Acta Psychologica 182 (2018) 154–165

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

Different patterns of modality dominance across development

Wesley R. Barnhart[a], Samuel Rivera[b], Christopher W. Robinson[c,*]

[a] ID/ASD Research Group, Nisonger Center, University Center for Excellence in Developmental Disabilities, USA
[b] Department of Psychology, The Ohio State University, USA
[c] Department of Psychology, The Ohio State University Newark, USA

ARTICLE INFO

Keywords:
Attention
Cross-modal processing
Development
Modality dominance

ABSTRACT

The present study sought to better understand how children, young adults, and older adults attend and respond to multisensory information. In Experiment 1, young adults were presented with two spoken words, two pictures, or two word-picture pairings and they had to determine if the two stimuli/pairings were exactly the same or different. Pairing the words and pictures together slowed down visual but not auditory response times and delayed the latency of first fixations, both of which are consistent with a proposed mechanism underlying auditory dominance. Experiment 2 examined the development of modality dominance in children, young adults, and older adults. Cross-modal presentation attenuated visual accuracy and slowed down visual response times in children, whereas older adults showed the opposite pattern, with cross-modal presentation attenuating auditory accuracy and slowing down auditory response times. Cross-modal presentation also delayed first fixations in children and young adults. Mechanisms underlying modality dominance and multisensory processing are discussed.

1. Introduction

Over the last forty years there has been a growing body of research examining how sensory modalities interact while processing multisensory information (e.g., sounds and pictures paired together). Under some conditions, presenting congruent information (e.g., stimuli provide converging details across modalities) to multiple sensory modalities facilitates processing (Fort, Delpuech, Pernier, & Giard, 2002; Giard & Peronnet, 1999; McGurk & MacDonald, 1976; Miller, 1982; see also Bahrick, Lickliter, & Flom, 2004; Spence & Driver, 2004 for reviews). However, there are many situations where multisensory information is incongruent in nature, with stimuli in one modality providing little to no details about stimuli presented to another modality. Research examining processing of incongruent information often shows modality dominance effects, with one sensory modality interfering with processing in a second modality (see Robinson & Sloutsky, 2010; Sinnett, Spence, & Soto-Faraco, 2007; Spence, 2009; Spence, Parise, & Chen, 2012, for reviews). Given that most of our experiences are multisensory in nature, it is important to examine how multisensory presentation affects processing of auditory and visual information at various points in development. The present study contributes to this research by investigating modality dominance effects across development.

There is a clear pattern within the young adult literature: when simultaneously presented with auditory and visual information, visual input often dominates (Colavita, Tomko, & Weisberg, 1976; Colavita & Weisberg, 1979; Egeth & Sager, 1977; Koppen, Alsius, & Spence, 2008; Ngo, Cadieux, Sinnett, & Soto-Faraco, 2011; Ngo, Sinnett, Soto-Faraco, & Spence, 2010; Sinnett et al., 2007; Sinnett, Soto-Faraco, & Spence, 2008). For example, in a typical Colavita task, participants are instructed to press one button when they see a light and press a different button when they hear a tone (Colavita, 1974). In said research, most trials are unimodal (only light or sound); however, some are cross-modal (light and sound are paired together). On these cross-modal trials, participants often respond incorrectly by only pressing the visual button as opposed to correctly pressing both buttons or a third button associated with cross-modal stimuli. Visual dominance has been extended to different tasks with a variety of attentional manipulations failing to reverse the effect (Ngo et al., 2010; see also Spence, 2009 for a review). While there is some evidence that auditory input can attenuate visual processing, these findings require significant modifications to the task or come from tasks that are temporal in nature and favor the auditory modality (Ngo et al., 2011; Robinson & Sloutsky, 2013; Shams, Kamitani, & Shimojo, 2000, 2002).

Numerous accounts have been put forward to explain visual dominance (see Sinnett et al., 2007; Spence, 2009; Spence et al., 2012, for reviews). For example, according to the modality appropriateness hypothesis, the modality that is most appropriate for a given task will

* Corresponding author at: Department of Psychology, The Ohio State University Newark, 1179 University Drive, Newark, OH 43055, USA.
E-mail address: robinson.777@osu.edu (C.W. Robinson).

https://doi.org/10.1016/j.actpsy.2017.11.017
Received 24 July 2017; Received in revised form 2 November 2017; Accepted 18 November 2017
0001-6918/

dominate, with the visual modality dominating spatial tasks and the auditory modality dominating temporal tasks (Welch & Warren, 1980). Attentional and sensory factors may also underlie modality dominance effects. For example, while auditory stimuli may automatically engage attention, an attentional bias to the visual modality might be needed to compensate for poor altering qualities of visual stimuli (Posner, Nissen, & Klein, 1976, but see Koppen & Spence, 2007a; Sinnett et al., 2007, which show that visual dominance cannot be reversed by directing adults' attention to the auditory modality). Other research has considered how quickly the auditory and visual stimuli engage attention, with the visual input possibly being detected first (Koppen & Spence, 2007b; Rutschmann & Link, 1964). More recent accounts posit sensory/network dominance, with interneurons between sensory systems being inhibitory in nature (i.e., activation of one inhibits the other and vice versa; Desimone & Duncan, 1995; Duncan, 1996; Spence et al., 2012). According to this view, visual dominance should be more likely to occur given that over half of the brain is devoted to processing visual information (Sereno et al., 1995), resulting in strong inhibition of other sensory systems.

However, it is unclear how the proposed mechanisms underlying visual dominance can account for developmental findings, which often show that auditory information often dominates processing of visual input (Lewkowicz, 1988a, 1988b; Nava & Pavani, 2013; Robinson & Sloutsky, 2004; Sloutsky & Napolitano, 2003; Sloutsky & Robinson, 2008). For example, using the Colavita visual dominance task with 6–7, 9–10, and 11–12 year-old children, 6–7 year-old's responses were consistent with auditory dominance (pressed sound button on cross-modal trials) while older children were consistent with visual dominance (pressed picture button on cross-modal trials). This shift from auditory to visual dominance (or increase reliance on visual information) dovetails with developmental research examining the McGurk effect (Massaro, 1984), Sound Induced Flash Illusion - SIFI (Nava & Pavani, 2013), Colavita-like task using semantically meaningful stimuli (Wille & Ebersbach, 2016), inductive generalization (Robinson & Sloutsky, 2004), and change detection (Napolitano & Sloutsky, 2004; Sloutsky & Napolitano, 2003).

While modality dominance effects appear to change across development, recent studies using more sensitive procedures show some support for auditory dominance in young adults (Dunifon, Rivera, & Robinson, 2016; Robinson, Chandra, & Sinnett, 2016; Robinson & Sloutsky, 2013). For example, using a modified oddball paradigm, Robinson et al. (2016) examined how quickly young adults discriminated a frequently presented AV pairing (standard) from infrequently presented auditory oddballs (auditory oddball paired with visual standard) and visual oddballs (visual oddball paired with auditory standard). When participants were instructed to quickly press the spacebar to any oddball, auditory dominance was found with cross-modal presentation slowing down visual but not auditory discrimination (Experiment 1). It is important to note that this effect reversed when participants had to not only detect a change, but also report what changed (i.e., press one button for visual oddball, a different button for auditory oddball, or a third button if both AV components changed (Experiment 2). This latter finding suggests that visual dominance might occur later in the course of processing by dominating the response/decision phase of processing, as opposed to disrupting early encoding of auditory information. However, one limitation of this study was that the auditory stimuli were simple tones paired with monochromatic and unfamiliar images.

Dunifon et al. (2016) extended these findings by using another variation of a change detection task while using semantically meaningful visual stimuli and more dynamic nonlinguistic sounds. Young adults had to quickly discriminate two visual stimuli (unimodal visual condition), two auditory stimuli (unimodal auditory condition), or two auditory-visual pairings (cross-modal condition). In addition to examining response times and accuracies, this study also examined visual fixations while participants were making discriminations. Simultaneously presenting the auditory and visual stimuli in the cross-modal condition was more likely to slow down visual response times (Experiments 1–3), even when participants were instructed to ignore the auditory stimuli (Experiment 2). Moreover, the presence of the sound also delayed the onset of first fixations to the visual stimulus (relative to the unimodal visual condition) and increased participants' mean fixation durations.

One potential explanation for auditory dominance is that sensory modalities may share a "pool" of available attentional resources and compete for these resources (see Duncan, Martens, & Ward, 1997; Eimer & Driver, 2000; Eimer & van Velzen, 2002; Pavani, Husain, Ládavas, & Driver, 2004; Sinnett et al., 2007; Wickens, 1984 for related discussions). Moreover, since auditory stimuli are often dynamic and transient in nature, it may be adaptive for the system to allocate greater attentional resources to auditory stimuli to ensure this information is processed before it disappears, especially early in stimulus presentation (see Robinson & Sloutsky, 2010 for a review). Thus, auditory dominance may stem from auditory stimuli automatically engaging attention early in the course of processing and attenuating or delaying visual processing. While this proposed mechanism predicts some of the developmental findings, it is unclear how to reconcile this account with the numerous studies that clearly show evidence of visual dominance (see Sinnett et al., 2007; Spence, 2009; Spence et al., 2012, for reviews) and if such an account can predict modality dominance effects in older adults.

It is well established that there are substantial changes to the sensory, motor, and cognitive systems into late adulthood (see Birren & Schaie, 2006, for a review), and it is unclear how these developmental/maturational changes affect multisensory processing and modality dominance. For example, previous research has documented that older adults are frequently outperformed by young adults on tasks of memory (e.g., Craik, 1994), motor response (e.g., Pratt, Chasteen, & Abrams, 1994), and executive control (e.g., Royall, Palmer, Chiodo, & Polk, 2004), and it is well established that there are also sensory declines in late adulthood (e.g., Corso, 1971; He, Dubno, & Mills, 1998; Schneider, Daneman, Murphy, & Kwong, 2000; Weale, 1975). However, to our knowledge, there is no research examining modality dominance effects in older adults. As such, it is unclear if the shift from auditory to visual dominance will continue, with older adults showing stronger visual dominance effects than children and young adults. However, it is also possible that visual dominance effects plateau, reverse to auditory dominance, or no dominance effects will be present. Pattern of dominance will shed light on how older adults prioritize different components of multisensory information and may predict which types of multisensory experiences facilitate or interfere with learning.

Potential support for a reversal to auditory dominance comes from multisensory integration research (DeLoss, Pierce, & Andersen, 2013; Laurienti, Burdette, Maldjian, & Wallace, 2006). For example, in DeLoss et al. (2013), young and older adults participated in the Sound Induced Flash Illusion task (SIFI) and had to ignore beeps and report how many flashes were presented on a computer screen. While the number of beeps presented influenced visual perception in both groups, the effect of beeps on visual perception was stronger in older adults. A parallel finding can be found when examining speeded responses to unimodal and cross-modal targets (Laurienti et al., 2006). Both young and older adults were faster to respond to cross-modal targets than visual targets, however, older adults appeared to benefit more from the presence of the sound. Thus, across both studies, auditory stimuli had a greater effect on visual processing and responding to targets in older adults. These effects might be related to the inhibitory deficit hypothesis, which posits that older adults may have difficulty filtering out cross-modal stimuli (Lustig, Hasher, & Zacks, 2007). Thus, auditory stimuli may be more likely to be combined with visual information (multisensory integration research) or facilitate/interfere with visual processing (modality dominance research) because these stimuli are more likely to be detected and encoded late in development (due to declines in filtering cross-modal information).

On the other hand, other studies suggest that visual dominance effects may continue to strengthen into late adulthood. For example, it is well established that visual information such as lip movements can contribute to speech perception (McGurk & MacDonald, 1976). Moreover, when examining McGurk effects in young and older adults, older adults seem to be relying more on visual information than younger adults (Sekiyama, Soshi, & Sakamoto, 2014). This reliance on visual information can come with a cost, as evidenced by research examining the impact of visual/auditory distractors on cross-modal selective attention in older adults (Guerreiro, Murphy, & Van Gerven, 2010; Van Gerven & Guerreiro, 2016). More specifically, Van Gerven & Guerreiro identified that visual distractors in a cross-modal selective attention task adversely impacted processing in the auditory modality for older adults, but not younger adults. Importantly, this modality-dependent impairment was not identified in reversed cases (e.g., impaired visual processing due to auditory distractors), nor in cases where both the targets and distractors were presented through a single sensory modality (Van Gerven & Guerreiro, 2016).

The primary aim of Experiment 1 was to determine if auditory dominance effects found in previous studies can be generalized to different classes of stimuli. In Experiment 1, young adults had to quickly discriminate two words, two pictures, or two word-picture pairings, and we recorded accuracies, response times, and visual fixations while participants completed the tasks. Based on previous research and on a proposed mechanism underlying auditory dominance (Robinson & Sloutsky, 2010), it was hypothesized that pairing pictures and words would slow down processing of the visual stimulus and have no negative effect on auditory processing. Experiment 1 expands on Robinson et al. (2016) by using a different methodology and stimuli to examine modality dominance. Recall that Robinson et al. (2016) found evidence of auditory dominance when using a modified oddball procedure with monochromatic, unfamiliar visual images (Experiments 1 and 3), whereas the present study used colorful drawings of familiar and more meaningful images (tree, rabbit, etc.) and recorded visual fixations while participants were completing the task. The present study also expands on Dunifon et al. (2016) by examining if auditory dominance effects are restricted to non-speech sounds or if these effects can be generalized to other classes of stimuli such as linguistic labels.

Experiment 2 expands on the first experiment by determining if patterns of modality dominance differ across development. Experiment 2 employed a similar methodology to examine patterns of modality dominance in children, young adults, and older adults. Based on previous research examining developmental differences in modality dominance in children and young adults (Nava & Pavani, 2013; Robinson & Sloutsky, 2004; Sloutsky & Napolitano, 2003), it was hypothesized that auditory dominance effects should be stronger in children; however, given the findings from Dunifon et al. (2016), it is possible that the present study would also find evidence of auditory dominance in young adults (i.e., slower visual discrimination and delayed visual fixations when images are paired with words). To the best of our knowledge, no research has examined modality dominance in older adults, and as such, it was unclear if older adults would show auditory, visual, or no dominance. That said, there is evidence from multisensory integration literature that sensory modalities become more integrated with age, with multisensory integration having stronger effects in older adults (DeLoss et al., 2013; Laurienti et al., 2006). However, the direction of these effects (auditory versus visual dominance) are less clear, with some research showing the effects of sounds on visual processing increasing with age (DeLoss et al., 2013; Laurienti et al., 2006) and other research showing increased reliance on visual information (Sekiyama et al., 2014; Van Gerven & Guerreiro, 2016).

2. Experiment 1: Modality dominance in young adults

2.1. Method

2.1.1. Participants

Thirty-eight undergraduate students (M = 19.52 years, Range 18.04–33.48, 20 Females) enrolled in an Introductory Psychology course participated in this experiment. Completion of the study granted participants with credit that served to fulfill a course requirement. All participants provided informed consent, had normal hearing and vision (self-reported), and were debriefed after completion of the study. An additional eight adults participated but were excluded from the final analyses due to poor eye tracking calibration.

2.1.2. Apparatus

Participants were centrally positioned and seated approximately 60 cm in front of an EyeLink 1000 Plus eye tracker with desktop mount and remote camera. The eye tracker computed eye movements at a rate of 500 Hz, and Experiment Builder 1.10.165 controlled the timing of stimulus presentations. Visual stimuli were presented on a BenQ XL2420 24″ monitor and auditory stimuli were presented via Kensington 33,137 headphones. Eye tracking data were collected and stored on a Dell Optiplex 7010 computer. Gaze fixation positions and durations were identified by the EyeLink system online during the experiment and recorded for offline analysis. The eye tracker, stimulus presentation computer, and eye tracking computer were stationed in a quiet testing room. A trained experimenter oversaw the entire duration of each participant's study and they manually started each trial when the participants fixated on a central stimulus.

2.1.3. Stimuli

Visual stimuli consisted of four pairs of images which were digitally constructed in Microsoft PowerPoint and exported as 600 × 600 bmp files (approximate size), see Fig. 1 for examples of visual stimuli and Areas Of Interest (AOI). The stimuli resembled the following real-world objects: cone of cotton candy, tree, globe, and rabbit, and each stimulus pair differed by two or four features. For example, as can be seen in Fig. 1, the diamond and circle could be used to differentiate the two trees; thus, these two features/AOIs were considered *relevant*. The heart and star were considered *irrelevant* because both trees shared these two features and therefore cannot be used to differentiate. Within an individual trial, one of the items from the pair was presented first for 1 s (Target), with a 1 s Inter-Stimulus Interval (ISI). The second item (Test) was presented until the participant made a response. Each item in the pair was equally likely to be the Target or Test item.

As with visual stimuli, auditory stimuli consisted of four word pairs. The auditory stimuli used were one- (e.g., paf vs. dax and ket vs. yun)

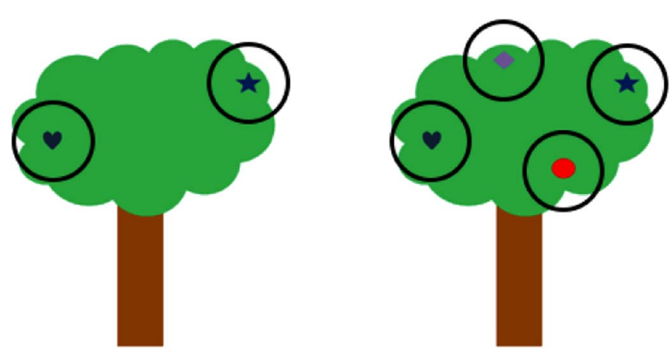

Fig. 1. Example of a visual pair used in Experiments 1 and 2. The circles around each feature denote the Areas of Interest and were not visible during the experiment. Here, the relevant AOIs are the red circle and purple diamond (right image), as these two features can be used to differentiate the stimuli. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and two-syllable nonsense words (e.g., lapo vs. vika and kuna vs. hoonae). Each word was individually spoken by a female experimenter and recorded using Cool Edit 2000. Words were approximately 700 ms in duration (range 480–1000 ms). Audio files were saved as 44.1 kHz wav files and presented to participants via headphones at approximately 65–68 dB. Each item in the pair was equally likely to be the Target or the Test item. Stimuli in the cross-modal condition were created by presenting images and words at the same time. While the visual Test image remained visible until participants made a response, the auditory Test stimulus was only presented for the stimulus duration (approximately 700 ms).

2.1.4. Design

Each participant completed three conditions and condition order was randomized: unimodal auditory (UA), unimodal visual (UV), and cross-modal (CM) conditions, with the UA and UV conditions serving as baselines. In the UA and UV conditions, participants were either presented with two words or two images, respectively, and they had to determine if the stimuli were exactly the same or different. In the CM condition they had to discriminate the same words and pictures; however, the auditory and visual information were presented at the same time. Discrimination in the cross-modal condition was compared to unimodal auditory and visual baselines. Visual dominance would be inferred if cross-modal presentation had a greater cost on auditory processing (greater slowdown in processing compared to UA baseline), and auditory dominance would be inferred if cross-modal presentation had a greater cost on visual processing (greater slowdown in processing compared to UV baseline). Increased response times in both modalities in the cross-modal condition would suggest increased task demands and no evidence that one modality dominated the other modality.

2.1.5. Procedure

Participants were positioned to face the eye tracker centrally with an approximate viewing distance of 60 cm. At the right side of each participant was the experimenter; s/he began the experiment by calibrating participants' eye measurements, a process that included a 9-point sequence of fixations followed by a 9-point validation. The initial calibration/validation process lasted approximately 1–5 min. After calibration, participants were presented with a screen that discussed the experiment's instructions. In the unimodal auditory and visual conditions participants were told they would hear two words or see two pictures and had to press 1 on the number pad if the stimuli were exactly the same and press 3 if the stimuli were different. Participants were also told to respond as quickly and as accurately as possible. There were 60 trials in each condition, half same trials and half different trials, and each trial began with drift correction (i.e., central fixation stimulus). See Fig. 2 for examples of the different trial types. In the cross-modal condition, participants were told that they would see two picture-word pairs and they were instructed to press 1 on the number pad if both the pictures and words were exactly the same ($AUD^1VIS^1 \rightarrow AUD^1VIS^1$). They were told to press 3 if the word changed ($AUD^1VIS^1 \rightarrow AUD^2VIS^1$), the picture changed ($AUD^1VIS^1 \rightarrow AUD^1VIS^2$), or if both components changed ($AUD^1VIS^1 \rightarrow AUD^2VIS^2$). There were 60 trials in the cross-modal condition, 15 of each of the trial types listed above, and each trial began with drift correction. Order of condition (auditory, visual, and cross-modal) was randomized for each participant, and as in the unimodal conditions, they were instructed to respond quickly and accurately.

2.2. Results and discussion

2.2.1. Behavioral analyses

Initial analyses focused on proportion of correct responses on Different Auditory and Different Visual trials. More specifically, discrimination of words in the cross-modal condition ($AUD^1VIS^1 \rightarrow AUD^2VIS^1$) was compared with discrimination of the same words in the unimodal condition ($AUD^1 \rightarrow AUD^2$). We also examined discrimination of visual images in the cross-modal condition ($AUD^1VIS^1 \rightarrow AUD^1VIS^2$) with discrimination of the same images in the unimodal condition ($VIS^1 \rightarrow VIS^2$). Proportion of correct response across the four trial types/conditions were submitted to a 2 Modality (Auditory vs. Visual) × 2 Presentation (Unimodal vs. Cross-modal) repeated measures ANOVA, which revealed a Modality × Presentation interaction, F (1,37) = 8.10, p = 0.007, η_p^2 = 0.18. See black bars (discrimination of words and pictures when presented unimodally) and grey bars (discrimination of words and pictures when presented cross-modally) in Fig. 3 for means and standard errors (see also top section of Table 1 for all means and standard errors). As can be seen in the figure, proportion of correct responses on auditory different trials was significantly lower in the cross-modal condition (M = 0.94, SE = 0.01) than in the unimodal condition, (M = 0.98 ms, SE = 0.01), t (37) = 2.69, p = 0.011. At the same time, cross-modal presentation had no cost on visual accuracy.

We also examined how cross-modal presentation affected response times. Response times greater than two standard deviations above the mean were removed from further analyses. The four means were submitted to a 2 Modality (Auditory vs. Visual) × 2 Presentation (Unimodal vs. Cross-modal) repeated measures ANOVA, see Fig. 4 for means and standard errors for new auditory and new visual trials (see also lower section of Table 1 for all means and standard errors). The ANOVA revealed a main effect of Modality, F (1,37) = 66.67, $p < 0.001$, η_p^2 = 0.64, a main effect of Presentation, F (1,37) = 10.62, $p < 0.005$, η_p^2 = 0.22, and the predicted Modality x Presentation interaction was also significant, F (1,37) = 13.18, $p < 0.001$, η_p^2 = 0.26. Planned comparisons showed slower visual response times in the cross-modal condition (M = 765 ms, SE = 20.56) than in the unimodal condition (M = 680 ms, SE = 20.57), t (37) = 4.64, $p < 0.001$. The slowdown in the auditory modality was less pronounced and did not reach significance, t (37) = 1.52, p = 0.14. Thus, consistent with auditory dominance, pairing words and pictures together negatively affected visual but not auditory processing.

2.2.2. Eye tracking analyses

According to the proposed mechanism underlying auditory dominance (Robinson & Sloutsky, 2010), auditory input may slow down or delay the onset of visual processing. To further examine this proposal, we directly compared patterns of fixations while participants were discriminating visual stimuli in the unimodal and cross-modal conditions. More specifically, we primarily focused on the latency of first fixations, which might be delayed in the cross-modal condition if auditory stimuli are initially competing with visual input. Moreover, disrupted visual processing and/or increased load could also result in longer fixations, more fixations, and/or a decrease in looking to relevant AOIs (AOIs that differentiate two stimuli).

Latency of first fixations to any part of the visual stimulus, mean fixation durations, proportion looking to relevant AOIs, and number of fixations were derived offline from fixations identified by the EyeLink system with custom MATLAB and Python software developed by the second author. Fixations initiated before the stimulus presentation or after responses were excluded. Latencies were defined as the first fixation start time to any location on the visual stimulus relative to the onset of the visual Test item. We computed each eye tracking variable on new Test trials in the unimodal visual condition ($VIS^1 \rightarrow VIS^2$) and compared these values with eye tracking when the same images were paired with words ($AUD^1VIS^1 \rightarrow AUD^1VIS^2$). See Table 1 for means, standard errors, and effect sizes. Overall, as can be seen in Table 2, latency of first fixations to the new visual stimuli were slower in the cross-modal condition compared to the unimodal condition, F (1, 37) = 4.92, p = 0.033, η_p^2 = 0.12. This delayed onset of first fixation replicates Dunifon et al. (2016) where the same images were paired with nonlinguistic sounds and suggests that this effect is not restricted to a particular class of auditory stimuli. While Dunifon et al. also found that

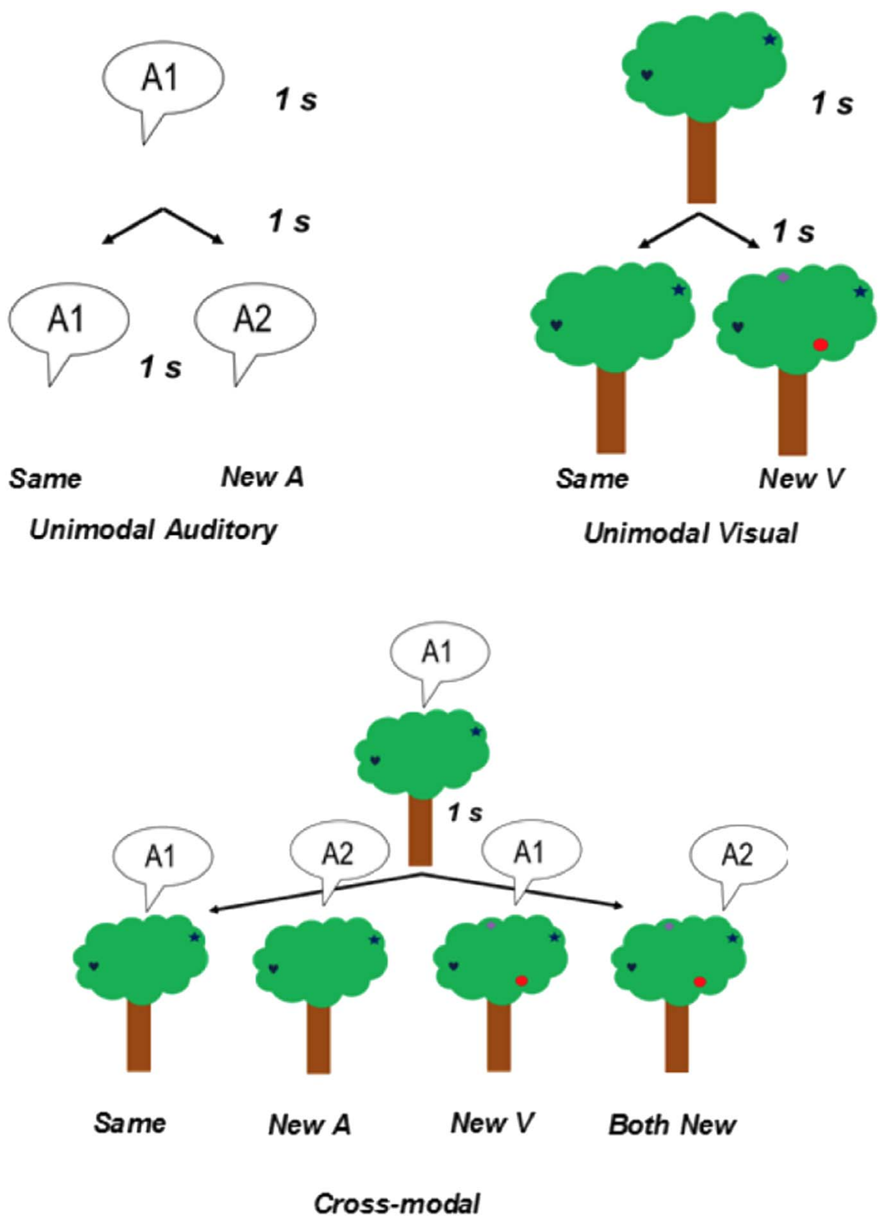

Fig. 2. Trial types from the unimodal auditory, unimodal visual, and cross-modal conditions. For the cross-modal condition, any trial other than "same" elicited the response of "different".

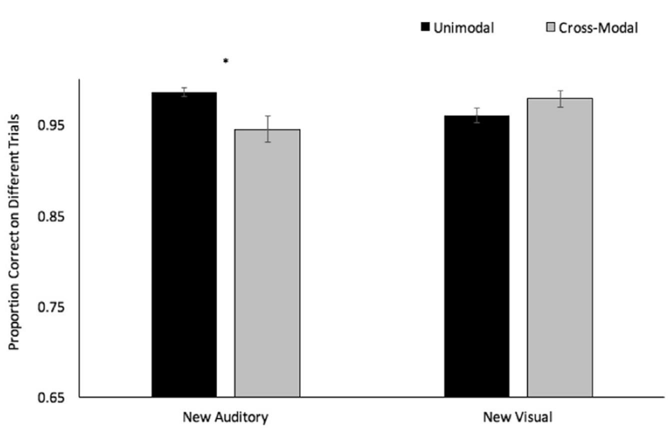

Fig. 3. Proportion correct on different trials in Experiment 1. Error bars denote Standard Errors. "*" denotes $p = 0.011$.

Table 1
Accuracies and response times in Experiment 1. Standard errors are in parentheses, and response times are in milliseconds.

Dependent measure	Same	New auditory	New visual	New both
Accuracy				
Unimodal auditory	0.98 (0.01)	0.99 (0.01)		
Unimodal visual	0.95 (0.01)		0.96 (0.01)	
Cross-modal	0.94 (0.02)	0.94 (0.01)	0.98 (0.01)	0.99 (0.01)
Response times				
Unimodal auditory	746 (27)	803 (26)		
Unimodal visual	689 (22)		680 (21)	
Cross-modal	740 (24)	834 (19)	765 (21)	706 (22)

cross-modal presentation increased mean fixation durations, this effect was in the right direction but did not reach significance in the current study, $p = 0.12$. In summary, Experiment 1 provides evidence of auditory dominance in adults. Simultaneously presenting spoken words and visual images resulted in asymmetric costs, with words slowing down visual processing and delaying first fixations, while there appeared to be no negative cost of visual input on auditory processing.

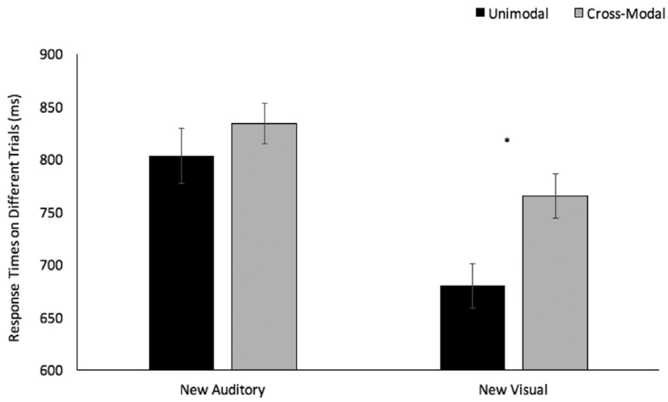

Fig. 4. Mean response times on different trials in Experiment 1. Error bars denote Standard Errors and "*" denotes $p < 0.001$.

Table 2
Eye tracking variables in Experiment 1. Means, Standard Errors, and Effect Sizes (η_p^2) comparing the unimodal visual and cross-modal visual conditions.

Dependent measure	Unimodal condition (SE)	Cross-modal condition (SE)	Effect size η_p^2
Mean fixation duration (ms)	272 (14)	296 (15)	0.06
Proportion looking to relevant	0.12 (0.01)	0.10 (0.01)	0.04
Number of fixations	1.20 (0.11)	1.21 (0.11)	0.00
Latency of first look (ms)*	332 (11)	374 (20)	0.12

Note: "*" denotes that repeated measures ANOVA, $p < 0.05$.

3. Experiment 2: Modality dominance across development

The primary goal of Experiment 2 was to examine how modality dominance effects change across development. To address this issue, we made minor modifications to the general procedure (discussed below) and examined how cross-modal presentation affected auditory and visual processing in children, young adults, and older adults. Based on previous research with children and young adults (Nava & Pavani, 2013; Robinson & Sloutsky, 2004; Sloutsky & Napolitano, 2003), it was expected that auditory dominance effects would decrease with age. Hypotheses underlying the effect of aging on modality dominance are less transparent. While numerous studies show that multisensory stimuli have larger effects in older adults (e.g., DeLoss et al., 2013; Laurienti et al., 2006)—possibly due to poorer filtering of incongruent multisensory stimuli (Lustig et al., 2007)—there is evidence suggesting that effects of auditory input on visual processing increase with age (e.g., DeLoss et al., 2013; Laurienti et al., 2006) and effects of visual input on auditory processing also appear to be more pronounced in older adults (Sekiyama et al., 2014; Van Gerven & Guerreiro, 2016).

3.1. Method

3.1.1. Participants, stimuli, and procedure

Nineteen children (M = 7.95 years, Range 5.94–12.91, 7 Females), 24 young adults (M = 19.25 years, Range 19.46–20.96, 13 Females), and 18 older adults (M = 77.22 years, Range 62.28–93.11, 10 Females) participated in Experiment 2. All children and young adults were run in a quiet room at The Ohio State University at Newark. Children were recruited by word of mouth and their guardians received a $10 gift card for their participation. Young adults were Introduction to Psychology students participating for course credit. Older adults were recruited by word of mouth and from a local adult living/recreational facilities. Approximately half of the older adults were tested in the laboratory at Ohio State Newark and the remaining older adults were tested in a quiet room at a Continuing Care Retirement Center (CCRC). All participants were independently-living community members. Those participants tested in the lab were recruited through word of mouth, whereas, those participants tested at the CRCC location were recruited through an existing partnership between the center and the University. All participants received a $10 gift card for their participation. The only criterion for participation was that participants had both hearing and vision that was considered normal or corrected to normal. Eight additional participants (four children, one young adult, and three older adults) were tested but excluded from the final data set. Two of the children were excluded due to a response bias and two were excluded due to poor eye tracking calibration. The young adult and two of the older adults were excluded due to poor calibration. An additional older adult was tested but excluded due to substantial hearing loss and a cochlear implant.

With the following exceptions, the design of Experiment 2 was identical to Experiment 1. First, we reduced the number of trials to make it easier for younger participants to complete the procedure. In Experiment 2, there were 32 unimodal trials (16 auditory and 16 visual) and 32 cross-modal trials (8 of each trial type). Second, we made it easier for participants to respond by using two, large StealthSwitch3 USB buttons to indicate Same or Different, as opposed to pressing 1 and 3 on the number pad. Third, the number of calibration and validation points used to ensure strength of eye tracking variables was reduced to five. Recall that in Experiment 1 we used a 9-point calibration. Finally, we included four practice trials at the beginning of each block. Practice trials were not included in the final analyses.

3.2. Results and discussion

3.2.1. Behavioral analyses

Proportion of correct responses on new auditory and new visual trials were submitted to a 3 Age (Children, Young Adults, vs. Older Adults) × 2 Modality (Auditory vs. Visual) × 2 Presentation (Unimodal vs. Cross-modal) mixed-factors ANOVA with age manipulated between subjects. The ANOVA revealed effects of Modality, F (1,58) = 4.18, $p = 0.045$, $\eta_p^2 = 0.07$, and Presentation, F (1,58) = 16.72, $p < 0.001$, $\eta_p^2 = 0.22$, and also revealed an Age x Modality interaction, F (2,58) = 5.39, $p = 0.007$, $\eta_p^2 = 0.157$. However, as can be seen in Fig. 5 (see also Table 3 for all means and standard errors), these effects were subsumed by an Age x Modality x Presentation interaction, F (2,58) = 5.39, $p = 0.007$, $\eta_p^2 = 0.157$. Modality dominance effects were examined by conducting separate Modality x Presentation ANOVAs for each age group. The analysis with children only revealed effects of Modality, F (1,18) = 9.07, $p = 0.007$, $\eta_p^2 = 0.34$, and Presentation, F (1,18) = 5.73, $p = 0.028$, $\eta_p^2 = 0.24$, with accuracies being higher in the unimodal and auditory conditions. For young adults, there was only an effect of Presentation, F (1,23) = 9.17, $p = 0.006$, $\eta_p^2 = 0.29$, with accuracy being higher in the unimodal condition. Thus, as with children, cross-modal presentation equally attenuated auditory and visual discrimination. The analysis with older adults revealed an effect of Presentation, F (1,23) = 9.17, $p = 0.006$, $\eta_p^2 = 0.29$ and the Modality × Presentation interaction was also significant, F (1,23) = 9.17, $p = 0.006$, $\eta_p^2 = 0.29$. As can be seen on the right side of Fig. 5, the three-way interaction is driven by older adults showing a different pattern, with cross-modal presentation decreasing auditory discrimination, t (17) = 3.98, $p < 0.001$, and having a marginally significant facilitation effect on visual processing, t (17) = − 1.94, $p = 0.069$. This asymmetric cost on auditory discrimination in older adults is consistent with visual dominance.

To examine how cross-modal presentation affected speed of processing, we submitted correct response times on new auditory and new visual trials to a 3 Age (Children, Young Adults, vs. Older Adults) × 2 Modality (Auditory vs. Visual) × 2 Presentation (Unimodal vs. Cross-modal) mixed-factors ANOVA. As in Experiment 1, response times greater than two standard deviations above the mean were removed

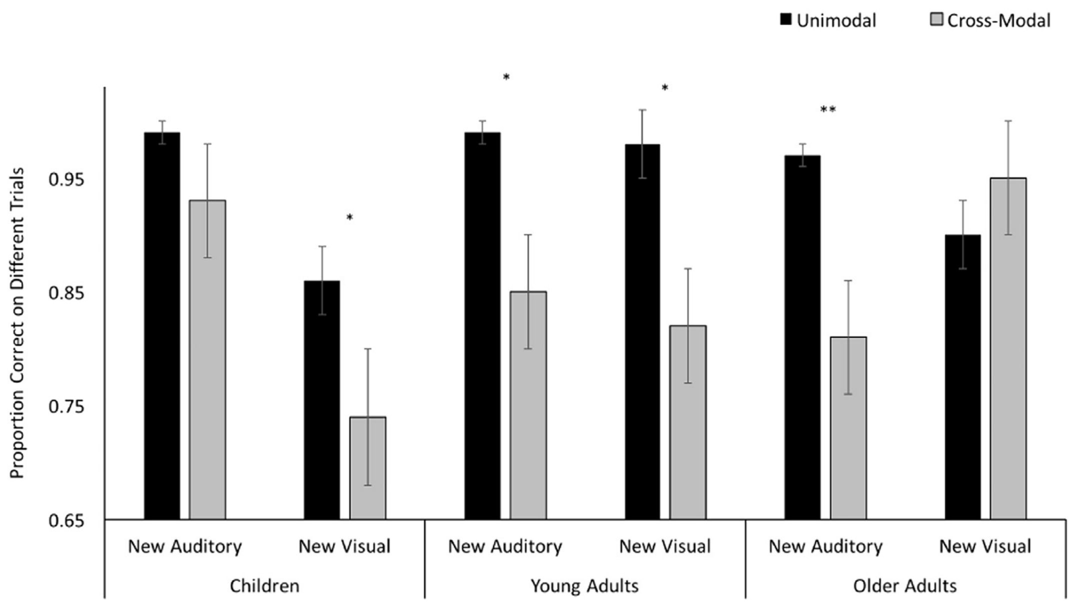

Fig. 5. Proportion correct on different trials in Experiment 2. Error bars denote Standard Errors. "*" denotes $p < 0.05$ and "**" denotes $p < 0.005$.

Table 3
Accuracies (and standard errors) in Experiment 2.

Age/condition	Same	New auditory	New visual	New both
Children				
Unimodal auditory	0.95 (0.02)	0.99 (0.01)		
Unimodal visual	0.89 (0.04)		0.86 (0.03)	
Cross-modal	0.95 (0.02)	0.93 (0.05)	0.74 (0.06)	0.98 (0.01)
Young adults				
Unimodal auditory	0.99 (0.01)	0.99 (0.01)		
Unimodal visual	0.95 (0.02)		0.98 (0.03)	
Cross-modal	0.97 (0.01)	0.85 (0.05)	0.82 (0.05)	0.98 (0.01)
Older adults				
Unimodal auditory	0.97 (0.01)	0.97 (0.01)		
Unimodal visual	0.88 (0.04)		0.90 (0.03)	
Cross-modal	0.94 (0.02)	0.81 (0.05)	0.95 (0.05)	0.99 (0.01)

Table 4
Response times in ms (and standard errors) in Experiment 2.

Age/condition	Same	New auditory	New visual	New both
Children				
Unimodal auditory	1364 (80)	1673 (84)		
Unimodal visual	1713 (146)		1559 (75)	
Cross-modal	1628 (109)	1771 (83)	1777 (77)	1675 (90)
Young adults				
Unimodal Auditory	976 (51)	1040 (76)		
Unimodal Visual	871 (47)		899 (68)	
Cross-modal	1004 (55)	1144 (75)	1058 (70)	1067 (48)
Older adults				
Unimodal auditory	937 (41)	997 (86)		
Unimodal visual	1197 (79)		1160 (77)	
Cross-modal	1076 (55)	1237 (85)	1052 (79)	957 (70)

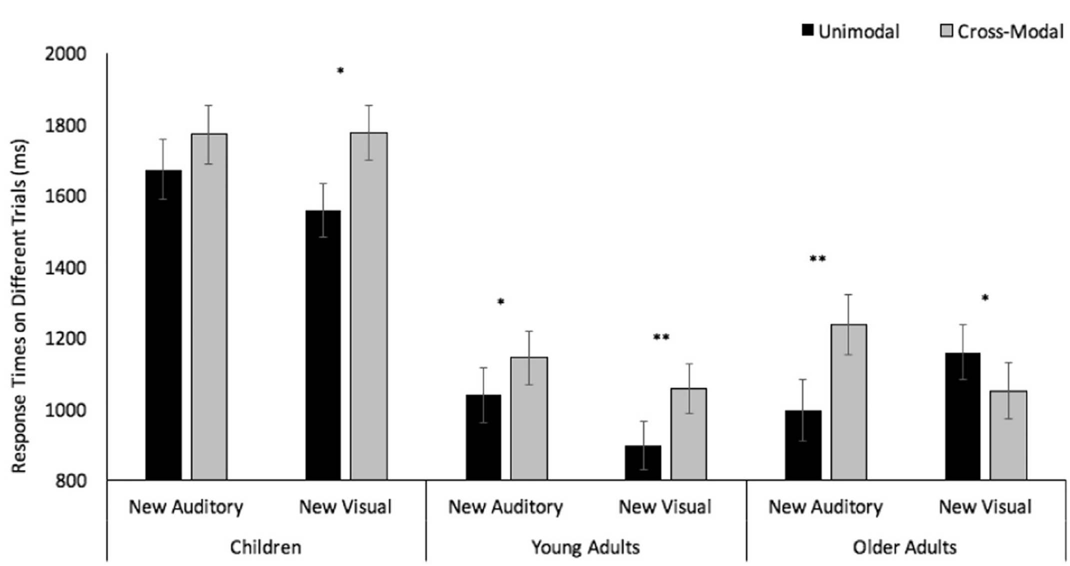

Fig. 6. Mean response times on different trials in Experiment 2. Error bars denote Standard Errors. "*" denotes $p < 0.05$ and "**" denotes $p < 0.005$.

from further analyses. Response times and Standard Errors are presented in Fig. 6 (see also Table 4). The analysis revealed main effects of Modality, F (1,57) = 7.62, $p = 0.008$, $\eta_p^2 = 0.118$, Presentation, F (1,57) = 26.53, ≤ 0.001, $\eta_p^2 = 0.318$, and Age, F (2,57) = 29.10, $p < 0.001$, $\eta_p^2 = 0.505$. Pairwise comparisons revealed that response times for all three age groups were significantly different, p's < 0.001, with young adults responding the fastest and children responding the slowest. The analysis also revealed a Modality x Age interaction, F (2,57) = 4.85, $p = 0.011$, $\eta_p^2 = 0.146$, and the Age x Modality x Presentation interaction was also significant, F (2,57) = 12.70, $p < 0.001$, $\eta_p^2 = 0.308$. Three separate Modality x Presentation ANOVAs were conducted for each age group. While there was some evidence that auditory input slowed down visual response times in children, t (18) = 3.96, $p < 0.001$, the ANOVA revealed no significant effects, p's > 0.10. The analysis in young adults only revealed effects of Modality, F (1,22) = 30.27, $p < 0.001$, $\eta_p^2 = 0.58$, and Presentation, F (1,22) = 12.37, $p = 0.002$, $\eta_p^2 = 0.36$, with response times being faster in the unimodal and visual conditions. Although the Modality x Presentation interaction did not reach significance, $p = 0.21$, effects of auditory input on visual processing ($\eta_p^2 = 0.35$) were almost twice as strong as effects of visual input on auditory processing ($\eta_p^2 = 0.19$). As with accuracy data, the three-way interaction was driven by older adults. As can be seen in Fig. 5, the ANOVA revealed a significant Modality x Presentation interaction, F (1,17) = 17.80, $p < 0.001$, $\eta_p^2 = 0.51$. While cross-modal presentation slowed down auditory responding, t (17) = 3.97, $p = 0.001$, it sped up visual responding, t (17) = −2.38, $p = 0.030$, a finding consistent with visual dominance.

Behavioral findings of Experiment 2 demonstrate weak evidence of auditory dominance in the younger groups with visual dominance effects more pronounced in older adults, however, it is important to note that young adults' accuracies differed across reported experiments. Recall that in Experiment 1, young adults' accuracies approached ceiling in all conditions, whereas cross-modal accuracies were lower in Experiment 2. We further examined this issue by looking at individual patterns in Experiment 2. In particular, it is possible that decreased discrimination in both modalities resulted from increased cognitive load, noisy sample, etc. If this is the case, then many of the participants should show a significant decrease in both auditory and visual accuracy. However, it is also possible that some participants prioritized visual processing (visual dominance with only decreased discrimination of auditory stimuli), while other participants prioritized auditory processing (auditory dominance with only decreased discrimination of visual stimuli).

To distinguish between these two possibilities, we categorized participants as auditory responders, visual responders, equal responders, or attenuated responders. If only visual accuracy in the cross-modal condition dropped by > 25% (compared to the unimodal baseline), then that participant was categorized as an auditory responder. If only auditory accuracy dropped by > 25%, then that participant was categorized as a visual responder. If neither modality dropped by > 25% or both modalities dropped by > 25%, then that participant was categorized as an equal responder or an attenuated responder, respectively. Seventeen of the young adults were categorized as equal responders, two as attenuated responders, three as auditory responders, and two as visual responders. When only focusing on equal responders, accuracy was > 95% across all conditions and none of the means differed, p's > 0.19. Furthermore, equal responders' response times replicated young adults in Experiment 1 with cross-modal presentation slowing down visual processing, t (16) = 2.29, $p = 0.036$, while having no significant effect on auditory processing, t (16) = 1.39, $p = 0.184$. Thus, when focusing only on participants with high overall accuracy, the finding is consistent with auditory dominance effects reported in Experiment 1.

We also used the same categorization scheme to classify kids and older adults as auditory, visual, equal, or attenuated responders. As can be seen in Table 5, most of the participants ably processed both modalities and were categorized as equal responders, and therefore there was not enough power to properly analyze how auditory and visual dominance changed with age. However, trending data were consistent with the notion that modality dominance differs across development from auditory to visual dominance as a function of age. More specifically, of the participants categorized as an auditory or visual responder, 100% of the children were classified as auditory responders, 60% of young adults were classified as auditory responders, and 100% of the older adults were categorized as visual responders.

Table 5
Responder types in Experiment 2. Values denote number of participants for each response type.

	Responder type			
Age	Auditory	Visual	Equal	Attenuated
Children	6	0	13	0
Young adults	3	2	17	2
Older adults	0	3	15	0

3.2.2. Eye tracking analyses

As in Experiment 1, we examined how cross-modal presentation affected the latency of first fixations (to any location on the visual stimulus), fixation durations, proportion looking to relevant AOIs, and number of fixations. The four dependent variables were submitted to separate 3 Age (Children, Young Adults, vs. Older Adults) × 2 Presentation (Unimodal New Visual vs. Cross-modal New Visual) mixed-factors ANOVA's. See Table 6 for means, standard errors, and effect sizes broken down by age and dependent measure.

The first analysis focused on latency of first fixation, revealing only an effect of Presentation, F (1,56) = 8.67, $p = 0.005$, $\eta_p^2 = 0.134$. Consistent with Experiment 1, cross-modal presentation delayed the onset of first fixations, with the first fixation being delayed in the cross-modal condition (M = 410 ms, SE = 29) compared to the unimodal condition (M = 334 ms, SE = 14). The second analysis focused on mean fixation durations. The analysis revealed a marginally significant effect of Presentation, F (1,56) = 3.24, $p = 0.077$, $\eta_p^2 = 0.055$ and a marginally significant Presentation x Age interaction, F (2,56) = 2.58, $p = 0.084$, $\eta_p^2 = 0.084$. The analysis also revealed an effect of Age, F (2,56) = 10.56, $p < 0.001$, $\eta_p^2 = 0.274$. Pairwise comparisons revealed that older adults made shorter fixations (M = 247 ms, SE = 22)

Table 6
Eye tracking variables and statistics in Experiment 2. Means, Standard Errors, and Effect Sizes (η_p^2) comparing the unimodal visual and cross-modal visual conditions.

Dependent measure	Unimodal condition (SE)	Cross-modal condition (SE)	Effect size η_p^2
Mean fixation duration (ms)			
Children	361 (25)	411 (28)	0.11
Young adults*	322 (24)	370 (29)	0.22
Older adults	258 (19)	236 (16)	0.08
Proportion looking to relevant			
Children	0.13 (0.02)	0.14 (0.02)	0.00
Young adults	0.12 (0.02)	0.10 (0.02)	0.02
Older adults	0.14 (0.02)	0.14 (0.02)	0.00
Number of fixations			
Children	4.05 (0.84)	3.55 (0.42)	0.02
Young adults	1.68 (0.24)	1.78 (0.25)	0.02
Older adults	3.23 (0.48)	2.55 (0.23)	0.10
Latency of first look (ms)			
Children*	306 (21)	404 (34)	0.31
Young adults+	336 (23)	448 (65)	0.15
Older adults	359 (29)	378 (32)	0.03

Note: "*" denotes that repeated measures ANOVA, $p < 0.05$ and "+" denotes $p = 0.069$.

than children (M = 386 ms, SE = 22) and young adults (M = 346 ms, SE = 20), p's $<$ 0.005. The third analysis focused on the percent looking to relevant AOI's, which yielded no significant effects. Finally, we examined how cross-modal presentation affected the number of fixations. The analysis revealed a main effect of Age, F (2,58) = 8.72, $p < 0.001$, $\eta_p^2 = 0.231$. Pairwise comparisons revealed that children (M = 3.80 ms, SE = 0.37) made more fixations than young adults (M = 1.73 ms, SE = 0.33), $p < 0.001$, and older adults (M = 2.89 ms, SE = 0.38) made marginally more fixations than younger adults, p = 0.078.

Overall, the child and young adult data provide some support for the claim that auditory input automatically engage attention and delay processing of visual information. However, it is unclear if this slowdown stems from the conflicting information (auditory input eliciting a same response and visual input eliciting a new response) or from any auditory stimulus engaging attention and increasing task demands and/or delaying visual processing. To address this issue, we compared the unimodal visual baseline conditions to cross-modal trials where both modalities elicited a new response (see Both New trials in Figs. 3 and 5). If the cross-modal conflict is driving the slowdown, then slower response times and delayed first fixations should only occur on New Visual trials where a new visual stimulus was paired with the old label. However, if any auditory stimulus engages attention, then there should also be a slowdown on Both New trials.

As can be seen below, the results are mixed. In Experiment 1, there was no evidence that changing both the pictures and words slowed down response times (p = 0.125) and first fixations were not delayed on Both New trials (p = 0.153) compared to the unimodal visual baseline. Thus, it appears that the conflicting auditory and visual information was responsible for the slowdown in Experiment 1. However, Experiment 2 revealed a different pattern. As can be seen in Fig. 5, young adults were significantly slower to respond when both word and picture changed (M = 1067 ms, SE = 48) compared to the unimodal visual baseline (M = 897 ms, SE = 47), t (23) = 5.02, $p < 0.001$, and young adults' first fixations on Both New trials (M = 407 ms, SE = 38) were also slower than the unimodal baseline (M = 312 ms, SE = 25), t (22) = 2.80, p = 0.011. It is not clear why the current study found a different pattern of results across experiments, however, Experiment 2 replicates Dunifon et al. (2016; Experiments 1 and 3) where response times to Both New items were also significantly slower. Although response times to Both New items were not significantly slower for children, p = 0.23, children's first fixations on Both New trials (M = 404 ms, SE = 48) were slower than the unimodal baseline (M = 306 ms, SE = 21), t (18) = 2.25, p = 0.037. Thus, while the findings are mixed, Experiment 2 provides some support for the claim that auditory stimuli may automatically engage attention and delay the onset of visual processing, as opposed to the slowdown resulting solely from the increased processing demands of processing competing auditory and visual information.

Finally, while young children and adults were sometimes slower to respond when both modalities changed at test, older adults showed a different pattern. For example, as can be seen in Fig. 5, older adults were significantly faster to respond when both modalities changed on Both New trials (M = 1022 ms, SE = 42) compared to the unimodal baseline (M = 1160 ms, SE = 60), t (17) = 3.84, $p < 0.001$. It is important to note that older adults were also faster at reporting visual changes when there was a conflict between auditory and visual information (i.e., see right side of Fig. 5 where response times on New visual trials were also faster than the unimodal visual baseline). This suggests that the facilitation effect was not driven by multisensory congruency (auditory and visual stimuli both being associated with a new response), but rather, from a more general auditory effect (e.g., increased arousal in multisensory conditions). However, this facilitation effect was only evident for processing of visual information, as auditory response times and accuracies both decreased under multisensory presentation.

4. General discussion

Many tasks require processing of multisensory information which can often result in modality dominance, with stimuli in one sensory modality disrupting or overshadowing processing in a second modality. The present study examined how modality dominance effects differ across development, testing children, young adults, and older adults. In Experiment 1, young adults were presented with either two words, two pictures, or two word-picture pairings and they had to determine if the two stimuli/pairings were exactly the same or different. Pairing the words and pictures together slowed down visual but not auditory processing (compared to respective unimodal baselines), which is consistent with reported auditory dominance effects typically found in younger populations (Nava & Pavani, 2013; Robinson & Sloutsky, 2004; Sloutsky & Napolitano, 2003). Latency of first fixations were also delayed when visual images were paired with spoken, nonsense words. This finding suggests that auditory input might be disrupting encoding of visual input. Moreover, these effects are not restricted to spoken words, as nonlinguistic sounds appear to have a similar effect on visual processing (Dunifon et al., 2016).

Experiment 2 utilized the same methodology to examine how modality dominance effects differ across children, young adults, and older adults. While cross-modal presentation equally attenuated auditory and visual accuracy for children and young adults, cross-modal presentation increased visual accuracy and decreased auditory accuracy in older adults - a pattern consistent with visual dominance. Analyses of response times showed a similar pattern, with a shift from auditory to visual dominance across development. More specifically, as can be seen in Fig. 5, pairing pictures and words together had stronger interference effects on children's and young adults' visual response times than auditory response times. In contrast, pairing words and pictures together sped up visual response times in older adults; however, prioritized visual processing came with a cost - slower responding to auditory information. Finally, as in Experiment 1, latency of first fixations were also delayed when images were paired with words compared to when images were presented in silence; however, these effects only reached significance for children and young adults (see Table 3).

Finding auditory dominance in an adult population is rare, especially when considering decades of research pointing to visual dominance (Colavita, 1974; Colavita et al., 1976; Colavita & Weisberg, 1979; Egeth & Sager, 1977; Koppen et al., 2008; Ngo et al., 2010; Ngo et al., 2011; Sinnett et al., 2007; Sinnett et al., 2008). We believe there are several factors that may underlie discrepancies across reported studies. First, while previous research using a similar task found no evidence of auditory dominance in young adults (e.g., Sloutsky & Napolitano, 2003), the present study did not solely investigate accuracies as the determinant of modality dominance. We also examined more sensitive measures such as response times and visual fixations. Second, many studies supporting visual dominance rely on the Colavita visual dominance task (Colavita, 1974) where participants make speeded, modality-specific responses to auditory and visual stimuli. This makes it possible for participants to bias their responding in favor of one modality. The present study attempted to circumnavigate a potential response bias by requiring participants to respond "same" or "different". These same/different responses are modality-independent as participants were required to respond "different" when either auditory or visual stimuli changed at test. We believe this manipulation best accounts for differences between the present study and visual dominance research. Additional support for this claim comes from Robinson et al.'s. (2016) cross-modal oddball task where auditory dominance was found when participants made the same response to changing auditory and visual oddballs (similar response demands to present task), and visual dominance was found when participants made separate responses to changing auditory, visual, and cross-modal oddballs.

Based on a potential mechanism underlying auditory dominance which posits that auditory stimuli automatically engaging attention and attenuate visual processing (Robinson & Sloutsky, 2010), it was hypothesized that pairing the pictures with words would slow down processing of the visual stimulus and have little to no negative effect on auditory processing. Moreover, it was also hypothesized that eye tracking variables such as latency of first fixation may also account for slower response times in cross-modal conditions, particularly when visual stimuli were paired with words. Children and adults' response time and eye tracking data in both of the reported experiments support this claim.

While additional research is needed, we believe these interference effects are happening early in course of processing, as opposed to interfering with subsequent processing during the response or decision phase. Support for this claim comes from several sources. First, previous research has shown that pairing sounds and pictures together slowed down visual but not auditory P300s in a passive oddball task (Robinson, Ahmar, & Sloutsky, 2010). It is important to note that participants in this study passively viewed auditory and visual oddballs; thus, these delays occurred even though participants did not make any decisions or responses to oddballs. Second, while children and adults have top-down oculomotor control, many fixations occur in a bottom-up, automatic manner. Finding delays in first fixations likely stemmed from automatic processing rather than a deliberate choice to withhold first fixations to an incoming visual stimulus. Finally, some (but not all) of these delays occur even when both auditory and visual modalities are providing complementary messages. For example, on some of the trials in the present study, both auditory and visual modalities changed and were associated with the same response. Adults in Experiment 2 and in Dunifon et al. (2016) were also slower to respond to these trials, and latency of first fixations in children and adults were also delayed (Experiment 2). These findings are consistent with the claim that auditory stimuli automatically engage attention and attenuate visual encoding, and they also suggest that modality dominance effects may not require direct competition between auditory and visual input (e.g., auditory stimulus elicits a same response and visual input requires a different response).

Numerous studies have examined the developmental trajectory of modality dominance in infants, children, and young adults (Lewkowicz, 1988a, 1988b; Nava & Pavani, 2013; Robinson & Sloutsky, 2004; Sloutsky & Napolitano, 2003). These studies typically show that infants and young children often exhibit auditory dominance with these effects decreasing or reversing (visual dominance) with age. The present findings are somewhat consistent with this research. While children and adults showed comparable decreases in auditory and visual accuracy when images were paired with words, pairing pictures and words together appeared to have a greater cost on visual response times and latency of first fixations were also delayed when images were paired with words and sounds (Dunifon et al., 2016). Thus, while we replicated previous research and found evidence of auditory dominance, we did not see a decrease in this effect into young adulthood. As previously mentioned, we believe the extended auditory dominance into adulthood stems from using more sensitive methodologies and removing modality-specific response demands, both of which are used in more traditional visual dominance tasks.

While there is a growing body of research examining the development of modality dominance across infancy, childhood, and early adulthood, we are not aware of any research examining how aging affects modality dominance. We believe that including a sample of older adults is the most novel and important aspect of the present study, as there are many sensory, motor, and cognitive changes that occur in late adulthood (see Birren & Schaie, 2006 for a review) and it is unclear how these changes affect processing of simultaneously presented auditory and visual information. We reviewed proposed mechanisms underlying auditory and visual dominance, and given the findings of the present study, it is doubtful that a single mechanism can appropriately account for the patterns of modality dominance unless it includes maturation or development as a factor. Changes in modality dominance in older adults could stem from numerous accounts: (a) general increase in multisensory integration, possibly to compensate for decreases in unimodal processing (DeLoss et al., 2013; Laurienti et al., 2006); (b) the visual sensory system inhibiting processing in other sensory modalities (Desimone & Duncan, 1995; Duncan, 1996; Spence et al., 2012), with these inhibition effects strengthening across development; or (c) attentional factors, with older participants having more difficulty filtering incongruent cross-modal information (Lustig et al., 2007).

While increased multisensory integration (DeLoss et al., 2013; Laurienti et al., 2006) and the inhibitory deficit hypothesis (Lustig et al., 2007) may differ in underlying causes (e.g., increased sensory interactions or poor attentional filtering, respectively), they both make the same predictions and only partially explain the findings of the present study. In particular, both accounts posit that older adults should show stronger facilitation effects when presented with congruent information because they are more likely to notice the cross-modal stimuli (inhibitory deficit hypothesis) and/or from better integration of simultaneously presented auditory and visual information (multisensory integration). We found support for these accounts: older adults were the only group to show facilitation effects in Experiment 2 (see Both New trials in Table 4). At the same time, increased intersensory interactions and/or poor filtering should result in stronger interference effects when modalities are presented with conflicting information (New Auditory and New Visual trials). We failed to find symmetrical interference effects in older adults, and in fact, older adults seemed to prioritize visual processing by showing faster and more accurate responses when presented with congruent and incongruent cross-modal information. Thus, multisensory integration/inhibitory deficit hypothesis accounts do not appear to completely account for the findings in older adults. An important consideration, however, is that much of the data from multisensory integration/inhibitory deficit hypothesis accounts come from procedures using selective attention tasks (e.g., ignore the beeps and report how many flashes you see), whereas data from the present study were collected with use of an immediate recognition task where participants had to divide their attention across sensory modalities.

A recently proposed mechanism underlying visual dominance suggests that sensory modalities inhibit each other (Desimone & Duncan, 1995; Duncan, 1996; Spence et al., 2012), and vision should be more likely to inhibit processing in other modalities because approximately 50% of the brain is dedicated to vision (Spence et al., 2012). This mechanism can account for prioritized visual processing–even when presented with incongruent cross-modal stimuli–and it may also explain the changes from young to older adults, assuming the visual sensory system undergoes a long, protracted development with visual inhibition continuing into late adulthood.

Additional support for increased sensory inhibition comes with work examining speech perception in older adults (Sekiyama et al., 2014). More specifically, Sekiyama et al. demonstrated that older adults show a heightened McGurk effect and subsequently rely on visual cues when presented with incongruent cross-modal information. Coupled with this research is the finding that older adults suffer when presented with a visual distractor on a cross-modal selective attention task (Van Gerven & Guerreiro, 2016). More specifically, such visual distractors delayed processing in the auditory modality in older adults, and when reversed, auditory distractors failed to delay processing in the visual modality. This asymmetrical cost highlights the potential impact of a visual system with stronger inhibitory connections, a system that has a developmental basis because such findings were not found in younger adults (Van Gerven & Guerreiro, 2016).

4.1. Limitations and future directions

The present study had several limitations regarding participant recruitment and stimulus selection/presentation. In regard to the former, we only screened children, young adults, and older adults on vision and hearing with a self-reported assessment. We did have a check to ensure normal vision and hearing (performance on unimodal visual and auditory trials); however, future research would be enhanced with formal screening protocol. Second, we did not formally screen for cognitive ability (e.g., IQ), cognitive decline, drug use, and/or psychiatric/neurological presentation, all factors that could have influenced findings. Third, while the average age of the children sampled in Experiment 2 was similar to the youngest group in Nava and Pavani (2013), children were older than those sampled in previous work utilizing a task similar to the one reported here (Sloutsky & Napolitano, 2003). As such, auditory dominance effects may be more pronounced in a younger population and future research replicating these effects should attempt to better capture age criteria outlined in previous developmental work. Finally, Experiment 2 shows different patterns of modality dominance across development. While this suggests that modality dominance shifts from auditory to visual dominance with age, future research will need to properly test this using longitudinal design.

There are also several stimulus-related issues, which should be considered in future research. First, there was nothing on the screen for participants to fixate on during unimodal auditory trials; therefore, there was no control to examine effects of cross-modal presentation on auditory processing. That said, Robinson et al. (2010) found no negative effects on auditory P300s and the present study found no delay in auditory response times in young adults (young adults) and children (Experiment 2), suggesting that cross-modal presentation is less likely to interfere with auditory processing. Second, the visual Test image remained visible until participants made a response while the auditory Test stimulus was only presented for the stimulus duration (approximately 700 ms), as opposed to stretching or repeating the word. It is doubtful that this manipulation affected findings of Experiment 1, as average responses were prior to or right around the same time as auditory offset, but this manipulation could have potentially had an effect in Experiment 2 as some of the mean response times occurred after the auditory stimulus disappeared. Third, visual stimuli in the present study were additive in nature (e.g., Tree 1 and 2 differed by adding two new features), whereas differences between words were more qualitative in nature. While manipulating stimulus features/dimensions that distinguish stimuli will impact discrimination and processing demands, using spatially distinct visual features was necessary for the present eye tracking experiment. Additional research needs to examine how stimulus manipulations such as these affect patterns of visual fixations; however, the behavioral findings appear to generalize to less meaningful stimuli that are not additive in nature (Robinson et al., 2016) and when using auditory stimuli that are additive (Parker & Robinson, 2017).

Another important component of the methodology was that we primarily focused on discrimination of auditory and visual information when presented cross-modally (i.e., same auditory/different visual and different auditory/same visual). We assumed that different responses on these trials stemmed from participants correctly detecting a change; however, it is also possible that they made a "different" response because they made an error on the other modality (e.g., participants correctly responded different on New Visual trials because they thought the auditory stimulus changed at test but it was identical to the first stimulus). While we cannot distinguish between these two possibilities in the present study, future research will need to systematically manipulate the testing phase to ensure that correct responses on New Auditory and New Visual trials stem from participants correctly identifying the changed component of the auditory-visual test item. Finally, in the present study, we sequentially presented auditory, visual, and cross-modal stimuli, thus requiring sensory/working memory to store the first item and then compare the currently presented stimulus with a representation of the first stimulus. To eliminate possible effects resulting from sensory modality differences in storage and retrieval, future research should also examine tasks that require processing of simultaneously presented information.

5. Conclusions

The present study examined the potential for differences in modality dominance effects across development. We found evidence for auditory dominance in children and young adults, a finding that runs contrary to a large body of research pointing to visual dominance. Pairing words and pictures together had a greater cost on visual processing than auditory processing (compared to respective unimodal baselines), which is consistent with reported auditory dominance effects typically found in younger populations. Additionally, the present study is one of the first to examine modality dominance effects in older adults, highlighting three considerable points: a) auditory dominance effects present in children and young adults differ as a function of age, with older adults showing clear evidence of visual dominance; b) older adults were the only group to benefit from cross-modal presentation, but only in the visual modality; and c) older adults also showed faster and more accurate responses when presented with conflicting auditory information, a finding that suggests the mere presence of the words increased performance possibly due to an overall increase in arousal. These findings have significant implications for real-world tasks that hinge on responding to and processing multisensory input, highlighting that auditory and visual input appear to function differently mechanistically across development.

References

Bahrick, L. E., Lickliter, R., & Flom, R. (2004). Intersensory redundancy guides the development of selective attention, perception, and cognition in infancy. *Current Directions in Psychological Science, 13*, 99–102.

Birren, J. E., & Schaie, K. W. (Eds.). (2006). *Handbook of the psychology of aging*(6th ed.). San Diego, CA: Elsevier.

Colavita, F. B. (1974). Human sensory dominance. *Perception & Psychophysics, 16*, 409–412.

Colavita, F. B., Tomko, R., & Weisberg, D. (1976). Visual pre-potency and eye orientation. *Bulletin of the Psychonomic Society, 8*, 25–26.

Colavita, F. B., & Weisberg, D. (1979). A further investigation of visual dominance. *Attention, Perception & Psychophysics, 25*, 345–347.

Corso, J. F. (1971). Sensory processes and age effects in normal adults. *Journal of Gerontology, 26*, 90–105.

Craik, F. (1994). Memory changes in normal aging. *Current Directions in Psychological Science, 3*, 155–158.

DeLoss, D., Pierce, R., & Andersen, G. (2013). Multisensory integration, aging, and the sound-induced flash illusion. *Psychology and Aging, 28*, 802–812.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193–222.

Duncan, J. (1996). Cooperating brain systems in selective perception and action. In T. Inui, & J. L. McClelland (Eds.). *Attention and performance XVI: Information integration in perception and communication* (pp. 549–578). Cambridge, MA: The MIT Press.

Duncan, J., Martens, S., & Ward, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature, 387*, 808–810.

Dunifon, C., Rivera, S., & Robinson, C. W. (2016). Auditory stimuli automatically grab attention: Evidence from eye tracking and attentional manipulations. *Journal of Experimental Psychology: Human Perception and Performance, 42*, 1947–1958.

Egeth, H. E., & Sager, L. C. (1977). On the locus of visual dominance. *Attention, Perception & Psychophysics, 22*, 77–86.

Eimer, M., & Driver, J. (2000). An event-related brain potential study of cross-modal links in spatial attention between vision and touch. *Psychophysiology, 37*(05), 697–705.

Eimer, M., & van Velzen, J. (2002). Crossmodal links in spatial attention are mediated by supramodal control processes: Evidence from event- related brain potentials. *Psychophysiology, 39*, 437–449.

Fort, A., Delpuech, C., Pernier, J., & Giard, M. H. (2002). Dynamics of cortico-subcortical cross-modal operations involved in audio–visual object recognition in humans. *Cerebral Cortex, 12*, 1031–1039.

Giard, M. H., & Peronnet, F. (1999). Auditory–visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience, 11*, 473–490.

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 renewed view. *Psychological Bulletin, 136*, 975–1022.

He, N., Dubno, J., & Mills, J. (1998). Frequency and intensity discrimination measured in a maximum-likelihood procedure from young aged normal-hearing subjects. *Journal of the Acoustical Society of America, 103*, 553–565. http://dx.doi.org/10.1121/1.421127.

Koppen, C., Alsius, A., & Spence, C. (2008). Semantic congruency and the Colavita visual dominance effect. *Experimental Brain Research, 184*, 533–546.

Koppen, C., & Spence, C. (2007a). Assessing the role of stimulus probability on the Colavita visual dominance effect. *Neuroscience Letters, 418*, 266–271.

Koppen, C., & Spence, C. (2007b). Audiovisual asynchrony modulates the Colavita visual

dominance effect. *Brain Research, 1186*, 224–232.
Laurienti, P., Burdette, J., Maldjian, J., & Wallace, M. (2006). Enhanced multisensory integration in older adults. *Neurobiology of Aging, 27*, 1155–1163.
Lewkowicz, D. J. (1988a). Sensory dominance in infants: 1. Six-month-old infants' response to auditory-visual compounds. *Developmental Psychology, 24*, 155–171.
Lewkowicz, D. J. (1988b). Sensory dominance in infants: 2. Ten-month-old infants' response to auditory-visual compounds. *Developmental Psychology, 24*, 172–182.
Lustig, C., Hasher, L., & Zacks, R. (2007). Inhibitory deficit hypothesis: Recent developments in a “new view”. In C. M. MacLeod, & D. S. Gorfein (Eds.). *Inhibition in cognition* (pp. 145–162). Washington, DC: American Psychological Association.
Massaro, D. W. (1984). Children's perception of visual and auditory speech. *Child Development, 55*, 1777–1788. http://dx.doi.org/10.2307/1129925.
McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*, 746–748.
Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology, 14*, 247–279.
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*, 1850–1870.
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.
Ngo, M. K., Cadieux, M. L., Sinnett, S., & Soto-Faraco, S. (2011). Reversing the Colavita visual dominance effect. *Experimental Brain Research, 214*(4), 607–618.
Ngo, M. K., Sinnett, S., Soto-Faraco, S., & Spence, C. (2010). Repetition blindness and the Colavita effect. *Neuroscience Letters, 480*, 186–190.
Parker, J. L., Robinson, C. W., Gunzelmann, G., Howes, A., Tenbrink, T., & Davelaar, E. (2017). Auditory and visual contributions to multisensory integration. *Proceedings of the 39th annual conference of the cognitive science society* (pp. 2858–2863). .
Pavani, F., Husain, M., Ládavas, E., & Driver, J. (2004). Auditory deficits in visuospatial neglect patients. *Cortex, 40*(2), 347–365.
Posner, M. I., Nissen, M. J., & Klein, R. M. (1976). Visual dominance: An information-processing account of its origins and significance. *Psychological Review, 83*, 157–171.
Pratt, J., Chasteen, A. L., & Abrams, R. A. (1994). Rapid aimed limb movements: Age differences and practice effects in component submovements. *Psychology and Aging, 9*, 325–334.
Robinson, C. W., Ahmar, N., & Sloutsky, V. M. (2010). Evidence for auditory dominance in a passive oddball task. In S. Ohlsson, & R. Catrambone (Eds.). *Proceedings of the 32nd annual conference of the cognitive science society* (pp. 2644–2649). Austin, TX: Cognitive Science Society.
Robinson, C. W., Chandra, M., & Sinnett, S. (2016). Existence of competing modality dominances. *Attention, Perception, & Psychophysics, 78*, 1104–1114. http://dx.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*, 1387–1401.
Robinson, C. W., & Sloutsky, V. M. (2010). Development of cross-modal processing. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(1), 135–141.
Robinson, C. W., & Sloutsky, V. M. (2013). When audition dominates vision: Evidence from cross-modal statistical learning. *Experimental Psychology, 60*(2), 113–121.
Royall, D. R., Palmer, R., Chiodo, L. K., & Polk, M. J. (2004). Declining executive control in normal aging predicts change in functional status: The freedom house study. *Journal of the American Geriatrics Society, 52*, 346–352.
Rutschmann, J., & Link, R. (1964). Perception of temporal order of stimuli differing in sense mode and simple reaction time. *Perceptual and Motor Skills, 18*, 345–352.
Schneider, B. A., Daneman, M., Murphy, D. R., & Kwong, S. (2000). Listening to discourse in distracting settings: The effects of aging. *Psychology and Aging, 15*, 110–125.
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*, 1–12.
Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., & Tootell, R. B. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science, 268*, 889–893.
Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions: What you see is what you hear. *Nature, 408*, 788. http://dx.doi.org/10.1038/35048669.
Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research, 14*, 147–152.
Sinnett, S., Soto-Faraco, S., & Spence, S. (2008). The co-occurrence of multisensory competition and facilitation. *Acta Psychologica, 128*, 153–161.
Sinnett, S., Spence, C., & Soto-Faraco, S. (2007). Visual dominance and attention: Revisiting the Colavita effect. *Perception & Psychophysics, 69*, 673–686.
Sloutsky, V. M., & Napolitano, A. (2003). Is a picture worth a thousand words? Preference for auditory modality in young children. *Child Development, 74*, 822–833.
Sloutsky, V. M., & Robinson, C. W. (2008). The role of words and sounds in visual processing: From overshadowing to attentional tuning. *Cognitive Science, 32*, 354–377.
Spence, C. (2009). Explaining the Colavita visual dominance effect. *Progress in Brain Research, 176*, 245–258.
Spence, C., & Driver, J. (2004). *Crossmodal space and crossmodal attention.* Oxford: Oxford University Press.
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. 529–556). Boca Raton, FL: CRC Press.
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.
Weale, R. A. (1975). Senile changes in visual acuity. *Transactions of the Ophthalmological Societies of the United Kingdom, 95*, 36–38. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1064207.
Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin, 88*, 638–667.
Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman, & R. Davies (Eds.). *Varieties of attention* (pp. 63–101). New York: Academic Press.
Wille, C., & Ebersbach, M. (2016). Semantic congruency and the (reversed) Colavita effect in children and adults. *Journal of Experimental Child Psychology, 141*, 23–33.