contrary to the prediction based on the hypothesis of Sinnett et al. (2008) we found that vision slowed response times to auditory targets and vice versa. Robinson et al. (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 our former two comparisons. No significant difference was found, suggesting visual and auditory stimuli slowed response times to the opposing modality to a similar extent.

4. 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 our model are likely to influence the effect size of the Colavita effect.

Furthermore, we showed that the Colavita effect may be modulated 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 and Pavani, 2013, n = 51; Wille and 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 and Sloutsky, 2004; Robinson et al., 2016; Robinson and Sloutsky, 2004, 2010; Sloutsky and Napolitano, 2003) and difficulty ignoring auditory distractions in childhood (Hanauer and Brooks, 2003). It should be noted that we used a binary categorisation of age group in our analysis ("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 we report 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.

Our response time analyses suggested that response times 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
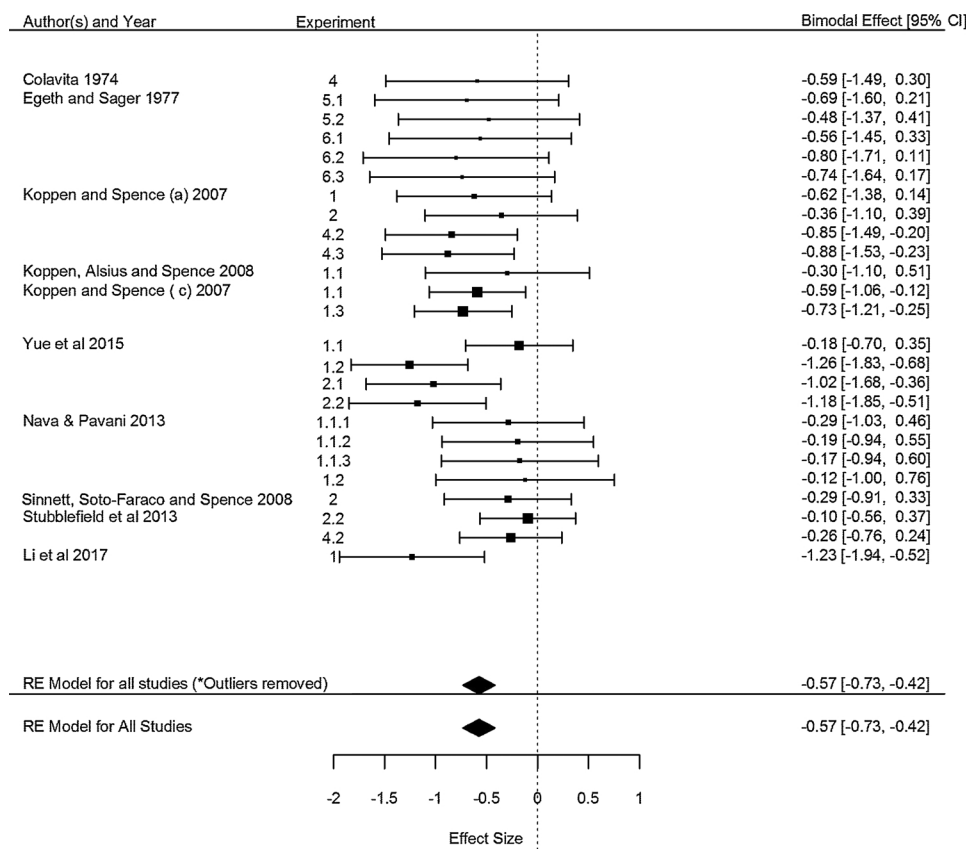
Fig. 5. Effect sizes and 95% confidence intervals for studies/experiments reporting response times (RT) for auditory targets under unimodal and bimodal conditions. Symbol size reflects sample size. Positive effect sizes indicate RT was faster under bimodal versus unimodal conditions. Negative effect sizes indicate RT was faster under unimodal versus bimodal conditions.

integration, whereby neural responses elicited from bimodal targets are greater than unimodal targets (i.e. additive; see Stanford et al., 2004). However, our findings indicated that response times were in fact *slower* under bimodal conditions. This finding appears contrary to both asymmetric and symmetric models of multisensory facilitation.

One likely explanation for our findings of 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.

Our 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 our 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 and 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 our 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 we found 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.

5. 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 paper you may be more distracted by an email pop-up versus 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.

Ethical conduct

The methodology included here was approved by the ethical review board of the School of Psychology, University of Nottingham, and conducted in accordance with the declaration of Helsinki.

Funding

References

Algina, J., Keselman, H.J., 2003. Approximate confidence intervals for effect sizes. Educ. Psychol. Meas. 63 (4), 537–553. https://doi.org/10.1177/0013164403256358.

Barnhart, W.R., Rivera, S., Robinson, C.W., 2018. Different patterns of modality dominance across development. Acta Psychol. 182, 154–165. https://doi.org/10.1016/j.actpsy.2017.11.017.

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

Colavita, F.B., 1974. Human sensory dominance. Percept. & Psychophys. 16 (2), 409–412. https://doi.org/10.3758/BF03203962.

Colavita, F.B., 1982. Visual dominance and attention in space. Bull. Psychon. Soc. 19 (5), 261–262. https://doi.org/10.3758/BF03330251.

Colavita, F.B., Weisberg, D., 1979. A further investigation of visual dominance. Percept. Psychophys. 25 (4), 345–347. https://doi.org/10.3758/BF03198814.

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

Cumming, G., 2012. Understanding the New Statistics: Effect Sizes, Confidence Intervals, and Meta- Analysis. Routledge, New York.

Cumming, G., Calin-Jagerman, R., 2017. Introduction to the New Statistics: Estimation, Open Science and Beyond. Routledge, New York.

Diederich, A., Colonius, H., 2004. Bimodal and trimodal multisensory enhancement: effects of stimulus onset and intensity on reaction time. Percept. Psychophys. 66 (8), 1388–1404. https://doi.org/10.3758/BF03195006.

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

Egeth, H.E., Sager, L.C., 1977. On the locus of visual dominance. Percept. Psychophys. 22 (1), 77–86. https://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. Br. Med. J. 315 (7109), 629–634. https://doi.org/10.1136/bmj.316.7129.469.

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. Exp. Brain. Res. 143 (4), 480–487. https://doi.org/10.1007/s00221-002-1017-9.

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

Gondan, M., Niederhaus, B., Rösler, F., Röder, B., 2005. Multisensory processing in the redundant-target effect: a behavioral and event-related potential study. Percept. Psychophys. 67 (4), 713–726. https://doi.org/10.3758/BF03193527.

Graven, S.N., Browne, J.V., 2008a. Auditory development in the fetus and infant. Newborn Infant. Nurs. Rev. 8 (4), 187–193.

Graven, S.N., Browne, J.V., 2008b. Visual development in the human fetus, infant, and young child. Newborn Infant. Nurs. Rev. 8 (4), 194–201.

Hanauer, J.B., Brooks, P.J., 2003. Developmental change in the cross-modal stroop effect. Percept. & Psychophys. 65 (3), 359–366. https://doi.org/10.3758/BF03194567.

Hartcher-O'Brien, J., Levitan, C., Spence, C., 2010. Extending visual dominance over touch for input off the body. Brain Res. 1362, 48–55. https://doi.org/10.1016/j.brainres.2010.09.036.

Hecht, D., Reiner, M., 2009. Sensory dominance in combinations of audio, visual and haptic stimuli. Exp. Brain. Res. 193 (2), 307–314. https://doi.org/10.1007/s00221-008-1626-z.

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. J. Neurosci. 35 (24), 9064–9077. https://doi.org/10.1523/JNEUROSCI.3760-14.2015.

Hullett, C.R., Levine, T.R., 2003. the overestimation of effect sizes from F values in meta-analysis: the cause and a solution. Commun. Monogr. 70 (1), 52–67. https://doi.org/10.1080/715114664.

Johnson, T.L., Shapiro, K.L., 1989. Attention to auditory and peripheral visual stimuli: effects of arousal and predictability. Acta Psychol. 72 (3), 233–245. https://doi.org/10.1016/0001-6918(89)90031-0.

Koppen, C., Spence, C., 2007a. Assessing the role of stimulus probability on the Colavita visual dominance effect. Brain Res. 1186 (1), 224–232. https://doi.org/10.1016/j.brainres.2007.09.076.

Koppen, C., Spence, C., 2007b. Audiovisual asynchrony modulates the Colavita visual dominance effect. Brain Res. 1186 (1), 224–232. https://doi.org/10.1016/j.brainres.2007.09.076.

Koppen, C., Spence, C., 2007c. Seeing the light: exploring the Colavita visual dominance effect. Brain Res. 1186 (1), 224–232. https://doi.org/10.1016/j.brainres.2007.09.076.

Koppen, C., Spence, C., 2007d. Spatial coincidence modulates the Colavita visual dominance effect. Neurosci. Lett. 417, 107–111. https://doi.org/10.1016/j.brainres.2007.09.076.

Koppen, C., Alsius, A., Spence, C., 2008. Semantic congruency and the Colavita visual dominance effect. Exp. Brain. Res. 184 (4), 533–546. https://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. Exp. Brain. Res. 196 (3), 353–360. https://doi.org/10.1007/s00221-009-1853-y.

Lakens, 2013. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Frontiers Psychology 4 (November), 1–12. https://doi.org/10.3389/fpsyg.2013.00863.

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. Front. Psychol. 8 (March), 1–8. https://doi.org/10.3389/fpsyg.2017.00303.

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. Cogn. Brain Res. 17 (2), 447–453. https://doi.org/10.1016/S0926-6410(03)00160-5.

Macaskill, P., Walter, S.D., Irwig, L., 2001. A comparison of methods to detect publication bias in meta-analysis. Stat. Med. 20 (4), 641–654. https://doi.org/10.1002/sim.698.

McGraw, K.O., Wong, S., 1992. A common language effect size statistic. Psychol. Bull. 111 (2), 361–365.

Meredith, M.A., Stein, B.E., 1986. Visual, auditory and somatosensory convergence on cells in superior colliculus results in multisensory integration. J. Neurophysiol. 56 (3), 640–662.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., The PRISMA Group, 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 6 (7). 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. Cogn. Brain Res. 14 (1), 115–128. https://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. Psychol. Add. Behav. 30 (1), 12–17. https://doi.org/10.1037/adb0000140.Alcohol-Related.

Moro, S.S., Steeves, J.K.E., 2012. No Colavita effect: equal auditory and visual processing in people with one eye. Exp. Brain. Res. 216 (3), 367–373. https://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. Neurosci. Lett. 556, 186–190. https://doi.org/10.1016/j.neulet.2013.09.064.

Napolitano, A.C., Sloutsky, V.M., 2004. Is a picture worth a thousand words? The flexible nature of modality dominance in young children. Child Dev. 75 (6), 1850–1870. https://doi.org/10.1111/j.1467-8624.2004.00821.x.

Nava, E., Pavani, F., 2013. Changes in sensory dominance during childhood: converging evidence from the Colavita effect and the sound-induced flash illusion. Child Dev. 84 (2), 604–616. https://doi.org/10.1111/j.1467-8624.2012.01856.x.

Ngo, M.K., Sinnett, S., Soto-Faraco, S., Spence, C., 2010. Repetition blindness and the Colavita effect. Neurosci. Lett. 480 (3), 186–190. https://doi.org/10.1016/j.neulet.2010.06.028.

Ngo, M.K., Cadieux, M.L., Sinnett, S., Soto-Faraco, S., Spence, C., 2011. Reversing the Colavita visual dominance effect. Exp. Brain. Res. 214 (4), 607–618. https://doi.org/10.1007/s00221-011-2859-9.

Noesselt, T., Bergmann, D., Hake, M., Heinze, H.J., Fendrich, R., 2008. Sound increases the saliency of visual events. Brain Res. 1220, 157–163. https://doi.org/10.1016/j.brainres.2007.12.060.

Occelli, V., Hartcher-O'Brien, J., Spence, C., Zampini, M., 2010. Assessing the audiotactile Colavita effect in near and rear space. Exp. Brain. Res. 203 (3), 517–532. https://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. Percept. Psychophys. 65 (1), 123–132.

Odgaard, E.C., Arieh, Y., Marks, L.E., 2004. Brighter noise: sensory enhancement of perceived loudness by concurrent visual stimulation. Cogn. Affect Behav. Neurosci. 4 (2), 127–132. https://doi.org/10.3758/CABN.4.2.127.

Osborn, W.C., Sheldon, R.W., Baker, R.A., 1963. Vigilance performance under conditions of redundant and nonredundant signal presentation. J. Appl. Psychol. 47 (2), 130–134. https://doi.org/10.1037/h0040874.

Robinson, C.W., Sloutsky, V.M., 2004. Auditory dominance and its change in the course of development. Child Dev. 75 (5), 1387–1401. https://doi.org/10.1111/j.1467-8624.2004.00747.x.

Robinson, C.W., Sloutsky, V.M., 2010. Development of cross-modal processing. Wiley Interdiscip. Rev. Cogn, Sci 1 (1), 135–141. https://doi.org/10.1002/wcs.12.

Robinson, C.W., Chandra, M., Sinnett, S., 2016. Existence of competing modality dominances. Atten. Percept. Psychophys 78 (4), 1104–1114. https://doi.org/10.3758/s13414-016-1061-3.

Shapiro, K.L., Johnson, T.L., 1987. Effects of arousal on attention to central and peripheral visual stimuli. Acta Psychol. 66 (2), 157–172. https://doi.org/10.1016/0001-6918(87)90031-X.

Shapiro, K.L., Egerman, B., Klein, R.M., 1984. Effects of arousal on human visual dominance. Percept. Psychophys. 35 (6), 547–552. https://doi.org/10.3758/BF03205951.

Sinnett, S., Spence, C., Soto-Faraco, S., 2007. Visual dominance and attention: the Colavita effect revisited. Percept. Psychophys. 69 (5), 673–686. https://doi.org/10.3758/BF03193770.

Sinnett, S., Soto-Faraco, S., Spence, C., 2008. The co-occurrence of multisensory competition and facilitation. Acta Psychol. 128 (1), 153–161. https://doi.org/10.1016/j.actpsy.2007.12.002.

Sloutsky, V.M., Napolitano, A.C., 2003. Is A picture worth a thousand words? Preference for auditory modality in young children. Child Dev. 74 (3), 822–833. https://doi.org/10.1111/1467-8624.00570.

Spence, C., 2009. Explaining the Colavita visual dominance effect. Prog. Brain Res. 176, 245–258. https://doi.org/10.1016/S0079-6123(09)17615-X.

Spence, C., Parise, C., Chen, Y.C., 2012. The Colavita visual dominance effect. In: Murray, M.M., Wallace, M.T. (Eds.), The Neural Bases of Multisensory Processes. Taylor & Francis, Boca Raton (FL), pp. 529–557.

Stanford, T.R., Wallace, M.T., Vaughan, W.J., Jiang, W., 2004. Crossmodal spatial interactions in subcortical and cortical circuits. In: Spence, C., Driver, J. (Eds.), Crossmodal Space and Crossmodal Attention. Oxford University Press, Oxford, pp. 25–50.

Stekelenburg, J.J., Keetels, M., 2016. The effect of synesthetic associations between the visual and auditory modalities on the Colavita effect. Exp. Brain. Res. 234 (5), 1209–1219. https://doi.org/10.1007/s00221-015-4363-0.

Stubblefield, A., Jacobs, L.A., Kim, Y., Goolkasian, P., 2013. Colavita dominance effect revisited: the effect of semantic congruity. Atten. Percept. & Psychophys. 75 (8), 1827–1839. https://doi.org/10.3758/s13414-013-0530-1.

Thompson, S.G., Higgins, J.P.T., 2002. How should meta-regression analyses be undertaken and interpreted? Stat. Med. 21 (11), 1559–1573. https://doi.org/10.1002/sim.1187.

Tsukiko, M., Desmarais, G., 2014. The impact of sensory dominance and congruency effects on multisensory integration. Can. J. Exp. Psychol. 68 (4), 303.

Van Damme, S., Crombez, G., Spence, C., 2009. Is Visual dominance modulated by the threat value of visual and auditory stimuli? Exp. Brain. Res. 193 (2), 197–204. https://doi.org/10.1007/s00221-008-1608-1.

Viechtbauer, W., 2010. Conducting meta-analyses in R with the metafor package. J. Stat. Softw. 36 (3), 1–48. https://doi.org/10.1103/PhysRevB.91.121108.

Welch, R.B., Warren, D.H., 1980. Immediate perceptual response to intersensory discrepancy. Psychol. Bull. 88 (3), 638–667. https://doi.org/10.1037/0033-2909.88.3.638.

Wille, C., Ebersbach, M., 2016. Semantic congruency and the (reversed) Colavita effect in children and adults. J. Exp. Child Psychol. 141, 23–33. https://doi.org/10.1016/j.jecp.2015.07.015.

Yue, Z., Jiang, Y., Li, Y., Wang, P., Chen, Q., 2015. Enhanced visual dominance in far space. Exp. Brain. Res. 233 (10), 2833–2843. https://doi.org/10.1007/s00221-015-4353-2.