

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Effects of Auditory Input on a Spatial Serial Response Time Task

Permalink

<https://escholarship.org/uc/item/7jj15188>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 38(0)

Authors

Robinson, Christopher W.

Parker, Jessica L.

Publication Date

2016

Peer reviewed

Effects of Auditory Input on a Spatial Serial Response Time Task

Christopher W. Robinson (robinson.777@osu.edu)

Department of Psychology, The Ohio State University at Newark
1179 University Dr, Newark, OH 43056, USA

Jessica L. Parker (parker.1026@osu.edu)

Department of Psychology, The Ohio State University at Newark
1179 University Dr, Newark, OH 43056, USA

Abstract

The current study examined how relevant and irrelevant auditory stimuli affect the speed of responding to structured visual sequences. Participants were presented with a dot that appeared in different locations on a touch screen monitor and they were instructed to quickly touch the dot. Response times sped up over time, suggesting that participants learned the visual sequences. Response times in Experiment 1 were slower when the dot was paired with random sounds, suggesting that irrelevant sounds slowed down visual processing/responding. Dots in Experiment 2 were paired with correlated sounds (both auditory and visual information provided location information). While the redundant intersensory information did not speed up response times, it did partially attenuate auditory interference. These findings have implications on tasks that require processing of simultaneously presented auditory and visual information and provide evidence of auditory interference and possibly dominance on a task that typically favors the visual modality.

Keywords: Cross-modal processing; Sensory Dominance; Attention.

Introduction

Many important tasks rely on detecting statistical regularities in the environment. For example, speech segmentation, word learning, and category learning require a person to abstract transitional probabilities of speech sounds, detect that some words co-occur with specific objects, and learn that some features are necessary or probabilistically relevant for a given category, respectively. Moreover, much of this learning appears to happen automatically with young infants quickly learning artificial categories (Younger & Cohen, 1983) and learning the transitional probabilities of speech sounds after two minutes of exposure to an artificial language (Saffran, Aslin, & Newport, 1996). Much of this learning can also happen without attention, with children learning probabilities of auditory sequences even when the primary task was visual in nature and they were not instructed to pay attention to the auditory sequences (Saffran, Newport, Aslin, Tunick, & Barrueco, 1997, but see Toro, Sinnett, & Soto-Faraco, 2005).

In addition to perceiving and abstracting statistical regularities, there are also occasions when the structure involves motor memory. For example, playing musical instruments, typing, swinging a golf club, driving a manual transmission, etc., require coordinated movements, which eventually become automated and consume few attentional

resources. Many studies have studied perceptual-motor learning in adults. For example, research using a Serial Response Time Task (SRTT) often consists of presenting visual information to spatially distinct locations, and participants are instructed to quickly respond to this information (e.g., Dennis, Howard, & Howard, 2006; Nissen & Bullemer, 1987; Song, Howard, & Howard, 2008). Unbeknownst to participants, the visual sequences are often structured and follow a statistical pattern, but see Vadillo, Konstantinidis, and Shanks (2016) for a recent review of implicit vs. explicit learning. While many studies have used variants of a SRTT, very little is known regarding how information from other sensory modalities affects learning of these structured visual sequences. Therefore, the primary goal of the current study is to examine how relevant and irrelevant auditory information affect learning and responding to visually structured sequences.

Over the last 40 years, there is a considerable amount of research showing that when simultaneously presented with auditory and visual information, the visual modality dominates the auditory modality (Colavita, 1974; Colavita, Tomko, & Weisberg, 1976; Colavita & Weisberg, 1979; Egeth & Sager, 1977). For example, in a classical Colavita task, participants are instructed to quickly respond to auditory and visual information by quickly pressing one button when they hear a sound and by pressing a different button when they see an image/flash (Colavita, 1974). On some of the trials, auditory and visual information are presented at the same time. Participants often miss these cross-modal trials by only pressing the visual button; therefore, it was concluded that the visual modality dominated the auditory modality. The Colavita visual dominance effect and variations of this task consistently point to visual dominance, with stimulus and attentional manipulations often weakening but not reversing the effect (see Sinnett, Spence, & Soto-Faraco, 2007; Spence, Parise, & Chen, 2012 for reviews). While numerous sensory, attentional, and motor mechanisms have been put forward to account for visual dominance, underlying mechanisms are poorly understood.

Recent findings provide some support for auditory dominance; however, these studies often test infants and children rather than adults (Lewkowicz, 1988a; 1988b; Robinson & Sloutsky, 2004; Sloutsky & Napolitano, 2003; Sloutsky & Robinson, 2008). For example, infants and children are often better at discriminating pictures when presented in silence than when the same pictures are paired

with sounds or words (Lewkowicz, 1988a; 1988b; Robinson & Sloutsky, 2004; 2010a; Sloutsky & Robinson, 2008). At the same time, the pictures appear to have no negative effect on auditory processing; thus, multisensory presentation attenuated visual but not auditory processing. To account for this finding, Robinson and Sloutsky (2010) have posited that sensory modalities may share the same pool of attentional resources and compete for attention. Furthermore, due to the transient and dynamic nature of auditory input, it may be adaptive to first allocate attention to auditory input before it disappears. Increased attention automatically deployed to the auditory modality may come with a cost - attenuated or delayed visual processing.

While this competition for attention explanation (Robinson & Sloutsky, 2010) may account for some of the developmental findings, there are only a few studies pointing to auditory interference in adults, and these studies do not use a traditional Colavita paradigm that require participants to quickly respond to multisensory information. For example, in sound induced flash illusion, participants are presented with a series of beeps and flashes (Shams, Kamitani, & Shimojo, 2000; 2002). There are often no response time constraints and participants are not asked to attend to (or report on) the auditory information. In these tasks, the auditory information influences the number of flashes reported. For example, if participants see two flashes but hear three beeps, they might report seeing three flashes.

Cross-modal presentation can also affect visual statistical learning. In a cross-modal statistical learning task, participants were presented with streams of auditory, visual, or cross-modal sequences (auditory and visual sequences were presented at the same time), and participants were either tested on the auditory or visual sequences (Robinson & Sloutsky, 2013). Increasing the task demands by randomizing one of the streams attenuated visual but not auditory statistical learning. In other words, participants learned the visual sequences when presented in silence or when paired with correlated sounds, but randomizing the auditory stream attenuated visual statistical learning. Randomizing the visual stream had no negative effect on auditory statistical learning.

One possible explanation that may account for the auditory interference/dominance effects in the sound induced flash illusion and cross-modal statistical learning tasks (Robinson & Sloutsky, 2013; Shams, Kamitani, & Shimojo, 2000; 2002) is that both tasks rely almost exclusively on temporal processing. According to the Modality Appropriateness Hypothesis (Welch & Warren, 1980), the modality that is most suitable for a given task will dominate. Given that the visual system is better at processing location information (Alias & Burr, 2004) and the auditory modality is better at processing temporal information (Burr, Banks, & Morrone, 2009), it is not surprising to see the auditory modality dominate in temporal tasks such as statistical learning (Conway & Christiansen, 2005).

In the current study, we employed a spatial SRTT to determine if auditory stimuli also affect the speed of responding to visual input on a task better suited for the visual

modality. Participants were presented with a sequence of dots that appeared in different locations on a touch screen monitor. Participants either heard a random sequence of sounds (Experiment 1), a correlated sequence of sounds (Experiment 2), or the visual sequences were presented in silence (Experiments 1 and 2). According to the Modality Appropriateness Hypothesis (Welch & Warren, 1980), the task should be well suited for the visual modality; thus, the auditory input should have no negative effect on visual response times. However, if auditory stimuli automatically engage attention and pull attention away from the visual modality (c.f., Robinson & Sloutsky, 2010), then it is possible that auditory stimuli will slow down visual responses and/or slow down learning rate on a visual-spatial task.

Experiment 1

Method

Participants Twenty-nine undergraduate students from The Ohio State University-Newark (12 Females, $M = 18.26$ years) participated in Experiment 1, from which they gained research assignment credit for the Introduction to Psychology course. One participant was tested but not included in the analyses due to a reported hearing loss.

Apparatus The experiment was created using OpenSesame software and ran on a Dell Optiplex 9010 computer with an Intel Core i7 processor. The visual stimuli were shown on a 22" Planar PXL2230 1920 x 1080 touch screen monitor. Participants used the touch screen to respond to the visual stimuli. The auditory stimuli were presented via Kensington KMW33137 headphones at approximately 65-68 dB.

Materials and Design The visual stimulus was the default fixation stimulus generated in OpenSesame. The fixation stimulus was a filled white circle (dot) with an 8 pixel radius and a 2 pixel hole, and it was presented on a black background. The dot appeared at 12 different locations on the monitor, and each participant saw two sequences, as represented by the xy coordinates in Table 1. Sequence order (sequence 1 vs. sequence 2) was randomized for each participant with approximately half of the participants seeing sequence one first, and the other half seeing sequence two first.

Condition (silent vs. sound) was also manipulated within subjects. In the sound condition, each visual stimulus was paired with a tone and tones were presented at the following 12 different frequencies: 200 Hz, 400 Hz, 600 Hz, 800 Hz, 1200 Hz, 1200 Hz, 1400 Hz, 1600 Hz, 1800 Hz, 2000 Hz, 2200 Hz, 2400 Hz, and 2600 Hz. For approximately half of the participants, the tones were paired with sequence one, and for the remaining participants, tones were paired with sequence two.

Procedure The experiment consisted of a silent condition and a sound condition. In the silent condition, the visual sequences were presented in silence, and in the sound condition, visual sequences were paired with the tones. The

fixation dot appeared in a repeating pattern of 12 locations (see sequences in Table 1), and the same sequence repeated 20 times in each condition, giving a total of 240 trials per condition, and 480 trials total in the experiment. Participants were instructed to touch the dot on the touch screen monitor as quickly as possible. The dot stayed on the screen until the participant touched that location. In the sound condition, participants were told that they would hear a sound, but that they were to respond only to the visual stimuli as in the silent condition. The tones were randomly paired with visual sequences and the pairings switched on every trial. For example, on trial 1, a 200 Hz tone may have been presented when the dot appeared in location 1. On the next trial, location 1 may have been associated with a 1600 Hz tone, etc. The order was counterbalanced among the participants so that approximately half of the participants received the sound condition first, and the other half received the silent condition first. In addition, the visual stimulus pattern was also counterbalanced among participants so that half of the participants experienced the pattern in reverse order.

XY Coordinates of the Visual Stimulus		
Stimulus	Sequence 1	Sequence 2
1	(283, 116)	(1350, 524)
2	(456, 564)	(936, 703)
3	(708, 826)	(1436, 764)
4	(1436, 116)	(484, 118)
5	(1720, 516)	(34, 328)
6	(1233, 732)	(284, 764)
7	(284, 764)	(1233, 732)
8	(34, 328)	(1720, 516)
9	(484, 118)	(1436, 116)
10	(1436, 764)	(708, 826)
11	(936, 703)	(456, 564)
12	(1350, 524)	(283, 116)

Table 1. Above is the 12 location pattern of the visual stimuli in Experiment 1.

Results and Discussion

Reaction times were calculated for each stimulus and mean response times were averaged for each trial (12 stimuli per trial). See Figure 1 for mean response times and standard errors across the 20 trials. As can be seen in the figure, responses sped up over time, suggesting that some learning occurred, and response times were generally faster in the silent condition.

Log transformed response times were submitted to a 2 (silent vs. sound) x 4 (block 1, block 2, block 3, block 4) repeated measures ANOVA. A block was defined as five trials (e.g., block 1 = trials 1-5, block 2 = trials 6-10, etc.). The results showed a significant effect of condition, $F(1, 28) = 4.65, p = 0.04$, which shows that response times in the sound condition ($M = 812.40$ ms, $SE = 20.26$) were slower than the silent condition ($M = 765.00$ ms, $SE = 16.93$). There was also a significant effect of time, $F(3, 84) = 32.71, p < .001$, showing that reaction time sped up across the blocks.

The means and (*SEs*) for blocks 1 – 4 were 832.50 ms (15.15), 802.60 ms (17.69), 764.50 ms (16.84), and 754.70 ms (15.87), respectively. Paired samples *t*-tests, using log transformed response times, were conducted between the blocks, showing block 1 was significantly slower than block 2, $t(28) = 4.26, p < 0.001$, and block 2 was significantly slower than block 3, $t(28) = 3.53, p = 0.001$. The difference between block 3 and block 4 did not reach significance. There was also a significant Condition x Block interaction, $F(3, 84) = 5.65, p = 0.001$. As can be seen in Figure 1, response times in the sound condition were significantly slower than the silent condition in block 1 (trials 1-5), $t(28) = -2.91, p = .007$, and marginally slower than silent in block 2 (trials 6-10), $t(28) = -1.98, p = .057$. The sound and silent conditions did not differ in blocks 3 or 4.

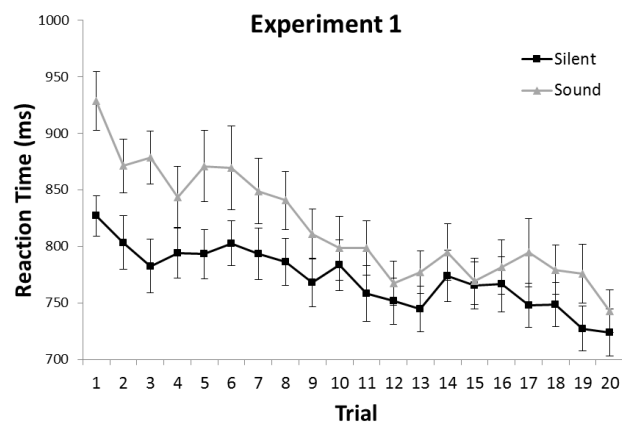


Figure 1. Mean response times across trials and condition. Error Bars denote Standard Errors.

In summary, the visual modality is typically well suited for processing of spatial information (Welch & Warren, 1980) and participants were clearly learning the visual sequences, as indicated by a speed up in response times. That said, the current experiment provides support for auditory dominance, with irrelevant tones slowing down responses to structured visual sequences. Even though participants were told to ignore the sounds, they couldn't, at least in the early stages of learning. By the end of the experiment, there was no significant difference in response times between the sound and silent condition. One possibility is methodological in nature and due to ceiling effects. However, it is also possible that this weakened interference stemmed from the visual task becoming more automated and less prone to cross-modal interference. It will be important to address this issue in future research.

Experiment 2

The goal of Experiment 2 was to determine if any auditory stimulus would slow down visual responses or if this effect was restricted to irrelevant tones. Participants in Experiment 2 were presented with the same two visual sequences used in Experiment 1 and one of the sequences was presented in

silence and one was paired with tones. In contrast to Experiment 1, the tones in this experiment were correlated with the location of the visual dots (e.g., every time a dot appeared in location 1, participants heard tone 1, etc.). If the presence of any auditory stimulus grabs attention and slows down visual processing, then the correlated/redundant auditory information may also slow down response times to the structured sequences. However, it is also possible that the interference is restricted to irrelevant or conflicting tones. If this is the case, then correlated sounds may not interfere with visual responses, with comparable response times in the silent and sound conditions. It is also possible that intersensory redundancy may also speed up processing (Colavita & Diederich, 2006; Giard & Peronnet, 1999), with response times in the correlated sound condition being faster than the unimodal visual baseline.

Method

Participants, Materials, and Procedure Thirty-one undergraduate student at The Ohio State University-Newark (13 Females, $M = 19.17$ years) participated in the study. In return, these students got credit for their research assignment in the Introduction to Psychology course. Data from one participant was not included due to reported hearing loss. Experiment 2 used the same visual and auditory stimuli used in Experiment 1. The procedure for Experiment 2 was exactly the same as the procedure for Experiment 1, but each visual stimulus location was paired with a specific auditory stimulus. Just as in Experiment 1, participants were instructed to touch the dot on the screen as quickly as possible when they see it appear.

Results and Discussion

Reaction times across the 20 trials are reported in Figure 2. As can be seen in the figure, response times sped up across training and auditory interference effects decreased compared to Experiment 1.

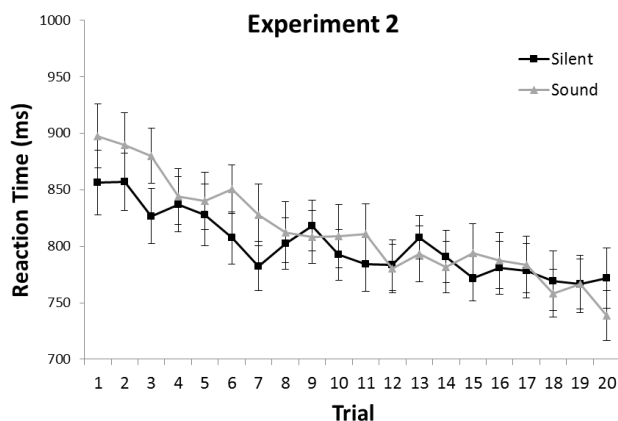


Figure 2. Mean response times across trials and condition. Error Bars denote Standard Errors.

Log transformed response times were averaged across five trials, and means were submitted to a 2 (silent vs. sound) x 4 (block 1, block 2, block 3, block 4) repeated measures ANOVA. The analysis only revealed an effect of time, $F(3, 90) = 55.74$, $p < .001$, showing that reaction time decreased across the blocks. The means and (*SEs*) for blocks 1–4 were 855.00 ms (19.98), 810.90 ms (19.67), 789.90 ms (18.69), and 770.60 ms (20.50), respectively. Paired samples *t*-tests, using log transformed data, were conducted between the blocks, showing block 1 to be significantly slower than block 2, $t(30) = 7.50$, $p < 0.001$. Block 2 was significantly slower than block 3, $t(30) = 4.04$, $p < 0.001$, and block 3 was significantly slower than block 4, $t(30) = 3.02$, $p = 0.005$. This displays a reliable pattern of learning across the blocks of the experiment. Although participants were initially faster in the silent condition at responding to visual input, the nonsignificant effect of condition and condition x block interaction suggest that cross-modal interference attenuated when using correlated sounds.

General Discussion

Many tasks require a person to divide attention across sensory modalities. The research on adults' processing of simultaneously presented auditory and visual information consistently points to visual dominance, with the visual modality dominating processing or responding (Colavita, 1974; Colavita, Tomko, & Weisberg, 1976; Colavita & Weisberg, 1979; Egeth & Sager, 1977). However, there are several recent studies highlighting situations where the presence of an auditory stimulus interferes, delays, or alters visual processing (Robinson & Sloutsky, 2013; Shams, Kamitani, & Shimojo, 2000; 2002). However, these studies employ tasks that rely almost exclusively on the processing of temporal information, which appears to be better suited for the auditory modality (Conway & Christiansen, 2005). The current study used a spatial SRTT, which should be better suited for the visual modality (Welch & Warren, 1980). In both of the reported experiments responding to a dot appearing in a predictable sequence sped up over time, suggesting that participants were learning the visual sequences. At the same time, response times were slower in Experiment 1 when the dot was paired with irrelevant tones compared to a silent condition, especially early in the experiment, which suggests that the auditory information was interfering with visual responses. Experiment 2 expands on this finding by showing that auditory interference is attenuated when using correlated sounds – sounds that also provided information regarding the location of the visual stimulus.

The finding that random sounds slowed down visual processing and/or responding more than correlated sounds is important and provides important insights into the nature of the cross-modal interference. For example, one explanation that can account for the findings in Experiment 1 is that the auditory and visual modalities share the same pool of resources and sensory modalities are competing for these resources (Robinson & Sloutsky, 2010). Moreover, because

auditory stimuli are dynamic and transient in nature it may be adaptive to allocate attention to this class of stimuli before they disappear. Thus, under high cognitive load conditions where resources are depleted or when examining speeded responses, auditory stimuli might automatically grab attention and attenuate or delay visual processing (Robinson & Sloutsky, 2010). The finding the auditory interference attenuated in Experiment 2 when sounds were correlated with visual information suggests that cross-modal interference effects may be occurring later in the course of processing, with only irrelevant and/or conflicting auditory information slowing down responding to visual input.

Several issues need to be examined in future research. First, irrelevant slowed down visual processing, and there was some evidence that correlated sounds weakened the effect. One possibility is that the effect is specific to auditory interference. However, it is also possible that irrelevant auditory stimuli increase task complexity or simply add conflicting information, which in turn slows down responses. For example, imagine a similar SRTT where dots vary in color. In the irrelevant condition, dot color varies randomly and does not predict spatial location. In the correlated/relevant condition, dot color is correlated with location. Finding that random colors also slow down learning would suggest that interference stems from increased task demands or any conflicting information, and the slowdown is not directly tied to auditory dominance per se. It will also be important to examine if interference effects are asymmetrical in nature - a signature pattern of modality dominance effects. The current study only examined the effect of sounds on visual sequence learning. Thus, to claim auditory dominance, future research will need to show that the presence of the visual sequence has no negative effect on auditory sequence learning.

Another issue that will need be addressed involves resolving the discrepancy between the current findings and the findings from the pip and pop effect (Van der Burg, Olivers, Bronkhorst, & Theeuwes, 2008). In the pip and pop task, participants are presented with numerous lines on a monitor and they have to quickly find a line that is perfectly vertical or perfectly horizontal, and then report out the line's orientation. The task is challenging when only presented visually; however, changing the color of the lines and synchronizing line color with an audible click significantly speeds up target detection. Interestingly, visual correlated cues do not speed up detection. This suggests that the clicking sound may result in automatic intersensory integration and highlight relevant visual information in complex visual scenes.

The intersensory integration account would predict that correlated sounds and possibly even random sounds should both facilitate target detection and speed up response times in the current study because of the close temporal presentation of the dot and sound. However, this was not the case. While this will have to be examined in future research, we believe the discrepant findings stem from the complexity of the visual scene. In the current speeded response task, the visual

stimulus should automatically pop out because there is only one stimulus presented at a time. As visual scenes become more complex, such as in the pip and pop task, redundant intersensory information may make certain parts of the visual scene become more salient. This account is somewhat consistent with the Intersensory Redundancy Hypothesis (Bahrick & Lickliter, 2000), which states that intersensory redundancy automatically directs attention to the redundant information. For example, when infants are required to learn amodal information such as the tempo of a hammer tapping on a table, presenting this information visually and auditorily facilitates learning compared to when the same tempo is only presented visually or auditorily (Bahrick & Lickliter, 2000).

In summary, while most of the research in adult populations points to visual dominance (see Sinnett, Spence, & Soto-Faraco, 2007; Spence, Parise, & Chen, 2012 for reviews), the current study provides some support for auditory interference on a task that is typically better suited for the visual modality. These findings have implications for tasks that require quick processing and responding to multisensory information and shed light on potential mechanisms underlying modality dominance effects.

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.
- 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.
- 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 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*(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.
- 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.
- 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.