Title

Auditory and Visual Contributions to Multisensory Integration

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 39(0)

Authors
Parker, Jessica L.
Robinson, Christopher W.

Publication Date
2017

Peer reviewed

Auditory and Visual Contributions to Multisensory Integration

Jessica L. Parker (parker.1026@osu.edu)
Department of Psychology, The Ohio State University at Newark
1179 University Dr., Newark, OH 43056, USA

Christopher W. Robinson (robinson.777@osu.edu)
Department of Psychology, The Ohio State University at Newark
1179 University Dr., Newark, OH 43056, USA

Abstract

Multisensory integration, or the merging of information from multiple sensory modalities, is important for many everyday tasks. One methodology used for examining this process is the Sound Induced Flash Illusion (SIFI), which presents participants with a number of flashes and either the same number of beeps (congruent) or a different number of beeps (incongruent), and requires the participant to respond by entering how many flashes they saw. The study expands on this research by examining the relative contributions of auditory and visual information on multisensory integration. While congruent and incongruent auditory stimuli affected visual perception (Experiment 1), there was little evidence that visual input affected auditory processing (Experiment 2). These findings support auditory dominance and modality appropriate hypothesis in adult populations and have implications on tasks that require integration across auditory and visual modalities.

Keywords: Crossmodal Processing; Multisensory Integration; Modality Dominance; Sound Induced Flash Illusion

Introduction

A majority of our daily experiences require people to process multisensory information. As a person walks down the street, for example, they may see a car driving by, hear the engine as it approaches, smell the exhaust and feel the breeze as the car passes. How information from the different sensory modalities is integrated and combined into a unitary percept is considered multisensory integration (Shams, 2000). Using the example above, these multisensory experiences are perceived as one percept (car) instead of having independent experiences. Given the evident impact of multisensory integration in our everyday experiences, it is important to understand the contributions of auditory and visual contribution to multisensory integration and factors that facilitate and inhibit multisensory integration.

Shams, Kamitani, and Shimojo (2000) developed a test of multisensory integration called the Sound Induced Flash Illusion (SIFI), where the number of beeps influences how many flashes people see. In their study, they presented one, two, or three flashes, and in the cross-modal condition, these flashes were paired with one, two or three beeps. Participants were then asked to report how many flashes they saw, regardless of how many beeps they heard. If the number of beeps exceeded the number of flashes, participants tended to overestimate the number of flashes (fission). If the number of beeps was less than the number of flashes, participants tended to underestimate flashes (fusion). Fission and fusion responses are implications of multisensory integration. This shows that the auditory information is being integrated with the visual information. Another study, using a different procedure and stimuli, tested perception of beeps to see if they could create a flash-induced sound illusion (Andersen, 2004). Under normal intensity levels, the visual flashes had no effect on auditory perception; however, they did find some evidence that flashes influenced beep perception when intensity levels of beeps were weakened to near threshold levels. This finding, in conjunction with Shams et al., may suggest that auditory information has a stronger effect on visual processing than vice versa; however, there were also numerous differences across studies; thus, making it difficult to make strong conclusions.

Why do auditory beeps affect participants' perception of the number of visual flashes? One possible explanation underlying this illusion is auditory dominance (Robinson & Sloutsky, 2010a). According to this account, auditory and visual stimuli compete for attentional resources; thus, increased attention to one modality might come with a cost - delayed or attenuated processing in the other modality. Moreover, because auditory stimuli are dynamic and transient, it may be adaptive to first allocate attention to the auditory modality before the information disappears. Most of the supporting research for auditory dominance comes from the developmental literature, where multisensory presentation attenuates visual processing more than auditory processing (Lewkowicz, 1988a; 1988b; Robinson & Sloutsky, 2004; 2010b; Sloutsky & Napolitano, 2003; Sloutsky & Robinson, 2008). According to this account, the beeps in SIFI may interfere with processing of the visual flashes; whereas, the visual input may have little effect on processing of the beeps. Another possible reason why beeps may affect flash perception is modality appropriateness hypothesis, which states that the modality that is more appropriate for the task is the one that dominates (Welch & Warren, 1980). Welch and Warren (1980) describe that with information processing, vision is dominant in spatial situations and audition is dominant for temporal judgements. When these two modalities are simultaneously presented and the task has a temporal aspect, studies have shown that audition becomes the dominant modality and can influence vision (Wada, Kitagawa, & Noguchi, 2003). Thus, both auditory dominance and modality appropriateness predict that auditory input should have a greater effect on visual perception than vice versa, especially when the task is temporal in nature (but see Tsay, 2013, where visual information affected judgements about musical performance).

Predicting that auditory information will have a greater effect on visual processing conflicts with much of the past research with adults that showed visual dominance, where the simultaneous presentation of auditory and visual information seems to inhibit auditory processing (see Sinnett, Spence, & Soto-Faraco, 2007, and Spence, Parise, & Chen, 2012). For example, when adults were required to press one button when they detected a visual stimulus, a different button when they detected an auditory stimulus, and a third button (or both buttons) when both stimuli were presented at the same time, participants often made errors on cross-modal trials by only pressing the visual button (Colavita, 1974). Thus, visual dominance tends to occur when adults are required to make speeded, modality-specific responses to auditory, visual, and crossmodal stimuli (Colavita, 1974; see Sinnett, Spence, & Soto-Faraco, 2007 for review). One possible explanation for visual dominance is that adults may have a visual response bias to compensate for the fact that visual input is less alerting than auditory (Posner et al., 1974). It is important to note that the current study testing the SIFI is different from some of the modality dominance studies because it requires quantity judgements (how many beeps or flashes), rather than requiring speeded, modality-specific responses to auditory or visual input.

The current study used a modified SIFI task to test both auditory and visual processing and expands previous research in three ways. First, the study expands SIFI research by examining the relative contributions of the auditory and visual modalities in multisensory integration, as opposed to only examining the effects of auditory input on multisensory integration or visual input on multisensory integration. Second, the current study expands SIFI research by using facilitation effects (greater accuracy on cross-modal congruent trials than unimodal trials) as a measure of multisensory integration. Do congruent auditory or visual stimuli increase the accuracy of beep/flash perception? Finally, the current study will contribute to the modality dominance literature by using quantity judgements and accuracy (how many beeps or flashes) rather than speeded, modality-specific responses to auditory and visual information. Experiment 1 examined the effect of beeps on perception of flashes (replicating most SIFI studies), and Experiment 2 tested the effect of flashes on perception of beeps. Based on auditory dominance and modality appropriate hypothesis (Robinson & Sloutsky, 2010a; Welch & Warren, 1980), it is expected that congruent and incongruent auditory information will have a stronger effect on visual perception, whereas, it is expected that visual input will have a weak or no effect on auditory perception.

Experiment 1

Method

Participants Participants for Experiment 1 included 24 young adults (18 to 35 years). Young adults were recruited from the Ohio State University, and received class credit for the Introductory Psychology course in return for their participation. Three participants with uncorrected hearing or vision, as self-reported, were excluded from the data analyses.

Apparatus The experiment was conducted on a 22" Dell PXL 2230 MW monitor with 1920 x 1080 resolution and Dell Optiplex 7040 systems with Intel Core i5 processors. Bose QuietComfort 25 Noise Cancelling headphones were used for auditory stimulus presentation. Stimulus timing and presentation and reaction time/accuracy data was collected using Direct RT software.

Materials The visual stimulus was a white circle 2º in visual angle in the center of the screen with a black background. Each flash has a 20 ms duration with a 50 ms Inter-Stimulus Interval (ISI) between consecutive flashes. The auditory stimulus was a sine wave presented at 3.5 kHz (no rise or decay ramps). Each beep lasted for 20 ms, and there was a 50 ms ISI in between consecutive beeps. Auditory stimuli were presented via headphones at approximately 50 dB. In the crossmodal condition, the first beep occurred 35 ms before the first flash, or vice versa. The beep first and visual first conditions were randomized among the participants. Figure 1 shows the timing of the stimuli. The stimuli and timing was modeled after the original SIFI study (Shams, Kamitani & Shimojo, 2000). Based on previous research and on preliminary analyses, the asynchronous timing had no significant effect on the SIFI.

Figure 1. Timeline of a 2 flash/2 beep stimuli. The dark blocks represent presentation of the stimuli and the grey represent ISI. The numbers above represent the time in ms from the beginning of the stimulus.

Design The experiment consisted of three blocks: visual unimodal, auditory unimodal, and crossmodal. There were five trials for each stimulus in the unimodal visual condition (2 flashes, 3 flashes, 4 flashes), and there were five trials for each possible stimulus in the unimodal auditory condition (2 beeps, 3 beeps, 4 beeps). There were also five trials of each possible stimulus in the crossmodal conditions (1 flash/1 beep, 1 flash/2 beeps, 1 flash/3 beeps, etc.). See Table 1 for all stimulus frequencies. Of the crossmodal trials, 15 were congruent and 30 were incongruent. Congruent trials had the same number of flashes and beeps, and incongruent had different numbers of flashes and beeps), which provided conflicting information.

Procedure In the unimodal auditory condition, participants heard 2, 3, or 4 beeps, and were asked to report how many beeps they heard. In the unimodal visual condition, they saw 2, 3, or 4 flashes and were asked to report how many they saw. In the crossmodal condition, participants were presented with 2, 3, or 4 beeps and/or 2, 3, or 4 flashes, and they were asked to report only how many flashes they saw. Each condition had a set of instructions before the trials and a conclusion to let the participant know when the condition was over. The order of condition was randomized among the participants, and each trial started within a condition started approximately 1000 ms after responding to the previous trial.

Unimodal	
Auditory	Visual
2 Beeps (5)	2 Flashes (5)
3 Beeps (5)	3 Flashes (5)
4 Beeps (5)	4 Flashes (5)
Crossmodal	
*2 Flashes/2 Beeps (5)	
*3 Flashes/3 Beeps (5)	
*4 Flashes/4 Beeps (5)	
2 Flashes/3 Beeps (5)	
2 Flashes/4 Beeps (5)	
3 Flashes/2 Beeps (5)	
3 Flashes/4 Beeps (5)	
4 Flashes/2 Beeps (5)	
4 Flashes/3 Beeps (5)	

Table 1: Experiment 1 trial types and frequencies. Note, "*" denotes congruent trials.

Results and Discussion

On each trial, participants reported how many flashes they perceived. Below we first report overall accuracies and then we report more traditional analyses focusing on actual responses (2, 3, or 4) and making a distinction between fission trials (more beeps than flashes) and fusion trials (fewer beeps than flashes).

Accuracy Each trial was classified as correct or incorrect. See left side of Figure 2 for means and standard errors of visual responses and the right side of Figure 2 for unimodal auditory accuracy. Analyses in Experiment 1 focus exclusively on visual responses. Using a 3 (number: 2, 3, 4) x 3 (trial type: congruent, incongruent, unimodal baseline) repeated measures ANOVA, a significant effect of condition was found, F (2,46) = 68.31, $p < .001$, $\eta_p{}^2 = .75$. Accuracies were lower on unimodal trials (M = .53, SE = .03) than congruent trials (M = .66, SE = .03), t (23) = -3.613, p = .001, which is consistent with facilitation effects. Interference effects were also found with higher accuracy on unimodal trials than incongruent trials (M = .21, SE = .09), t (23) = 8.42, $p < .001$. Also, accuracy was also higher on congruent trials than incongruent trials, t (23) = 9.91 $p < .001$.

The ANOVA also revealed a significant effect of number, F (2,46) = 17.05, $p < .001$, $\eta_p{}^2 = .43$, with accuracy decreasing as the number of flashes increased. There was significantly higher accuracy on the 2-flash trials (M = .56, SE = .03) than on the 4-flash trials (M = .31, SE = .04), t (23) = 4.47, $p < .001$. The 3-flash trials (M = .53, SE = .03) also had a higher accuracy than the 4-flash trials, t (23) = 5.02, $p < .001$. Finally, the analyses also revealed a trial type x number interaction, F (4,92) = 5.32, p = .001, $\eta_p{}^2 = .188$. As can be seen in Figure 2, cross-modal facilitation effects (congruent > unimodal) was most pronounced when presented with four flashes, and interference effects (unimodal > incongruent) decreased, with the strongest interference on 2-flash trials.

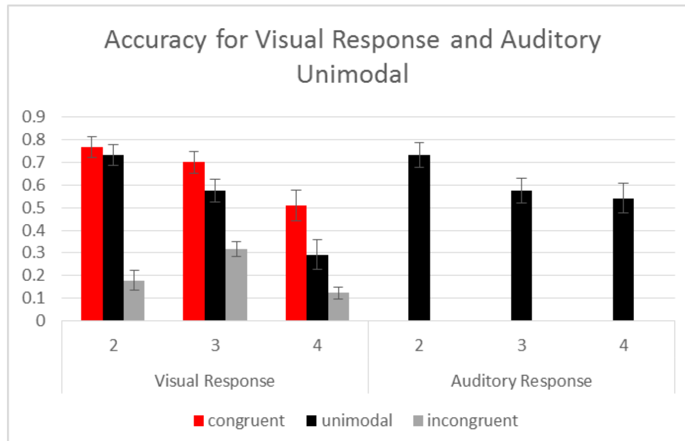

Figure 2. Accuracies across number and trial type. Error Bars denote Standard Errors.

The remaining analyses focus on actual responses (2, 3, or 4), not accuracies. On fission trials, there were more beeps than flashes, and on fusion trials, there were fewer beeps. Moreover, for fission and fusion cross-modal trials, we could only test two of the three numbers. For example, as can be seen in Table 1, there were no fission trials for 4 flashes and no fusion trials for 2 flashes because there were no trials where we presented five or one auditory stimulus, respectively. Thus, we used two 2 x 2 repeated measures ANOVA's to test for fission and fusion effects.

Fission Actual responses were collected on each trial and only fission trials and unimodal trials were submitted to a 2 (trial type: unimodal, fission) x 2 (number: 2, 3) repeated measures ANOVA. See left side of Figure 3 for means and standard errors. A significant effect of trial type was found, F (1,23) = 71.41, p <.001, $\eta_p{}^2 = .76$, suggesting that auditory input affected perception of flashes. In particular, participants reported more flashes on fission trials (M = 3.23, SE = .08) than unimodal trials (M = 2.54, SE = .06), which was expected since there were more beeps than flashes. There was also a significant effect of number, F (1,23) = 60.15, $p < .001$, $\eta_p{}^2 = .72$. Not surprisingly, participants on 2-flash trials (M =

2.68, *SE* = .05) reported fewer flashes than on 3-flash trials (*M* = 3.09, *SE* = .07).

Fusion Actual responses were collected on each trial and only fusion trials and unimodal trials were submitted to a 2 (trial type: unimodal, fusion) x 2 (number: 3, 4) repeated measures ANOVA. See right side of Figure 3 for means and standard errors. A significant effect of trial type was found, *F* (1,23) = 15.87, *p* = .001, η_p^2 = .41, which suggests that the number of beeps affected perception of flashes. Participants reported fewer flashes on fusion trials (*M* = 2.56, *SE* = .03) than on unimodal trials (*M* = 2.95, *SE* = .09), which was expected since there were fewer beeps than flashes. There was also a significant effect of number of stimuli on response, *F* (1,23) = 62.16, *p* < .001, η_p^2 = .73, with participants reporting fewer flashes on 3-flash trials (*M* = 2.57, *SE* = .045) than 4-flash trials (*M* = 2.94, *SE* = .06).

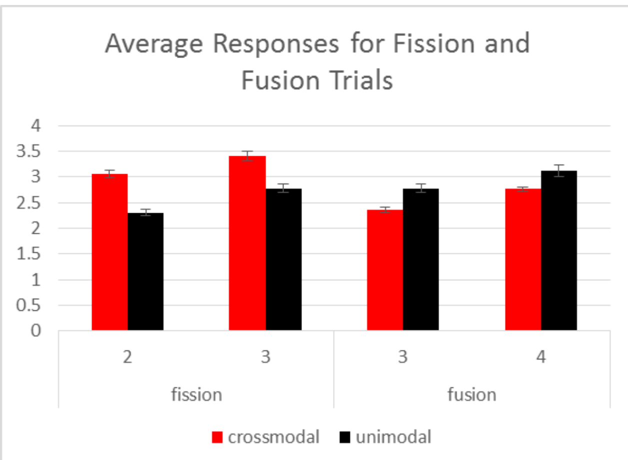

Figure 3. Actual responses across number and trial type. The left side of the figure denotes fission trials and the right side denotes fusion trials. Error Bars denote Standard Errors.

Experiment 2

The purpose of Experiment 2 was to test if the relative contribution of auditory and visual information on multisensory integration was symmetrical or asymmetrical. In cross-modal trials of Experiment 2, participants were asked to report how many beeps they heard. It was hypothesized that the effects would be asymmetrical, with visual input in Experiment 2 having little to no effect on auditory processing.

Method

Participants, Materials, and Procedure Experiment 2 was identical to Experiment 1, with the exception that we tested effects of flashes on beep perception; thus, the same participants from Experiment 1 were told to respond to report out how many beeps they heard, regardless of how many flashes they saw. To ensure that they were paying attention to the visual stimuli and did not shut their eyes in the cross-modal condition, a green visual stimulus (small green square) was presented for each possible trial type, and participants were asked to hit the space bar instead of 2, 3, or 4 when they saw the green stimulus. Five participants were removed because they did not detect the green catcher stimulus on at least 75% of the trials.

Results and Discussion

Accuracy See left side of Figure 4 for mean accuracies and standard errors on auditory response trials and right side of Figure 4 for unimodal visual responses. Experiment 2 focused exclusively on auditory responses. Using a 3 (number: 2, 3, 4) x 3 (trial type: baseline, congruent, incongruent) repeated measures ANOVA, a significant effect of number of beeps presented was found, *F* (2,46) = 19.15, *p* < .001, η_p^2 = .45. Based on the data, there was significantly higher accuracy on the 2-flash trials (*M* = .80, *SE* = .03) than on 3-flash trials (*M* = .67, *SE* = .04), *t* (23) = 2.75, *p* = .011, and 4-flash trials (*M* = .48, *SE* = .05), *t* (23) = 5.22, *p* < .001. The 3-flash trials also had a higher accuracy than 4-flash trials, *t* (23) = 4.09, *p* < .001. In addition, a condition x number interaction was observed, *F* (4,92) = 3.36, *p* = .013, η_p^2 = .13. No differences were found across trial types for 2- and 4-flash trials; however, congruent and incongruent trials both exceeded the baseline on 3-flash trials, *ts* (23) > -1.94, *ps* < .033 (one-tailed).

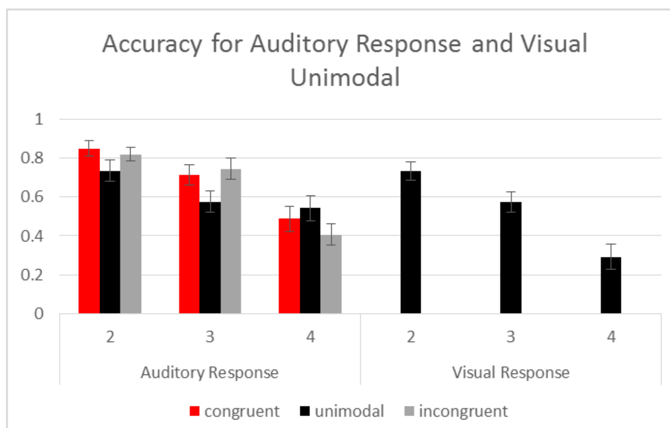

Figure 4. Accuracies across number and trial type. Error Bars denote Standard Errors.

Fission Actual responses were collected on each trial and only fission trials and unimodal trials were submitted to a 2 (trial type: unimodal, fission) x 2 (number: 2, 3) repeated measures ANOVA. See left side of Figure 5 for means and standard errors. Using a 2 (condition: unimodal, fission) x 2 (number: 2, 3) repeated measures ANOVA, a significant effect of number of stimuli presented on response was found, *F* (1,23) = 160.07, *p* <. 001, η_p^2 = .87. The 2-flash trials (*M* = 2.24, *SE* = .04) had a significantly lower response than the 3-flash trials (*M* = 2.96, *SE* = .06), *t* (23) = -12.65, *p* < .001.

There was no effect of trial type, suggesting that flashes did not affect beep perception.

Fusion Actual responses were collected on each trial and only fusion trials and unimodal trials were submitted to a 2 (trial type: unimodal, fusion) x 2 (number: 3, 4) repeated measures ANOVA. See right side of Figure 5 for means and standard errors. A significant effect of number of stimuli presented on response was found, F (1,23) = 60.30, $p < .001$, $\eta^2 = .72$. The 3-flash trials (M = 2.95, SE = .06) had a significantly lower response than the 4-flash trials (M = 3.40, SE = .08), t (23) = -7.77, p < .001. Again, there was no effect of trial type, suggesting that the flashes did not affect beep perception.

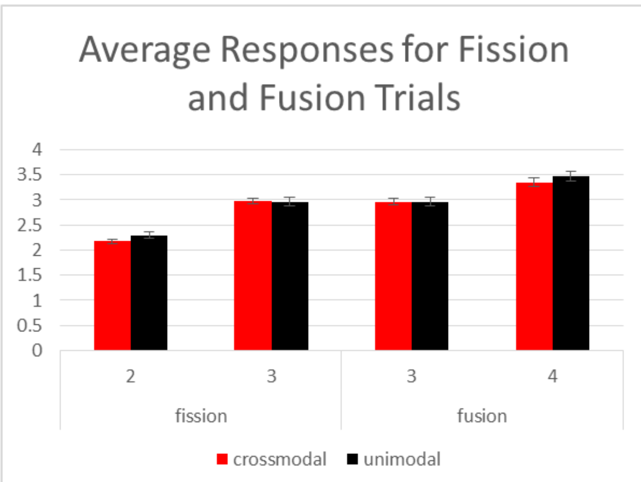

Figure 5. Actual responses across number and trial type. The left side of the figure denotes fission trials and the right side denotes fusion trials. Error Bars denote Standard Errors.

General Discussion

Many tasks require processing and integration of multisensory information. The primary goal of the current study was to examine relative contributions of auditory and visual information on multisensory integration. In Experiment 1, we hypothesized that auditory information would have a strong effect on visual processing, as seen in Shams et al. (2000). The results of Experiment 1 supported this hypothesis. In particular, when auditory and visual information provided the same information (congruent trials in Figure 2), adults were more accurate at reporting the number of flashes. Moreover, incongruent trials also affected visual perception. Participants overestimated the number of flashes when the flashes were paired with more beeps (fission trials in Figure 3) and underestimated the flashes when paired with fewer beeps (fusion trials in Figure 3). In Experiment 2, it was hypothesized that the visual information would not have as strong of an effect on the auditory processing, based on auditory dominance (Robinson & Sloutsky, 2010a) and the modality appropriateness hypothesis (Welch & Warren, 1980). The results of Experiment 2 supported this hypothesis, as most of the analyses showed that the visual information had no effect on auditory processing.

This expands the SIFI research by observing the effects of both auditory and visual information on multisensory integration. According to our knowledge, previous research has only focused on effects of auditory input on visual processing or vice versa; thus, these studies cannot determine if effects are symmetrical. The current study also used facilitation effects as a measure of multisensory integration. Facilitation effects were observed, and performance on the congruent trials was better than performance on the unimodal trials (baseline). These effects are seen in the visual responding condition with auditory input facilitating visual processing, but were not seen in the auditory responding condition. This asymmetry is consistent with both auditory dominance (Robinson & Sloutsky, 2010a) and the modality appropriateness hypothesis (Welch & Warren, 1980).

This study also expands modality dominance literature by measuring quantitative responses, rather than just response times, as seen in visual dominance research. Visual dominance has been observed for the past forty years in adults, showing that visual input often dominates auditory processing when making speeded, modality specific responses (e.g., Colavita, 1974). The findings of the current study were not tied to speeded modality specific responses, but were associated with accuracy of quantitative judgments. The findings support auditory dominance and modality appropriateness hypothesis and show that auditory input has a larger effect on visual processing than vice versa. Future research could take further measures to separate these two findings, as it cannot be distinctly determined whether the results are an effect of auditory dominance or modality appropriateness. Finally, it will be important to examine the role of stimulus intensity on multisensory integration, as changes in unimodal sensitivity may underlie developmental changes in multisensory integration. In particular, increased multisensory integration with age might stem from older adults compensating for weakened unimodal processing (DeLoss, Pierce, & Anderson, 2013). While Anderson (2004) found that weakening the auditory stimulus to near threshold increased visual effects on multisensory integration, weakening both modalities tends to decrease the SIFI (Parker & Robinson, in prep), and it is unclear how weakened auditory stimuli affect multisensory integration.

In summary, most of our experiences are multisensory in nature and it is important to understand how auditory and visual information contributes to multisensory integration. Future research needs to examine how this ability changes across the lifespan.

References

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, *14*(3), 257-262.

Bahrick, L. E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental psychology*, *36*(2), 190.

Burr, D., Banks, M.S., & Morrone, M.C. (2009). Auditory dominance over vision in the perception of interval duration. *Experimental Brain Research*, *198*(1), 49-57.

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.

Colonius, H., & Diederich, A. (2006). The race model inequality: Interpreting a geometric measure of the amount of violation. *Psychological Review*, *113*, 148–154.

Conway, C.M., & Christiansen, M.H., (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(1), 24-39.

DeLoss, D., Pierce, R., Andersen, G. (2013). Multisensory Integration, Aging, and the Sound-Induced Flash Illusion. *Psychology and Aging, 28*, 802-812.

Dennis, N.A., Howard, J.H., & Howard, D.V. (2006). Implicit sequence learning without motor sequencing in young and old adults. *Experimental Brain Research*, *175*, 153–164.

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

Giard, M.H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience*, *11*(5), 473-490.

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

Nissen, M.J., Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19,* 1–32.

Parker, J.L., & Robinson, C.W. (in prep). *Aging in Multisensory Integration*. Manuscript to be submitted for publication.

Robinson, C. W., & Sloutsky, V. M. (2013). When audition dominates vision: Evidence from cross-modal statistical learning. *Experimental Psychology*, *60*, 113-121.

Robinson, C. W., & Sloutsky, V. M. (2010). Development of cross-modal processing. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*, 135-141.

Robinson, C. W., & Sloutsky, V. M. (2004). Auditory dominance and its change in the course of development. *Child Development*, *75*, 1387-1401.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: *Listening (and learning) out of the corner of your ear. Psychological Science*, 8, 101–105.

Shams, L., Kamitani, S., Shimojo, S. (2000). Illusions. What you see is what you hear. *Nature*, *408*, 788.

Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research*, *14*(1), 147-152.

Sinnett, S., Spence, C., & Soto-Faraco, S. (2007). Visual Fdominance 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*(3), 822-833.

Song, S., Howard, J.H., & Howard, D.V. (2008). Perceptual sequence learning in a serial reaction time task. *Experimental Brain Research*, *189*, 145–158.

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.

Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition*, *97*, B25-B34.

Tsay, C. (2013). Sight over sound in the judgement of music performance. *Proceedings of the National Academy of Sciences, 110(36), 14580-14585. Doi;10.1073/pnas. 1221454110*

Vadillo, M. A., Konstantinidis, E., & Shanks, D. R. (2016). Underpowered samples, false negatives, and unconscious learning. *Psychological Bulletin and Review*, *23*, 87-102.

Van der Burg, E., Olivers, C.N., Bronkhorst, A. W., & Theeuwes, J. (2008). Pip and pop: nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 1053-1065.

Wada, Y., Kitagawa, N., Noguchi, K. (2003). Audio-visual Integration in temporal perception. *International Journal of Psychology*, *50*, 117-124.

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, *88*, 638-667.

Younger, B. A., & Cohen, L. B. (1983). Infant perception of correlations among attributes. *Infant Development*, *54*, 858-860.