The role of modality congruence in the presentation and recognition of task-irrelevant stimuli in dual task paradigms.

Maegen Walker (maegenw@hawaii.edu)
Department of Psychology, University of Hawaii at Manoa
2530 Dole Street, Honolulu, HI 96822 USA

Andrew Dewald (adewald@hawaii.edu)
Department of Psychology, University of Hawaii at Manoa
2530 Dole Street, Honolulu, HI 96822 USA

Scott Sinnett (ssinnett@hawaii.edu)
Department of Psychology, University of Hawaii at Manoa
2530 Dole Street, Honolulu, HI 96822 USA

Abstract

Explicitly presented task-irrelevant targets are facilitated in a later recognition test, provided they frequently appear synchronously with targets from a previously presented relevant task (Dewald & Sinnett, 2013). This dual task paradigm was used to test the relationship between the modality of which a primary task was presented, and the modality of a subsequently presented secondary task. Earlier findings suggest that cross-modal presentations lead to higher facilitation rates for items that were previously aligned with auditory targets when compared with only unimodal (auditory or visual) presentations. The current study extends these findings to conditions where the primary task is presented visually, while testing later word recognition in either the same (visual), different (auditory), or across (audiovisual) modalities. Overall, target-aligned information was recognized at significantly higher rates than non-aligned information for all three recognition tests. Critically, when comparing the magnitude of facilitation, cross-modal presentation resulted in the highest degree of facilitation.

Keywords: Attention, Dual Task Paradigms, Implicit Learning, Inattentional Blindness

Introduction

The degree to which task-irrelevant and unattended stimuli are processed has been investigated using paradigms commonly employed in the exploration of visual perceptual learning and inattentional blindness (Dewald & Sinnett, 2011a, 2011b; Dewald, Sinnett, & Doumas, 2011, 2013; Rees, Russell, Firth, & Driver, 1999; Seitz & Watanabe, 2003, 2005; Sinnett, Costa, & Soto-Faraco, 2006; Tsushima, Sasaki, & Watanabe, 2006; Tsushima, Seitz, & Watanabe, 2008; Watanabe, Nàñez & Sasaki, 2001). This body of work has demonstrated that unattended, task-irrelevant stimuli may be processed if presented at the same time as a task-relevant target in an attention-demanding task. Interestingly though, these studies yield seemingly contradictory results, indicating that later recognition of the unattended stimuli will either be facilitated (Dewald & Sinnett, 2011b; Dewald et al., 2013; Seitz & Watanabe, 2003; Watanabe et al., 2001) or inhibited (Dewald & Sinnett 2011a; Dewald et al., 2011; Tsushima et al., 2006; Tsushima et al., 2008) depending on various factors such as the frequency of exposure and whether the stimuli was presented above or below the threshold for explicit awareness.

Regardless of this documented dichotomy of a facilitatory or inhibitory relationship, the various paradigms employed are fundamentally the same in that participants are required to pay attention to specific stimuli while simultaneously ignoring irrelevant information. Later recognition of the irrelevant items is then assessed via some variation of a surprise recognition task. Crucially, the irrelevant items in the recognition task could have originally appeared at the same time (*target-aligned*), or not (*non-aligned*), as a task-relevant target presented amongst the attended stimuli in the primary task. Initial research suggested that facilitation or inhibition was dependent on whether the irrelevant target item was presented either below or above (respectively) the threshold for explicit awareness. For instance, the seminal finding by Watanabe et al. (2001) used a presentation of dynamic random dot (DRD) displays (the irrelevant stimuli) in which a small subset (5%) of the otherwise randomly moving dots (subthreshold to explicit awareness) moved coherently in the same direction. While exposed to this display, participants also engaged in an attention demanding target identification task (i.e., letter identification) presented in a rapid serial visual presentation (RSVP). Subsequently, a motion identification task was presented that required participants to determine the direction of moving dots in a similar DRD display. Ultimately, the findings demonstrated that identification for dots that moved in the same direction as subthreshold motions, temporally aligned (*target-aligned*) with the presence of the task-relevant target in the letter identification task, was significantly better than for motions not paired with task-relevant targets (see also, Seitz & Watanabe, 2003). These results were taken as evidence that frequent exposure to task-irrelevant stimuli that are temporally aligned with task-relevant targets, results in learning effects for the irrelevant information, provided that it was originally presented below awareness and in synchronous temporal pairing with a target task.

Given that Watanabe et al. (2001) and Seitz and Watanabe (2003) only found facilitation for motion presented at subthreshold levels, a naturally ensuing question would be what happens with suprathreshold presentations using the same paradigm. Precisely addressing this question, Tsushima et al. (2008) systematically varied the salience of the *target-aligned*, but irrelevant, motion coherence in an attempt to determine the effects of stimulus saliency on irrelevant target learning rates. Remarkably, the facilitation that was previously observed for subthreshold presentations disappeared after suprathreshold presentations. This suggests that when task-irrelevant stimuli are presented in synchronous temporal pairing at a salience level that is *above* the threshold for explicit awareness, diminished learning effects may be observed.

Recently, Dewald et al. (2011, 2013) extended this work with a more complex and salient stimuli (words), by utilizing a procedure employed by Rees et al. (1999; see also Sinnett et al., 2006) in which participants viewed a Rapid Serial Visual Presentation (RSVP) stream of words superimposed on top of pictures. Participants were required to attend to the pictures (ignore the superimposed words) and identify an immediate picture repetition. After this repetition detection task, participants were given a surprise recognition test in which they were asked to identify a series of words from the experiment intermixed with novel foil words. Paralleling the findings of Tsushima et al. (2008), the resulting data supported that words, temporally aligned with picture repetitions, were recognized at levels significantly *below* chance when compared with *non-aligned* words, which suggests that *target-aligned* words were actually inhibited. These findings have also been extended to the auditory modality. For instance, Dewald and Sinnett (2011a) found similar inhibitory results for *target-aligned* spoken words paired with a stream of sounds, while performing an isomorphic version of the experiment in the auditory modality.

Given that previous research only presented a frequently occurring, unchanging, single motion (see Tsushima et al., 2008; Watanabe et al., 2001), Dewald et al. (2013) lowered the total number of presented words such that only one unchanging word was paired with picture repetitions in the RSVP stream (highly frequent exposure). Interestingly, recognition rates for both *target-aligned* and *non-aligned* words were significantly *above* chance. However, *target-aligned* words were unambiguously recognized at rates significantly higher than *non-aligned* words. These findings suggest that, in addition to explicit or implicit presentations and synchronization with an attended target, the rate of exposure is also a critical element in understanding how irrelevant but *target-aligned* items are processed.

An invariant feature of all of these investigations, as well as a feature of seminal studies of attention and perception, is that tests of perception for unattended stimuli (i.e., recognition) are always provided in a congruent modality to the initial exposure (Broadbent, 1954, 1961; Cherry, 1953). It is plausible that different sensory modalities of presentation between exposure and recognition could foster a different trend in the findings. Indeed, this was precisely what happened when Dewald and Sinnett (2012) tested this very notion. That is, when the initial presentation of the repetition detection task was presented in an auditory modality, facilitation for *target-aligned* words surfaced when the surprise test was presented in the auditory modality but not when presented visually. Interestingly, cross-modal presentation of the surprise recognition test lead to a significant enhancement in the magnitude of facilitation. With the exception of the previous work in the auditory modality (Dewald & Sinnett, 2012), all research involving this paradigm has presented the surprise recognition task in the visual modality regardless of modality presentation during the initial repetition detection task. It is still unknown if these findings will extend to the dominate sensory modality in humans, vision (Chandra, Robinson & Sinnett, 2011; Colavita, 1974; Posner, Snyder & Davidson, 1980; Sinnett, Spence & Soto-Faraco, 2007).

In order to test the robustness of facilitated word recognition for cross-modal presentations, the same paradigm employed by Dewald and Sinnett, (2012) was used here, with the main difference being that the primary task was presented in the visual modality as opposed to the auditory modality. In order to test if facilitation levels are modulated by whether the surprise recognition test occurs in the same or different modality of the primary task, or if cross-modal presentations enhance the effect, the recognition task was presented in the visual or auditory modality, or across modalities (respectively). If the modality of presentation for the primary and surprise tasks is of crucial importance, then we should see enhanced recognition for *target-aligned* words when they are presented in a congruent modality (i.e., vision) in the surprise recognition test when compared with an incongruent modality (i.e., audition). Furthermore, if cross-modal presentations do indeed enhance later recognition of the unattended visual stimuli, we should observe the greatest levels of facilitation when words in the surprise recognition test are presented across modalities.

Methods

Participants

Eighty-three participants (56 females, mean age of 20.6) were recruited from the University of Hawai'i at Manoa in exchange for course credit. Each participant completed the same visual repetition detection task, but were randomly assigned to three different types of surprise recognition tests: visual only (n=28), auditory only (n=27), or cross-modal (n=28). Participants were naïve to the experiment and had normal or corrected to normal vision and hearing.

Stimuli

A total of 50 pictures (on average 5 to 10 cm's) were selected from the Snodgrass and Vanderwart (1980) picture database. These pictures were randomly rotated +/-30 degrees from their original orientation to ensure that the identification task was sufficiently demanding in each version of the experiment (see Rees et al., 1999). Each picture was superimposed with one of eight high frequency English words selected from the MRC psycholinguistic database (Wilson, 1998; average length of 5 letters with a range of 4-6; frequency of 361 per million, range 135-782). Care was taken to ensure that picture-word combinations did not have any semantic relationship. The words were superimposed over the rotated pictures in bold, capitalized letters and presented in Arial font (24 points).

For the exposure stage, a stream of 960 combined picture-word items (height and width not exceeding 10cm) was created. Repeated pictures in the rapid serial visual presentation (RSVP) stream acted as the task relevant-targets in the identification task. The RSVP stream was broken into eight blocks of 120 trials. The presentation was pseudorandomized so that in each block an immediate picture repetition occurred an average of one out of every eight trials, creating a mean of 15 task-relevant targets (picture repetitions) per block. This resulted in a total of 120 trials of exposure to a task-relevant target and a repeated task-irrelevant target word.

Of the eight words that were superimposed over the pictures in the 960 trial RSVP stream, one was randomly selected to appear in temporal alignment with the task-relevant target. In other words, a single word was selected and always paired with the presentation of an immediately repeated picture target. Eight iterations of this experiment were created for which each of the eight words acted as the word that was aligned with the picture repetitions. To control for any possible differences that may have existed with regard to individual word saliency, the presentation was randomized between participants (average of 10 participants per word). This was done to replicate the dependent measure and parallel the quantity of items and exposure to irrelevant stimuli employed by Dewald and Sinnett (2012; see also Watanabe et al., 2001) but with a much larger sample size.

The later surprise recognition test consisted of a total of sixteen words, eight of which came from the previously viewed visual stream, while the other eight consisted of never before seen foil words. Recall that one of the eight previously presented words exclusively appeared only with picture repetitions. The foil words were never used in the exposure stage of the experiment, but were taken from the same database (average frequency of 236 per million; range of 165 to 399). Words that were temporally aligned with task-relevant targets (picture repetitions) will be referred to as *target-aligned words* and those aligned with non task-relevant targets (non-repeating pictures) will be referred to as *non-aligned words* (see Figure 1) (see also Dewald et al., 2012).

The word recognition tasks were randomized and presented by DMDX software (Forster & Forster, 2003) one at a time, in either the visual or auditory modality, or across modalities. For the visual presentation the words were written in bold, capitalized letters in Arial font at a size of 24 points (i.e., identical to their initial presentation in the repetition detection task), and remained on the screen until a response was made. For auditory presentations a native English speaker's voice was recorded reading the list three times, after which three blind listeners chose the best exemplar of each spoken word (a fourth listener was recruited in order to break a tie when needed). The selected recordings were edited to have the same length of presentation (350 ms) and average amplitude (see Sinnett et al., 2006). The auditory surprise recognition task was presented from two speakers, equidistant to the computer screen and the next word was not spoken until a response was made. Cross-modal presentations involved the written word on the screen with the spoken word presented simultaneously. In the cross-modal presentation the word remained on the screen until a response was made.

Procedure

Participants were instructed to ignore the superimposed words (attend only to the pictures) and respond when they saw a picture immediately repeat in the RSVP stream by pressing the 'G' key on the keyboard of the computer. Each item in the picture-word presentation was presented for 350 ms with a 150-ms inter-stimulus interval (ISI; blank screen) between each item for a stimulus onset asynchrony (SOA) of 500 ms (see Figure 1). Before the first experimental block, a training block of eight trials was given and repeated until participants were familiar and comfortable with the task.

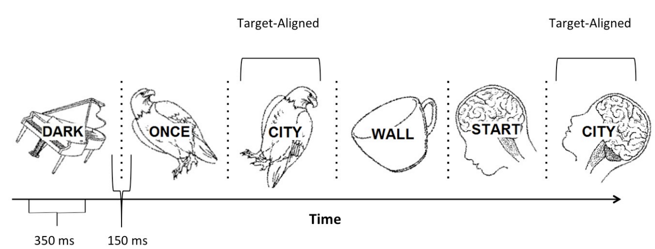

Figure 1: Schematic representation of the task. See text for details.

Immediately after the repetition detection task, the surprise word recognition test was administered to all participants (randomized across participants, visual only (n=28), auditory only (n=27), or cross-modal (n=28)). Participants were instructed to press the "B" key if they had seen the word during the repetition detection task or, instead, the "V" key if they had not seen the word before.

Results

Immediate Repetition Accuracy Overall performance accuracy (across all conditions) of immediate target repetition revealed that participants were successful at detecting target repetitions in the primary task, [hit rate: 74% vs. miss rate: 26%, $t(82) = 15.87$, $p < .001$].

Overall Recognition Accuracy Across all conditions, participants were able to recognize the unattended words during the repetition detection task statistically better than chance (both the *target-aligned* and *non-aligned* words). Collapsed over all surprise tasks, overall performance for word recognition was better than chance [79%, $SE = .019$, $t(73) = 15.29$, $p < .001$]. Recognition for the *target-aligned* words [89%, $SE = .036$, $t(72) = 10.60$, $p < .001$] and the *non-aligned* words was also better than chance [58%, $SE = .029$, $t(72) = 2.93$, $p < .005$]. Critically, *Target-aligned* words were recognized more accurately than *non-aligned* words [$t(72) = 8.71$, $p < .001$].

In order to assess whether later word recognition was modulated by the modality that the surprise test was presented in, as well as target alignment, a two-way, repeated measures, ANOVA was conducted with modality (visual, auditory or cross-modal) as a between subjects factor and target alignment (*target-aligned* or *non-aligned*) as a within subjects factor. No main effect for modality presentation type was observed, [$F(2, 70) = .90$, $p = .413$]. However, a main effect for target alignment demonstrated that *target-aligned* word recognition (79%) was significantly better than *non-aligned* (58%), [$F(1, 70) = 78.97$, $p < .001$]. An interaction was not observed [$F(2, 70) = 2.41$, $p = .097$] (see Figure 2).

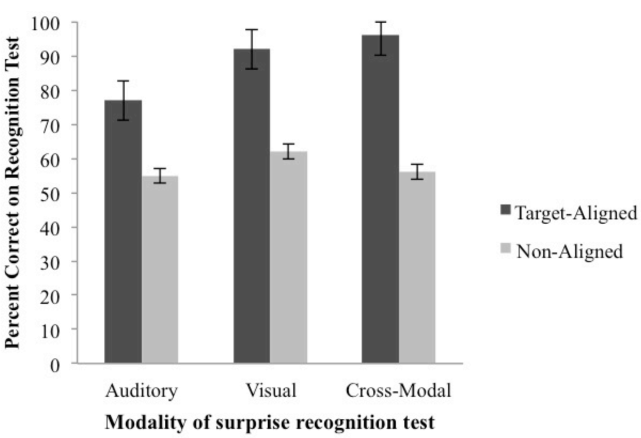

Figure 2: Recognition rates for target-aligned words (dark grey bar) and non-aligned words (light grey bar) according to modality of recognition test.

To further assess the magnitude of the enhancement for alignment in each modality, planned comparisons of the enhancement for *target-aligned* words across each modality of presentation were conducted (see Figure 3). No significant differences in the magnitude of alignment facilitation were observed when comparing the visual presentation (29%, $SE = .061$) with the auditory [20%, $SE = .066$, $t(47) = .98$, $p < .330$] or cross-modal condition [39%, $SE = .045$, $t(47) = 1.32$, $p < .191$]. However, the magnitude of enhancement in the auditory modality was significantly less than that of the cross-modal condition [$t(46) = 2.36$, $p < .01$].

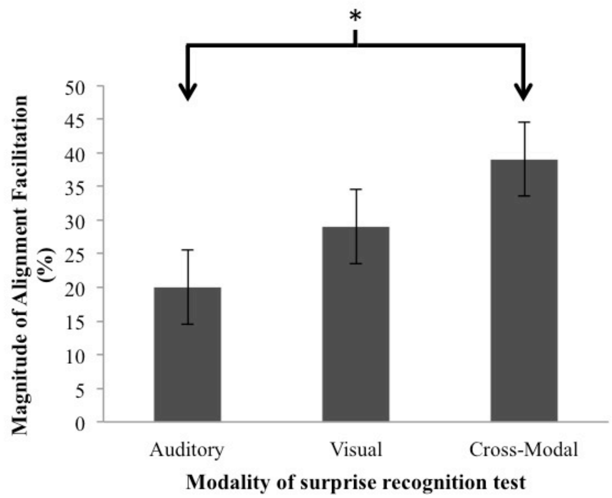

Figure 3: Alignment facilitation for each recognition test. Cross-modal facilitation was significantly better than auditory.

To assess the difference in performance on the three recognition tests, respective of modality, overall performance and individual analysis on each recognition test is provided next.

Visual Surprise Recognition Test (VR*)* Overall recognition performance for when the surprise test was presented in the visual modality only was 77%, which was significantly different from chance [$SE = .041$, $t(49) = 6.61$, $p <$ 001]. Recognition for both *target-aligned* and *non-aligned* words, respectively, was better than chance, [92%, $SE = .055$, $t(24) = 7.58$, $p < .001$] and [62%, $SE = .044$, $t(24) = 2.77$, $p < .011$]. Recognition for *target-aligned* words was significantly better than *non-aligned* words, [$t(24) = 4.86$, $p < .001$].

When comparing the correct rejection (CR) of foil words against word recognition, in the visual modality, performance for both *target-aligned* and *non-aligned* words were again considered. For *target-aligned* words versus correct rejection of foil words, there was no significant difference, [*target-aligned:* 92% vs. CR: 94%, $SE = .02$, $t(24) = .32$, $p = .751$]. When comparing recognition rates for *non-aligned* words and correctly rejecting foil words there was a significant difference, [*non-aligned:* 62%, $SE = .044$ vs. CR: 94%, $t(24) = 7.53$, $p < .001$]. Further exemplifying overall word recognition accuracy, there were significantly fewer false alarms (FA) when compared to correct rejection of foil words, [FA: 6%, *SE* = .020 vs. CR: 94%, $t(24) = 21.39$, $p < .001$], *target-aligned* words [92% $t(24) = 15.47$, $p < .001$], *non-aligned* words [62%, $t(24) = 10.27$, $p < .001$], and chance [$t(24) = 21.39$, p < .001]

Auditory Surprise Recognition Test (AR) Overall recognition performance for word recognition when the surprise test was presented in the auditory modality only was 68%, which was significantly different from chance, [*SE* = .054, $t(47) = 3.40$, $p <$ 001]. Recognition for *target-aligned* words was better than chance, [79%, *SE* = .084, $t(23) = 3.44$, $p < .002$]. Interestingly, recognition rates for *non-aligned* words were not significantly different from chance, [57%, *SE* = .061, $t(23) = 1.25$, $p = .222$]. Recognition for *target-aligned* words was significantly better than *non-aligned* words, [$t(23) = 3.14$, $p < .005$].

For *target-aligned* words versus correct rejection of foil words, there was no significant difference, [*target-aligned*: 79%, *SE* = .084 vs. CR: 84%, *SE* = .045, $t(23) = .51$, $p = .613$], while the same comparison for *non-aligned* words showed a significant difference, [*non-aligned:* 57%, *SE* = .061 vs. CR: 84%, $t(23) = 3.14$, $p < .005$]. Furthermore, there were significantly fewer false alarms (FA) when compared to the correct rejection of foil words, [FA: 15%, *SE* = .045 vs. CR: 84%, $t(23) = 7.48$, $p <$ 001], *target-aligned* words, [79%, $t(23) = 6.99$, $p < .001$*], non-aligned* words [57%, $t(23) = 6.18$, $p < .001$] and chance [$t(23) = 7.48$, $p < .001$].

Cross-Modal Surprise Recognition Test (CR) Overall, performance for word recognition when the surprise test was presented across both modalities was 75%, which was significantly different from chance [*SE* = .042, $t(47) = 6.02$, $p < .001$]. Recognition for *target-aligned* words was better than chance, [95%, *SE* = .041, $t(23) = 11.00$, $p < .001$. The recognition rate for *non-aligned* words was not significantly different from chance, [55%, *SE* = .048, $t(23) = 1.22$, $p < .233$]. Recognition for *target-aligned* words was significantly better than *non-aligned* words, [$t(23) = 8.76$, p < .001].

For *target-aligned* words versus the correct rejection of foil words, there was no significant difference, [*target-aligned*: 95%, *SE* = .041 vs. CR: 91%, *SE* = .025, $t(23) = .80$, $p = .431$], while the same comparison for *non-aligned* words showed a significant difference, [n*on-aligned*: 55%, *SE* = .048 vs. CR: 91%, $t(23) = 6.04$, $p < .001$]. Finally, there were significantly fewer false alarms (FA) when compared to correct rejection of foil words,[FA: 8%, SE = .025 vs. CR: 91%, $t(23) = 16.21$, $p < .001$], *target-aligned* words [95%, $t(23) = 19.11$, $p <$ 001.], *non-aligned* words [55%, $t(23) = 9.43$, $p < .001$] and chance [$t(23) = 16.21$, p < .001.]

Discussion

There are several important findings that merit discussion. First, this study successfully replicated previous findings from Dewald et al. (2012). Specifically, the findings demonstrate the robustness of perceptual learning for task irrelevant stimuli, provided that the stimuli are temporally aligned with targets in a previously presented and attended task.

An equally important finding from this study pertains to the main effect for target alignment. That is, there was an enhanced recognition for *target-aligned* words when compared with *non-aligned* words. While this has been demonstrated previously, it nonetheless lends further support to the idea that temporal alignment with the attended target task and length of exposure to the unattended stimuli are the driving forces behind visual perceptual learning.

There was a critical finding when directly comparing the magnitude of facilitation across modalities. Given previous findings (see Dewald & Sinnett, 2012), we had expected the magnitude difference between *target-aligned* and *non-aligned* words to be greatest after cross-modal presentations (for the surprise test) followed by congruent, then incongruent primary task to surprise test presentations. This hypothesis was partially supported as the magnitude improvement for *target-aligned* vs. *non-aligned* words, in the surprise test, was significantly greater after cross-modal presentations when compared with auditory presentations (39% vs. 20%). Although, it should be noted that visual presentations fell directly in the middle (29%), and was not significantly different form the other conditions. This trend in the overall pattern, coupled with the significant differences between auditory and cross-modal presentations, indicates that cross-modal presentations might bolster an already facilitated recognition of unattended irrelevant stimuli. Moreover, this falls in concert with previous investigations indicating that the cross-modal presentation of information may have decreased task difficulty. This may be due to the existence of individualized attentional reservoirs for different modality systems (Duncan et al., 1997; Sinnett et al., 2006; Soto-Faraco & Spence, 2002; Treisman & Davies, 1973; Wickens, 1980, 1984), which may account for the observed differences between the cross-modal and auditory recognition tests.

There are several contributing factors that may also be able to account for the increased facilitation for target-aligned words after cross-modal presentations when compared with auditory presentations. First, the audio quality and rate of presentation in the auditory condition may have led to lower recognition rates due to participants having difficulty understanding the spoken words. However, this is unlikely given that the same stimuli were used in Dewald and Sinnett (2012), where auditory performance exceeded visual performance. A more likely candidate would be that once the word was spoken in the auditory condition, participants were left with only a blank screen. Contrarily, in the visual and cross-modal conditions the written word was left on the screen until participants made a decision. It could be argued that this resulted in reduced word exposure in the auditory condition while participants decided whether or not they had seen it in the previously completed experiment. This may also have been a contributing factor in the cross-modal condition such that if a participant was unable to understand the spoken word they may have attended solely to the visually presented written word on the screen. Further investigations should attempt to address issues of auditory quality and exposure rates to stimuli in the recognition tests to limit the discrepancy between visual and auditory presentation of the words.

Finally, we have extended investigations on cross-modal stimulus pairing that offers areas of continued exploration regarding the role that modality presentation plays in memory recall and recognition rates for the unattended stimuli. To further explore the impact modality presentation has on perceptual learning of stimuli, later studies might consider focusing on continued manipulations of the initial task to include a cross-modal presentation of stimuli with varying modality presentation in the surprise recognition tests.

References

Broadbent, D. E. (1954). The role of auditory localization in attention and memory span. *Journal of experimental psychology*, *47*(3), 191.

Broadbent, D. E., & Gregory, M. (1961). On the recall of stimuli presented alternately to two sense-organs. *Quarterly Journal of Experimental Psychology*, *13*(2), 103-109.

Chandra, M., Robinson, C. W., & Sinnett, S. (2011). Coexistence of multiple modal dominances. In *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 2604-2609).

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the acoustical society of America*, *25*, 975.

Colavita, F. B. (1974). Human sensory dominance. *Perception & Psychophysics*, *16*(2), 409-412.

Dewald, A.D., & Sinnett, S. (2013). Speed facilitation in the absence of enhanced recognition for target-aligned but irrelevant stimuli under cross-modal presentations. *Proceedings of the Thirty-Fifth Annual Conference of the Cognitive Science Society,* 2183-2188.

Dewald, A. D., & Sinnett, S. (2012). Enhanced Performance for Recognition of Irrelevant Target-Aligned Auditory Stimuli: Unimodal and Cross-modal Considerations. *Proceedings of the Thirty-Fourth Annual Conference of the Cognitive Science Society,* 294-299.

Dewald, A. D., Sinnett, S., & Doumas, L. A. (2013). A window of perception when diverting attention? Enhancing recognition for explicitly presented, unattended, and irrelevant stimuli by target alignment. *Journal of Experimental Psychology: Human Perception and Performance*, 39(5), 1304-1312.

Dewald, A. D., & Sinnett, S. (2011a). An inhibited recognition performance for explicitly presented target-aligned irrelevant stimuli in the auditory modality. In *Proceedings of the Thirty-Third Annual Conference of the Cognitive Psychology Society*, 1158-1163.

Dewald, A. D., & Sinnett, S. (2011b). A Multimodal Investigation of Recognition Performance for Target-Aligned but Irrelevant Stimuli. In *Proceedings of the Thirty-Third Annual Conference of the Cognitive Science Society*, 1164-1169.

Dewald, A.D., Sinnett, S., & Doumas, L.A.A. (2011). Conditions of directed attention inhibit recognition performance for explicitly presented target-aligned irrelevant stimuli. *Acta Psychologica*, 138, 60-67.

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

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*(1), 116-124.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of experimental psychology: General*, *109*(2), 160.

Rees, G., Russell, C., Frith, C. D., & Driver, J. (1999). Inattentional blindness versus inattentional amnesia for fixated but ignored words. *Science*, 286, 2504-2507.

Seitz, A.R., Kim, R., & Shams, L. (2006). Sound facilities visual learning. *Current Biology*, 16(14), 1422-1427.

Seitz, A. R. & Watanabe, T. (2003). Psychophysics: Is subliminal learning really passive? *Nature*, 422, 36.

Seitz, A. R. & Watanabe, T. (2005). A unified model for perceptual learning. *Trends in Cognitive Science*, 9(7), 329-334.

Sinnett, S., Costa, A., & Soto-Faraco, S. (2006). Manipulating inattentional blindness within and across sensory modalities. *Quarterly Journal of Experimental Psychology*, 59(8), 1425-1442.

Sinnett, S., Spence, C., & Soto-Faraco, S. (2007). Visual dominance and attention: the Colavita effect revisited. *Perception & Psychophysics*, *69*(5), 673-686.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 174–215.

Soto-Faraco, S., & Spence, C. (2002). Modality-specific auditory and visual temporal processing deficits. *The Quarterly Journal of Experimental Psychology: Section A*, *55*(1), 23-40.

Treisman, A. M., & Davies, A. (1973). Divided attention to ear and eye. *Attention and performance IV*, 101-117.

Tshushima, Y., Sasaki, Y., & Watanabe, T. (2006). Greater disruption due to failure of inhibitory control on an ambiguous distractor. *Science*, 314, 1786-1788.

Tsushima, Y., Seitz, A. R., & Watanabe, T. (2008). Task-irrelevant learning occurs only when the irrelevant feature is weak. *Current Biology*, 18(12), 516-517.

Watanabe, T., Náñez, J. E., & Sasaki, Y. (2001). Perceptual learning without perception. *Nature*, *413*(6858), 844-848.

Wickens, C. D. (1980). The structure of attentional resources. *Attention and performance VIII*, *8*.

Wickens, C. D. (1984). Varieties of attention. *Varieties of attention*, 63-102.

Wilson, M. D. (1988). The MRC psycholinguistic database: Machine readable dictionary, version 2. *Behavioural Research Methods, Instruments and Computers*, 20, 6-11.