Title

Auditory Stimuli Disrupt Visual Detection in a Visuospatial Task

Permalink

https://escholarship.org/uc/item/4078t9fs

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 41(0)

Authors

Robinson, Christopher W.
Laughery, Dylan

Publication Date

2019

Auditory Stimuli Disrupt Visual Detection in a Visuospatial Task

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

Dylan Laughery (laughery.12@osu.edu)
Department of Psychology, The Ohio State University at Newark
1179 University Dr., Newark, OH 43056, USA

Abstract

The current study used an eye tracker to examine how auditory input affects the latency of visual fixations and speeded responses on a Serial Response Time Task (SRTT). In Experiment 1, participants viewed a sequence of visual stimuli that appeared in different locations on a computer monitor and the same sequence repeated throughout the experiment. The visual sequence was either presented in silence or paired with uncorrelated sounds (i.e., sounds did not predict visual target location). Participants made more fixations and were more likely to fixate on the visual stimuli when visual sequences were presented in silence than when paired with sounds. Participants in Experiment 2 were presented with the same sequences, but they also had to determine if each visual stimulus was red or blue. The presence of auditory stimuli had no effect on accuracy (red vs. blue), however, there was some evidence that auditory stimuli delayed the latency of first fixations to the visual stimuli and discriminating the images as red or blue was also slower relative to the unimodal visual baseline. While visual stimuli often dominate auditory processing on spatial tasks, the current findings show that auditory stimuli can also slow down visual detection on a task that is better suited for the visual modality. These findings are consistent with a potential mechanism underlying auditory dominance effects, which posits that auditory stimuli may attenuate and/or delay the encoding of visual information.

Keywords: Attention, Multisensory Processing, Auditory Dominance

Introduction

Over the last 40 years, there has been a considerable amount of research examining how individuals process and integrate multisensory information (see Bahrick, Lickliter, & Flom, 2004; Calvert, Spence & Stein, 2004; Robinson & Sloutsky, 2010; Spence, Parise, & Chen, 2012; Stein & Meredith, 1993, for reviews). Much of this research focuses on multisensory integration where information from different sensory modalities is quickly, if not automatically, bound into a multisensory percept in which processing and responding to these multisensory percepts is often faster and more efficient than responding to the unisensory information (Bahrick, Flom, & Lickliter, 2002; Fort, Delpuech, Pernier, & Giard, 2002; Giard & Peronnet, 1999; Miller, 1982). For example, localizing a visual stimulus paired with a sound is often faster than localizing a visual stimulus presented in silence.

However, there are also many situations where multisensory information is arbitrary in nature and information presented to one sensory modality is unrelated to the information presented to the other sensory modality (e.g., listening to music while visually navigating traffic). Under these situations, multisensory presentation can sometimes disrupt encoding, learning, and/or responding, with one sensory modality dominating processing of the other sensory modality. For example, modality dominance research in adults often shows that when auditory and visual stimuli are presented simultaneously, visual input often dominates processing of auditory information (Colavita, 1974; Sinnett, Spence, & Soto-Faraco, 2007; see also Spence et al., 2012, for a review).

There is recent evidence of auditory dominance in adults (Barnhart, Rivera, & Robinson, 2018; Dunifon, Rivera, & Robinson, 2016; Robinson, Moore, & Crook, 2018), however, research pointing to auditory dominance in adult populations typically relies on temporal tasks (e.g., Parker & Robinson, 2018; Robinson & Sloutsky, 2013; Shams et al., 2000; 2002). More specifically, while the auditory modality can sometimes dominate visual processing on temporal tasks, the visual modality typically dominates auditory processing on spatial tasks (Welch & Warren, 1980). These findings suggest that modality dominance effects are flexible in nature and vary as a function of response demands (Robinson, Chandra, & Sinnett, 2016), nature of the task (Welch & Warren, 1980), and signal strength (Alias & Burr, 2004).

Given that auditory dominance effects are less prevalent in the adult literature, the primary goal of the current paper was to focus on these effects. One potential mechanism underlying auditory dominance is that sensory modalities might be competing for attention (Robinson & Sloutsky, 2010; see also Duncan, Martens, & Ward, 1997; Eimer & Driver, 2000; Sinnett et al., 2007; Wickens, 1984, for related discussions). Moreover, because auditory stimuli are often dynamic and transient in nature, it would be adaptive to first allocate attention to this information before it disappears. Attentional resources automatically deployed to the auditory modality might come with a cost - disrupted or delayed visual processing. There is some support for this claim from studies using temporal and recognition tasks (Barnhart et al., 2018; Dunifon et al., 2016; Parker & Robinson, 2018; Robinson et al., 2018, Robinson & Sloutsky, 2013; Shams et al., 2000; 2002), however, a stronger test of this proposed mechanism would be to examine if auditory stimuli also delay visual processing on a visuospatial task, a task better suited for the visual modality.

A recent study presented adults with a SRTT, which was administered on a touch screen computer (Robinson & Parker, 2016). As in previous research using variations of this paradigm (e.g., Dennis, Howard, & Howard, 2006; Nissen & Bullemer, 1987; Song, Howard, & Howard, 2008), Robinson and Parker (2016) presented visual information to spatially distinct locations, and participants had to quickly respond to this information (i.e., they had to touch each stimulus when it appeared on the touch screen monitor). Unbeknownst to participants, the visual sequences were structured and followed the same sequence throughout the experiment. Motor responses sped up over the time suggesting that, at some level, participants were learning the sequences. More relevant to the current study, motor responses to the visual stimuli were slower when the visual stimuli were paired with uncorrelated sounds (i.e., sounds that did not predict/correspond with location of the visual stimulus).

The current study expands on this research by using variations of a SRTT administered on an eye tracker to examine patterns of visual fixations over time. In both reported experiments, participants were shown two visual sequences of 12 stimuli, and the same sequences repeated throughout the experiment. In one condition, the sequence was presented in silence (unimodal condition) and in the other condition, the visual sequence was paired with sounds that were not correlated with the spatial location of the visual stimuli (cross-modal condition). Participants either counted the number of visual stimuli (Experiment 1) or they responded to each stimulus by quickly making a distinction on whether visual stimulus was red or blue (Experiment 2). If auditory stimuli are disrupting visual detection/encoding, then latency of first fixations to the visual stimuli should also be delayed. However, if auditory stimuli are disrupting later stages of visual processing (e.g., response/decision phase), then auditory interference should only be found in Experiment 2 when participants are making explicit responses to each stimulus.

Experiment 1

Method

Participants Forty undergraduate students (*M* = 19.41 years, *SD* = 1.61 years, 22 Females, one person did not disclose gender or age information) from The Ohio State University at Newark participated in the experiment for course credit. Data from 11 other participants were excluded from the study due to technical difficulties such as poor calibration, software crashes, etc.

Apparatus Participants were centrally positioned and seated approximately 65 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” 1920 x 1080 monitor and auditory stimuli were presented via Kensington 33137 headphones. Eye tracking data were collected and stored on a Dell Optiplex 7010 computer. Gaze fixation positions, Areas Of Interest (AOIs), and fixations were identified by the EyeLink system and data were exported using Data Viewer. The eye tracker, stimulus presentation computer, and eye tracking computer were stationed in a quiet testing room and a trained experimenter oversaw the entire duration of each participant’s study.

Materials and Design Visual stimuli were solid red and blue circles (100 pixels in diameter) and were presented on a white background. Visual stimuli were presented for 700 ms and were presented one right after another with no interstimulus interval. Auditory stimuli were 6 sine waves (500 - 3000 Hz, each stimulus increasing by 500 Hz) and 6 sawtooth waves (250 - 2750 Hz, each increasing by 500 Hz) Auditory stimuli were created in Audacity and were presented via headphones at a comfortable level - approximately 65 dB. Auditory stimuli were presented for 500 ms, and the auditory and visual stimuli shared the same onset.

The experiment consisted of two within-subjects conditions: a unimodal condition and a cross-modal condition. We presented two visual sequences of 12 distinct circle locations that repeated 20 times (see Figure 1 for sequences). In the unimodal condition, the sequence was presented in silence, and in the cross-modal condition, the visual sequence was paired with sounds. The color of the circles in both conditions was random (not correlated with the location of the circle), as were the sounds in the cross-modal condition. The order of the two sequences and the sequence-condition pairings were counterbalanced across participants.

Procedure Participants were told that they would see red and blue circles appear one at a time in different locations across the screen. They were instructed to look at the circles as they appear and respond by pressing a USB button placed in front of them after every 10 circles that they saw. Participants were not told that the circles would appear in the same sequence of 12 locations, however, they were informed that the study was split into two parts, a silent condition and a sound condition. Participants were given a consent form and demographics form to fill out before the study began.

After completing the consent and demographic forms, participants were calibrated on the eye tracker. Drift correction occurred every 50 stimuli (approximately 40 s), and we recalibrated the eye tracker every 100 stimuli (approximately 80 s). When the experiment concluded, the participants were given a three-question survey. On each item, they had to determine if they thought the order of the visual sequence, the order of the visual sequence paired with the sounds, and the order of the auditory sequence, was random or followed a pattern which repeated throughout the experiment. Question order was counterbalanced across participants (e.g., participants who received the unimodal condition first were first asked about the unimodal sequence and vice versa).

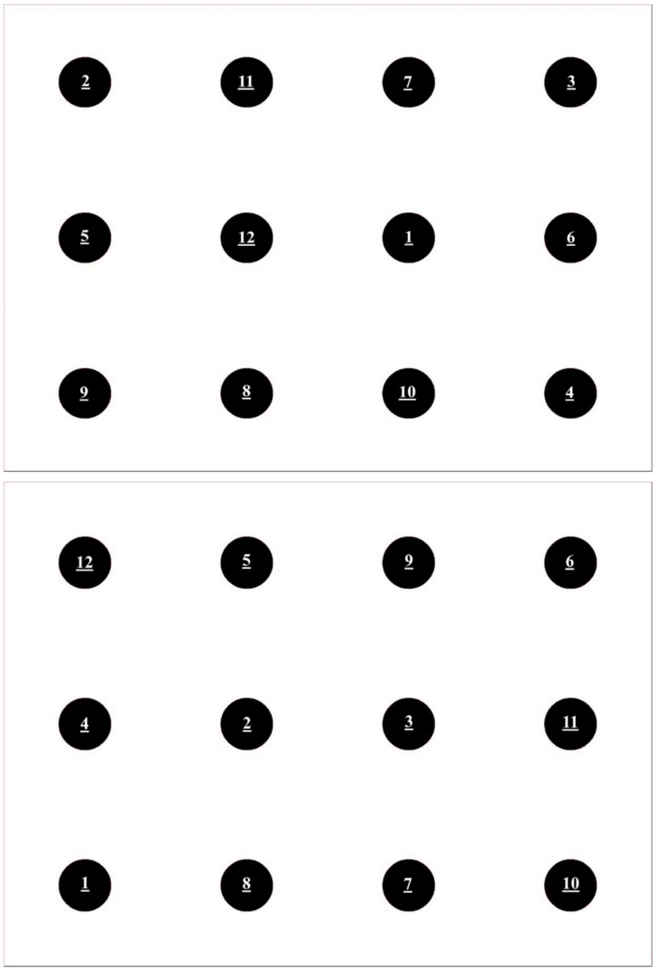

Figure 1: Order of the two visual sequences.

Results

We examined three different eye tracking variables which might be affected by the presence of auditory stimuli. First, we examined the latency of first fixations, which are sometimes slowed down on recognition tasks when visual stimuli are paired with sounds or words (Barnhart et al., 2018; Dunifon et al., 2016). Second, we examined the proportion of stimuli where participants made a fixation to the target location. If auditory stimuli disrupt visual encoding, there should be fewer fixations to the visual targets in the sound condition. Finally, we examined the number of fixations on each trial, however, these predictions are less clear. For example, attention automatically deployed to the auditory modality (or away from the visual modality) could reduce the overall number of fixations or it could make the visual task more challenging and require more fixations before detecting the visual target.

Each participant reported in the final sample completed the unimodal condition and the cross-modal condition, and in each condition, participants were presented with an ordered sequence of 12 visual stimuli, which repeated 20 times. Each sequence of 12 stimuli was considered as a trial, and to reduce noise, we created four blocks by averaging across five trials (60 stimuli). Thus, each condition consisted of four blocks of five trials, with 60 stimuli per block (e.g., Block 1 = first 60 stimuli, Block 2 = 61 - 120, etc.).

Every 700 ms a visual stimulus appeared in one of 12 pre-specified locations on the monitor and we recorded the latency first fixation to the visual stimulus (timestamp of first fixation to AOI - timestamp of stimulus onset). AOIs were created in Data Viewer and were 300 x 300 pixel squares centered around each visual stimulus. We submitted the mean latency of first fixations to a 2 (condition: unimodal, cross-modal) x 4 (block: blocks 1-4) repeated measures ANOVA. Mean latency of first fixations across condition and time ranged from 256 - 261 ms. There were no significant effects and the interaction did not reach significance, *p*s > .31.

We also examined the proportion of stimuli where participants made a fixation to the AOIs. If a participant made a fixation to the location of the target from stimulus onset to stimulus offset, then we coded that stimulus as a 1. If a participant did not make a fixation to the AOI during this time window, then we coded that stimulus as a 0. Proportions of fixations to the AOIs were averaged within each block and we submitted these values to a 2 (condition: unimodal, cross-modal) x 4 (block: blocks 1-4) repeated measures ANOVA. The analyses only revealed a marginally significant effect of condition, F (1, 39) = 3.95, p = .054, η_p^2 = .092, with participants making a higher proportion of fixations to the AOIs in the unimodal condition (M = .94, SE = .01) than in the cross-modal condition (M = .92, SE = .02).

The number of fixations from stimulus onset to stimulus offset (to any location on the monitor) was collected and we submitted these values to a 2 (condition: unimodal, cross-modal) x 4 (block: blocks 1-4) repeated measures ANOVA. The analysis only revealed an effect of condition, F (1, 39) = 5.66, p = .022, η_p^2 = .127, with participants making more fixations in the unimodal condition (M = 2.42, SE = .03) than in the cross-modal condition (M = 2.39, SE = .03).

Finally, at the end of the experiment, participants completed a three-item questionnaire. Random and Patterned responses for the unimodal and cross-modal conditions were analyzed using a McNemar's Chi-square. The McNemar Chi-square was significant (N = 40), p = .049, and one sample binomial tests compared to chance revealed that a majority of the participants thought the unimodal visual sequences were random (M = 68% reported random, p = .04), whereas, only 45% of the participants indicated that the visual sequences paired with sounds were random, which did not differ from chance, p = .64. Forty-five percent of participants also reported that the order of the auditory sequence was random.

In summary, while previous research demonstrated that auditory stimuli can slow down first fixations on recognition tasks (Barnhart et al., 2018; Dunifon et al., 2016) and slow down motor responses on a touch screen SRTT (Robinson & Parker, 2016), the current study found only weak support for auditory interference. More specifically, participants in the current study were slightly less likely to make a fixation to the visual stimulus when it was paired with a sound and they also made fewer fixations (to any location on the monitor). However, unlike Robinson and Parker (2016), there was no evidence that participants learned the visual sequences. Recall that latency of first fixations to the target locations did not speed up across training, whereas, motor responses sped up in Parker and Robinson (2016). Finally, while a majority of participants thought the unimodal visual sequences were random, participant responses did not differ from chance when sequences were paired with sounds. It is unclear if the uncorrelated sounds increased the perceived structure of visual input or if the sounds simply increased chance responding. However, if the sounds did increase the perceived structure of visual sequences, it did not result in faster or more fixations to the visual targets.

Experiment 2

The primary aim of the Experiment 2 was to further examine possible effects of auditory stimuli on visual sequence learning. Are interference effects restricted to tasks that require an explicit response? To address this aim, we presented participants with structured visual sequences in silence or paired with sounds, however, in contrast to Experiment 1, participants were required to make a response to each visual stimulus (i.e., indicate if the visual target was red or blue). If auditory stimuli interfere with visual processing during the decision/response phase, as opposed to disrupting encoding, then response times should slow down in the cross-modal condition in Experiment 2 while having no negative effect on the latency of first fixations. However, slowed response times and delayed first fixations would be consistent with the claim that auditory stimuli are disrupting visual encoding (Robinson & Sloutsky, 2010a).

Experiment 2 was not originally designed to examine the effects of engagement on sequence learning, however, requiring participants to make an explicit response to each stimulus should make the task more engaging. Thus, it is also possible to examine if poor engagement could account for the lack of learning in Experiment 1. While visual sequence learning on SRTT and statistical learning tasks are often thought to be implicit in nature and not dependent on attention (e.g., Nissen & Bullemer, 1987; Saffran, Newport, Aslin, & Tunick, 1997), it is possible that learning would be more robust if participants were more engaged throughout testing. Requiring participants to indicate if each visual stimulus is red or blue should make the task more engaging, which could result in better learning (i.e., faster response times and/or fixations across time).

Method

Participants, Materials, Design, and Procedure Thirty undergraduate students (M = 20.19 years, SD = 2.51 years, 20 Females) from The Ohio State University at Newark participated in the experiment for course credit. Data from eight other participants were excluded from the study due to technical difficulties, such as software/system crashes, computer lagging, or poor calibrations.

The procedure and design of Experiment 2 were identical to Experiment 1, except that in Experiment 2, a choice response task paradigm was used. Participants were required to make a color distinction with each stimulus by responding with one of two external USB buttons, labeled "RED" and "BLUE" respectively. Participants were instructed to respond as fast and as accurate as possible. The left-right locations of the buttons were counterbalanced across participants.

Results

As in Experiment 1, we examined the latency of first fixations, the proportion of stimuli where participants made a fixation to the visual target, and the number of fixations between stimulus onset and stimulus offset, however, we also examined response times and accuracies on the primary task.

First, as in Experiment 1, we submitted the mean latency of first fixations to a 2 (condition: unimodal, cross-modal) x 4 (block: blocks 1-4) repeated measures ANOVA. The analyses revealed a marginally significant effect of condition, $F(1, 29) = 3.62$, $p = .067$, $\eta_p^2 = .111$, and a significant time x condition interaction, $F(3, 87) = 3.28$, $p = .025$, $\eta_p^2 = .102$. While latency of first fixations were numerically faster across all blocks in the unimodal condition, simple effects with Bonferroni adjustments revealed that the difference between unimodal and cross-modal means only reached significance in block 3, $p = .012$ (see Figure 2).

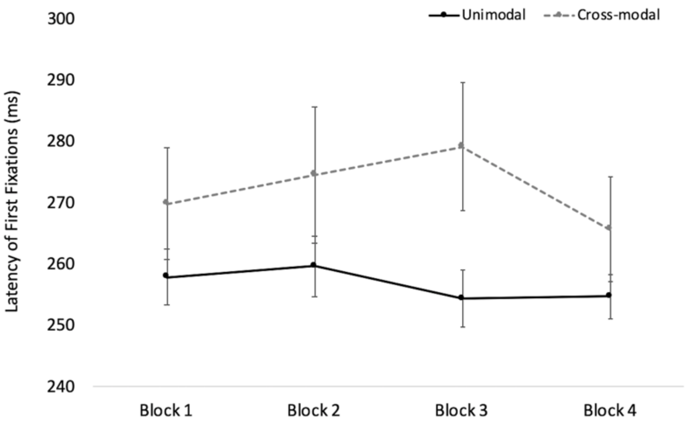

Figure 2. Latency of First Fixations (ms) across condition and time. Error bars denote Standard Errors.

Response times were also submitted to a 2 (condition: unimodal, cross-modal) x 4 (block: blocks 1-4) repeated measures ANOVA. The analysis revealed an effect of condition $F(1, 29) = 5.90$, $p = .022$, $\eta_p^2 = .169$, with response times being faster in the unimodal condition (M = 584 ms, SE = 19.85) than in the cross-modal condition (M = 624 ms, SE = 28.56). The analysis also revealed an effect of time, $F(3, 87) = 16.04$, $p < .001$, $\eta_p^2 = .356$. See Figure 2 for response times across condition and time. Pairwise comparisons with Bonferroni adjustments revealed that mean response times on Block 1 ($M = 637$ ms, $SE = 27.92$) were significantly slower than Block 2 ($M = 600$ ms, $SE = 24.18$), Block 3 ($M = 592$ ms, $SE = 22.72$), and Block 4 ($M = 587$ ms, $SE = 19.48$), ps $< .001$. Blocks 2-4 did not differ, ps $> .56$. Also note that accuracies (i.e., discriminating red vs. blue stimuli) exceeded .96 across all conditions with no significant effects or interactions, ps $> .22$.

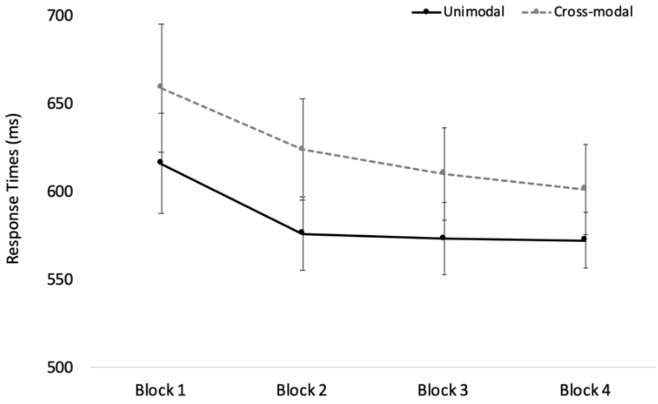

Figure 3. Response times (ms) across condition and time. Error bars denote Standard Errors.

The proportion of stimuli where participants made a fixation to the AOI were submitted to a 2 (condition: unimodal, cross-modal) x 4 (block: blocks 1-4) repeated measures ANOVA. Mean proportion of fixating to the visual stimuli ranged from .93 - .97, and the analysis revealed no significant effects or interactions, ps $> .32$.

The mean number of fixations between stimulus onset and offset were submitted to a 2 (condition: unimodal, cross-modal) x 4 (block: blocks 1-4) repeated measures ANOVA. The analysis only revealed a marginally significant effect of time, $F(3, 87) = 2.62$, $p = .056$, $\eta_p^2 = .083$. Pairwise comparisons with Bonferonni adjustments revealed that participants made more fixations on block 1 ($M = 2.24$, $SE = .05$) than on block 2 ($M = 2.18$, $SE = .05$), $p = .023$. Block 1 did not differ from block 3 ($M = 2.20$, $SE = .05$) or block 4 ($M = 2.19$, $SE = .05$), and blocks 2-4 did not differ, ps $> .323$.

Finally, we also examined responses on the three-item questionnaire, which was administered at the end of the study. Two participants did not complete the questionnaire. Patterned responses for the unimodal visual and visual sequence paired with sounds were analyzed using a McNemar's Chi-square. The McNemar Chi-square was not significant (N = 28), $p > .99$, Binomial tests compared to chance revealed that 78% of the participants thought the order of the unimodal visual sequence was random, different from chance, $p = .004$, and 82% of the participants reported that the order of the visual sequence paired with sounds was also random, different from chance, $p = .001$.

General Discussion

In both reported experiments, participants were shown two visual sequences, and each sequence repeated 20 times over the course of the experiment. One sequence was presented in silence, whereas, the other sequence was paired with sounds, which were not correlated with the location of the visual stimulus. In Experiment 1, participants simply counted the number of visual stimuli, pressed a button after every 10 stimuli, and we examined visual fixations throughout the procedure. Experiment 2 was more engaging, as participants were required to quickly determine if each visual stimulus was red or blue.

Auditory interference effects were found in both experiments. More specifically, in Experiment 1 when participants counted the number of visual stimuli, participants were more inclined to fixate on the visual stimuli in the unimodal condition and also made more overall fixations in the unimodal condition. When participants had to determine if each visual stimulus was red or blue, both latency measures showed some evidence of a slowdown/delay in the cross-modal condition. More specifically, latency of first fixations to the visual stimuli was slower in the cross-modal condition compared to the unimodal baseline, especially in block 3 (see Figure 2). In addition, overall response times were also slower in the cross-modal condition than in the unimodal condition.

The current study contributes to modality dominance research in the following ways. First, most research examining modality dominance in adults often points to visual dominance, with the visual modality dominating auditory processing (Spence, 2009; Spence et al., 2012, for reviews). While the current study did not examine the effects of visual input on auditory processing, the findings provide support for auditory dominance with auditory stimuli slowing down visual fixations and responding. These findings are remarkable given that spatial tasks are typically better suited for the visual modality (Welch & Warren, 1980). Moreover, the current study examined latency of first fixations as well as response times. If sounds were simply interfering during the response/decision phase, then only response times should have been slowed down. Finding evidence that first fixations to the stimuli were also delayed suggests that interference effects are happening early in the course of visual processing (i.e., during the detection phase).

While these findings shed light on the dynamics of multisensory processing, there are some limitations to the current study. First, while response times sped up in Experiment 2, there was no evidence in the eye tracking data that participants were learning the sequences. There are several reasons why learning may have not occurred. First, in both reported experiments, the color of the visual stimuli added noise to the sequences and the sounds in the cross-modal conditions also added additional noise (i.e., participants may have focused on these irrelevant variables and failed to learn the sequences). However, this additional information should not have affected sequence learning if the task is assessing implicit learning. It is also possible that participants were learning the sequences, but we failed to capture this learning because we primarily focused on participants' responses to visual stimuli and not on their anticipations (fixations before stimulus onset). These possibilities need to be addressed in future research.

In summary, the current study demonstrates that sounds can disrupt visual stimulus detection and response times. These effects have implications on tasks that require processing of multisensory information and shed light on possible mechanisms underlying auditory dominance effects.

References

Alias, D., & Burr., D. (2004). The Ventriloquist Effect Results from Near-Optimal Bimodal Integration. *Current Biology*, *14*, 257 - 262.

Bahrick, L. E., Flom, R., & Lickliter, R. (2002). Intersensory redundancy facilitates discrimination of tempo in 3-month-old infants. *Developmental Psychology*, *41*(4), 352-363.

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.

Barnhart, W. R., Rivera, S., & Robinson, C. W. (2018). Different patterns of modality dominance across development. *Acta Psychologica*, *182*, 154-165.

Calvert, G., Spence, C., & Stein, B. (2004). *The handbook of multisensory processe*s. Cambridge, MA: MIT Press.

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

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.

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

Dunifon, C. M., 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*(12), 1947-1958.

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.

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*(5), 473-490.

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, *14*, 247-279

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. (2018). Changes in multisensory integration across the lifespan. *Psychology and Aging*, *33*(3), 545-558.

Robinson, C.W., Chandra, M., & Sinnett, S. (2016). Existence of competing modality dominances. *Attention, Perception, & Psychophysics*, *78*, 1104-1114.

Robinson, C. W., Moore, R. L., & Crook, T. A. (2018). Bimodal presentation speeds up auditory processing and slows down visual processing. *Frontiers in Psychology*, *9*, 1-10.

Robinson, C.W., & Parker., J.L. (2016). Effects of auditory input on a spatial serial response time task. In Papafragou, A., Grodner, D., Mirman, D., & Trueswell, J.C. (Eds.) *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 2237-2242). Austin, TX: Cognitive Science Society.

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.

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 dominance and attention: Revisiting the Colavita effect. *Perception & Psychophysics*, 69, 673–686.

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.

Stein, B. E. & Meredeth, M. A. (1993). *The merging of the senses*. Cambridge, MA. MIT Press.

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.