Independent effects of bottom-up temporal expectancy and top-down spatial attention. An audiovisual study using rhythmic cueing

Alexander Jones *

Department of Psychology, Middlesex University, London, UK

Selective attention to a spatial location has shown enhanced perception and facilitate behavior for events at attended locations. However, selection relies not only on where but also when an event occurs. Recently, interest has turned to how intrinsic neural oscillations in the brain entrain to rhythms in our environment, and, stimuli appearing in or out of sync with a rhythm have shown to modulate perception and performance. Temporal expectations created by rhythms and spatial attention are two processes which have independently shown to affect stimulus processing but it remains largely unknown how, and if, they interact. In four separate tasks, this study investigated the effects of voluntary spatial attention and bottom-up temporal expectations created by rhythms in both unimodal and crossmodal conditions. In each task the participant used an informative cue, either color or pitch, to direct their covert spatial attention to the left or right, and respond as quickly as possible to a target. The lateralized target (visual or auditory) was then presented at the attended or unattended side. Importantly, although not task relevant, the cue was a rhythm of either flashes or beeps. The target was presented in or out of sync (early or late) with the rhythmic cue. Results showed participants were faster responding to spatially attended compared to unattended targets in all tasks. Moreover, there was an effect of rhythmic cueing upon response times in both unimodal and crossmodal conditions. Responses were faster to targets presented in sync with the rhythm compared to when they appeared too early in both crossmodal tasks. That is, rhythmic stimuli in one modality influenced the temporal expectancy in the other modality, suggesting temporal expectancies created by rhythms are crossmodal. Interestingly, there was no interaction between top-down spatial attention and rhythmic cueing in any task suggesting these two processes largely influenced behavior independently.

Keywords: entrainment, crossmodal, endogenous, exogenous, attention, expectancy, hazard function

INTRODUCTION

Our sensory system is constantly exposed to vast amounts of information. To efficiently deal with this information and guide behavior we need to select, prioritize and predict certain events and stimuli over others. The collective term for this selective mechanism is known as attention. There are many forms of attention and one of the most extensively researched is how we focus our attention towards different locations in space. Spatial attention research is typically divided into endogenous attention, which is under voluntary control and exogenous attention which is bottom-up and stimulus driven. The most common method to explore the behavioral effects of endogenous and exogenous attention has been using the Posner cueing task (Posner, 1980). The participant's task is to respond as quickly as possible to a target, usually presented peripherally to the left or right. In an endogenous version, the targets are preceded by a cue, usually centrally located, informing the most likely location of the target (70–80% likelihood). In an exogenous version, the cue, usually

peripheral, does not give any indication of where the target may appear, however, the cue nevertheless typically elicits effects on target processing (Santangelo and Spence, 2008). Endogenous spatial attention has been studied extensively within and across modalities. Selective attention to a spatial location has shown to enhance perceptual processing (e.g., Mangun and Hillyard, 1990; Yeshurun and Carrasco, 1998) as well as facilitate response times (e.g., Posner et al., 1980) to visual stimuli at attended as compared to unattended locations (for a recent review see Carrasco, 2014).

Predictions about events in our environment rely not only on *where* something happens but also *when* an event occurs. Similar to spatial attention, focusing attention to a specific moment in time influences perception and biases our actions (Nobre, 2010). Temporal expectation can be generated in different ways and similar to voluntary spatial attention, instructive cues have been used to manipulate temporal expectations (e.g., Coull and Nobre, 1998; Naccache et al., 2002; Davranche et al., 2011; Zanto et al.,

2011; Rohenkohl et al., 2014; see Nobre and Rohenkohl, 2014 for a recent review). Coull and Nobre (1998) used a symbolic cue to indicate when an upcoming target would likely appear, either 300 ms or 1500 ms after cue onset. In this detection task they found behavioral benefits when the cue appeared at a temporally anticipated compared to an unexpected time interval. Zanto et al. (2011) extended these findings showing similar benefits of voluntary orienting to targets at a particular point in time using event-related potentials (ERPs) as well as behavioral discrimination and Go-NoGo tasks. Moreover, Correa et al. (2005) showed high temporal expectancies increased perceptual sensitivity (d') for detecting visual targets. Temporal cueing studies in the auditory modality have also shown effects of perceptual modulation by temporal cues. Several studies have observed a modulation of the early N1 component, suggested to originate from the primary auditory cortex, in response to temporally expected compared to unexpected tones (Lange and Röder, 2006; Lange et al., 2006; Lampar and Lange, 2011). There is thus mounting evidence showing that voluntary directing attention to a specific point in time influences both perception and modulates behavior.

Temporal expectancies can also be created by rhythms, something which commonly appears in our environment. For example the rhythm of breathing or our heartbeat, or the swaying of a tree, the sound and movement of walking or waves on a beach, the rhythmic structure of speech, or of course the rhythm in music. Rhythmic cueing has been used to investigate how external rhythms influences perception and performance. For example, Jones et al. (2002) presented participants with a standard tone which was followed by a sequence of tones presented in a rhythm. The participant's task was to judge whether a target tone had the same pitch as the standard tone. They found that performance accuracy was better when the target tone was presented *in* compared to *out of* sync with the preceding rhythm. Auditory perceptual discrimination has consistently shown to be better when stimuli coincide with the rhythm and perceptual performance deteriorates if the stimuli is presented too early or too late in relation to the rhythm (Jones et al., 2006; see Jones, 2010 for a review). Similarly, response times have also been reported to be improved for stimuli occurring on the beat of a particular rhythm compared to an asynchronous rhythm using both auditory (e.g., Sanabria et al., 2011) or visual stimuli (Doherty et al., 2005; Rohenkohl et al., 2012; Cravo et al., 2013).

More recently the concept of rhythmic cueing has seen an increased research interest from a more neuroscientific viewpoint in that intrinsic brain operations are profoundly rhythmic (Raichle, 2010). Groups of neurons in the brain fluctuate rhythmically together and create oscillations with different frequencies which can be measures using electroencephalogram (EEG). These self-generated brain oscillations have shown to modulate responses and influence motoric, perceptual and cognitive processes (Buzsaki, 2006; Thut and Miniussi, 2009). It has for example been shown that the threshold of detecting visual stimuli fluctuates over time along with the phase of ongoing EEG activity (Busch et al., 2009). Importantly, the neural oscillation can also entrain to external rhythms aligning the firing pattern according to rhythms in our environment (Arnal and Giraud, 2012). In

other words, neurons start to fire in synchrony with external rhythms. Moreover, entrainment to particular rhythms has been suggested to underlie selective attention (Lakatos et al., 2013; Calderone et al., 2014). For example Lakatos et al. (2008) presented monkeys with auditory and visual interleaved rhythms and found selectively attention to one stream amplified neural responses to events in that stream. Moreover, entrainment has been shown to increase with participant effort (Lakatos et al., 2013). This further indicates entrainment can also be modulated by higher level processes, such as attention. Similar to spatial attention, temporal expectancies can be bottom-up or top-down. However, it remains to be fully established to what extent rhythmic cueing and entrainment occurs unintentionally, in a purely bottom-up fashion. Recent evidence suggests temporal expectancies do still occur as a result of a rhythm even though the rhythm is detrimental to the task, suggesting automatic effects of rhythmic stimuli in the absence of top-down processes (Breska and Deouell, 2014).

Predicting where or when an event will occur has independently been shown to influence perception and drive behavior. However, space and time are not dimensions which occur in isolation in our environment, yet only a handful of studies have explored these two types expectation together. Doherty et al. (2005) manipulated both temporal and spatial expectancies by presenting participants with a ball which moved from left to right across a screen. Towards the right side of the screen there was a section which occluded the ball before reappearing. Doherty et al. found that response times were faster when the ball reappeared behind the occluding band in sync with the preceding rhythm. Similarly, response times were faster when the ball reappeared in the spatial location which was predicted by the balls trajectory across the screen. The individual effects were also additive showing faster response times when both temporal and spatial expectancies matched, an additive effect also demonstrated on the visual P1 component. Recently Rohenkohl et al. (2014) also investigated the synergy between spatial and temporal expectancies. In their task a symbolic visual arrow simultaneously indicated the likely location of a target as well as the likely time point when to expect the target. Unlike Doherty et al. they found an interaction between spatial and temporal effects. Temporal expectations improved visual perception, but only at spatially attended and not unattended locations. Importantly, in both studies (see also Tang et al., 2013) participants were asked to use both types of expectancies to increase performance. That is, both temporal and spatial expectancies were generated top-down, or, in Doherty et al. (2005) study using a rhythmic cue, a likely mix of stimulus and voluntary temporal attention. What remains less clear is how stimulus driven temporal expectancies, created by external rhythms, are affected by top-down spatial attention.

Crossmodal spatial attention effects have been extensively reported and shown to enhance perceptual processing and facilitate behavior (Vroomen and de Gelder, 2000; Spence and Driver, 2004). However, less is known how entrainment operates across modalities. In a study by Lakatos et al. (2007) it was observed that somatosensory inputs can reset the phase of the neural oscillations in primary auditory cortex of macaque monkeys, and

in turn, auditory stimuli are enhanced or suppressed according to when in the oscillation they appear (see Kayser et al., 2008, for similar results in humans). This observation indicating that oscillations show crossmodal effects at a neural level. Moreover, recently Miller et al. (2013) also showed a crossmodal effect of entrainment whereby eye movements towards a visual target were faster if they occurred in sync with a preceding rhythm of tones.

The present study investigated how voluntary spatial attention affected automatic effects of rhythmic cueing using a simple detection task. Participants performed a typical Posner cueing task where an informative cue indicated to which side the target was most likely to appear. In addition, the cue consisted of four or five stimuli presented in a rhythm and the target was presented in or out of sync with this rhythm. Importantly this rhythm and the timing of the target was not task relevant. This novel paradigm allowed independent manipulation of top-down spatial attention and bottom-up temporal expectancies in order to investigate whether these represent dependent or independent mechanisms in driving behavior (as measured by response times). Furthermore, this study aimed to investigate whether rhythmic stimuli in one modality automatically influence the temporal expectancy in another modality. In separate tasks, participants were either presented with a visual cue and a visual target (VV), auditory cue and auditory target (AA) or in a crossmodal setting with a visual cue and an auditory targets (VA), or auditory cue and visual target (AV). Taken together, this study explored how endogenous spatial attention and stimulus-driven temporal expectancies, two processes which have independently shown to modulate perception and behavior, affected behavior in both a unimodal and crossmodal setting.

METHODS

PARTICIPANTS

The study consisted 16 participants in each task, 64 in total (16 males; 13 right-, 3 left-handed, and 48 females; 43 right-, 5 left-handed). The participants were naive to the study and participated voluntarily or in return for course credits. The participant number was based on similar behavioral studies (e.g., Lawrence and Klein, 2013). Each participant only took part in one of the following four tasks: AA, AV, VA, VV. Due to excessive responses to catch trials and/or an inability to perform the task two participants were removed and replaced in the AV task, one from the VA, and two from the VV. The study was approved by the Middlesex University ethics committee and all participants provided written informed consent.

STIMULI AND MATERIALS

The stimuli were presented and data collected using E-Prime v2 software (Psychology Software tools) run on a PC. Visual stimuli (fixation cross, visual cues and targets) were presented on a 17 inch monitor (1280 × 1024 pixels). A black fixation cross was presented in the middle of the screen. The visual cue consisted of an X above, below, to the left and right of the fixation cross creating the appearance of a larger cross in the center (see **Figure 1** for details). Visual targets (three black Xs) were presented to the left or right side of the monitor. The
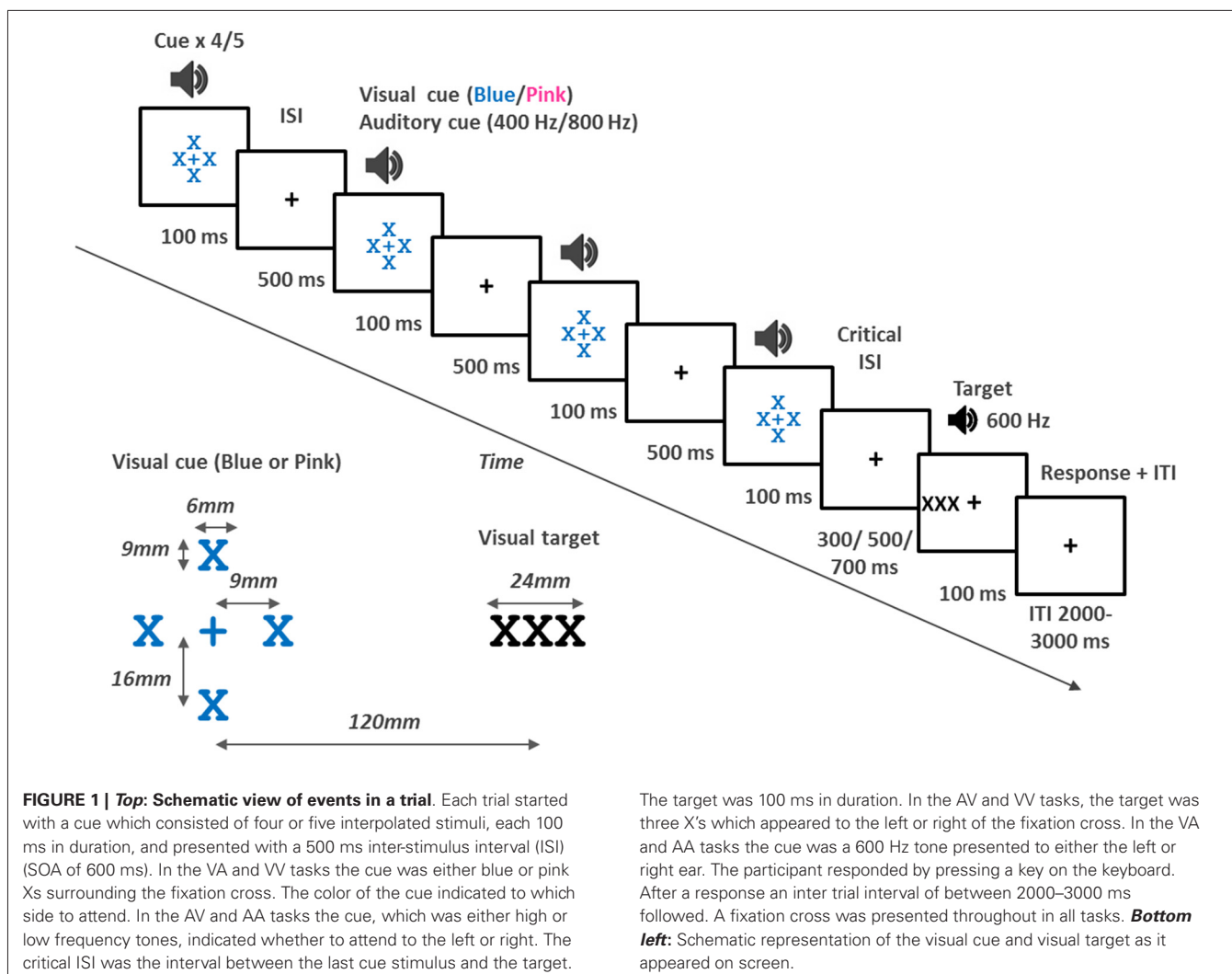
font was Courier new. The participant was seated with their eyes approximately in line with the fixation cross and approximately 400–500 mm away from the screen. The visual angle for the target typically ranged between 18.15 and 14.7°. Auditory stimuli were presented via headphones (Audio 355, Plantronics). Auditory stimuli were 100 ms in duration with a 5 ms rise and fall time. The cue consisted of either low tones (400 Hz) or high frequency tones (800 Hz) and always presented in stereo. Targets were presented to only one ear and were 600 Hz. A keyboard was used to collect response times. The *down arrow* key was positioned in a straight line behind the fixation cross.

DESIGN AND PROCEDURE

On each trial a rhythmic cue was presented. In the VA and VV task the color of the flashes (pink or blue) indicated whether the participant was to direct attention to the left or right. In the AV and AA tasks, the cue was either high or low tones, and this indicated which side attention was to be directed. A target then appeared at the attended (75%) or unattended side (25%) and the participant was to respond as quickly as possible by pressing the keyboard once a target appeared. In the AV and VV tasks, the target appeared to the left or right of the fixation cross and in the VA and AA tasks, the target was a tone presented to the left or right ear. The target appeared either in sync with the rhythm (the cue) or out of sync (early or late). The participant was not informed about this and it was not relevant to the task.

Each task consisted of five blocks with a total 260 trials (52 trials per block). Out of the 260 trials, 180 were attended (69%) and 60 unattended (23%), and 20 catch trials (8%). The weighting of targets, excluding catch trials was, 75/25 for attended/unattended targets. There was an even distribution of early, sync, and late trials. That is, for attended trials, there were 60 early, 60 sync, and 60 late trials per participant, and 20 unattended trials for each of the early, sync and late conditions. For half of the trials the cue consisted of four stimuli, and for half of the trials the rhythmic cue included five stimuli. Prior to the experiment the participant ran a practice block.

The participant was seated in an experimental booth in front of a PC monitor. In the tasks including auditory stimuli (all but the VV task) the participant wore headphones. Each trial started with the presentation of the rhythmic cue which consisted of four or five stimuli presented every 600 ms (see **Figure 1** for events in a trial). More specifically, the first of the rhythmic stimuli (the cue) was presented for 100 ms followed by an inter-stimulus interval (ISI) of 500 ms. The cue stimuli were presented four of five times creating a rhythm of 100 ms stimulus every 600 ms. After the last of the interpolated stimuli a target was presented. The critical ISI, preceding the target was 300 ms (early), 500 ms (sync), or 700 ms (late). The target was then presented for 100 ms. Participants responded with their dominant hand by pressing the *down arrow* key on the keyboard. If no response was recorded the trial terminated after 2000 ms. There was a random inter-trial interval of 2000–3000 ms. A centrally located fixation cross was presented throughout and participants were explicitly instructed to keep their gaze on this fixation cross at all times.

FIGURE 1 | Top: Schematic view of events in a trial. Each trial started with a cue which consisted of four or five interpolated stimuli, each 100 ms in duration, and presented with a 500 ms inter-stimulus interval (ISI) (SOA of 600 ms). In the VA and VV tasks the cue was either blue or pink Xs surrounding the fixation cross. The color of the cue indicated to which side to attend. In the AV and AA tasks the cue, which was either high or low frequency tones, indicated whether to attend to the left or right. The critical ISI was the interval between the last cue stimulus and the target.

The target was 100 ms in duration. In the AV and VV tasks, the target was three X's which appeared to the left or right of the fixation cross. In the VA and AA tasks the cue was a 600 Hz tone presented to either the left or right ear. The participant responded by pressing a key on the keyboard. After a response an inter trial interval of between 2000–3000 ms followed. A fixation cross was presented throughout in all tasks. **Bottom left**: Schematic representation of the visual cue and visual target as it appeared on screen.

In the tasks with an auditory cue (AA and AV), for half the participants high frequency tones indicated to attend to the left and low tones indicated attend to the right. This allocation was reversed for the other half. Similarly, for half the participants who performed a task with a visual cue (VA and VV), a blue Xs indicated attend to the left and pink Xs attend to the right, and the reverse for the other half of participants.

The RT data was Log_{10}-transformed and submitted to a mix design ANOVA with the factors Task (AA, AV, VA, VV), Spatial attention (attended, unattended), Temporal expectancy (early, sync, late), and Rhythm count (four stimuli, five stimuli). Following the overall analysis, each task was analyzed separately.

RESULTS

SUMMARY
The results showed that participants responded faster to attended compared to unattended trials in all four tasks. There was also an effect of temporal expectancy in all tasks but the unimodal auditory task (AA). In the two cross-modal tasks (AV and VA) the targets were faster when in sync and late targets compared

to early targets. In the visual task the late targets were faster than both in sync and early targets. Although clear effects of spatial attention and temporal expectancy were observed there was no evidence of an interaction between these two factors in any task. Thus suggesting temporal expectancy and spatial attention affected response times independently.

Overall analysis including task
A mixed design ANOVA, including Task as a factor, showed a main effect of Spatial attention ($F_{(1,60)} = 36.32$, $p < 0.001$, $\eta_p^2 = 0.38$) with faster RTs for attended (325.2 ms) compared to unattended trials (367.6 ms). There was a main effect of Temporal expectancy ($F_{(2,120)} = 27.27$, $p < 0.001$, $\eta_p^2 = 0.31$) and follow up pairwise-comparisons (Bonferroni corrected) showed that sync (342.8 ms) and late targets (339.7 ms) were significantly faster than early targets (356.6 ms) (both p's < 0.001). There was no difference between sync and late targets ($p = 0.18$). There was a main effect of Rhythm count ($F_{(1,60)} = 42.32$, $p < 0.001$, $\eta_p^2 = 0.41$) with targets preceded by four stimuli in the rhythm (352.8 ms) were slower compared to if the rhythmic cue contained five stimuli
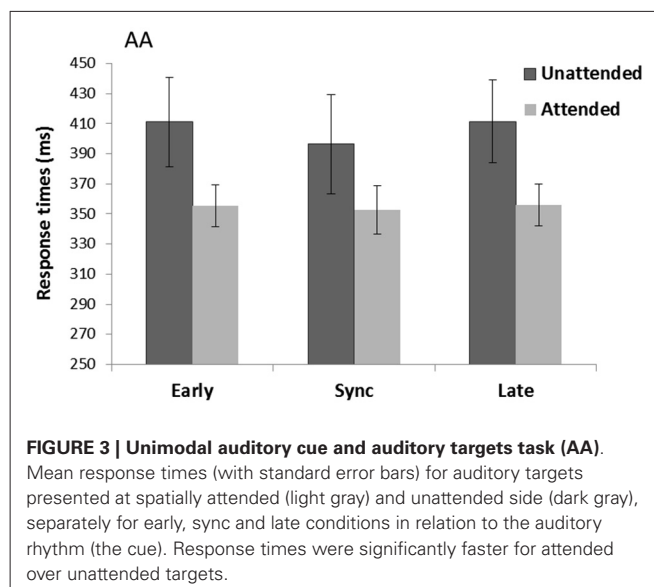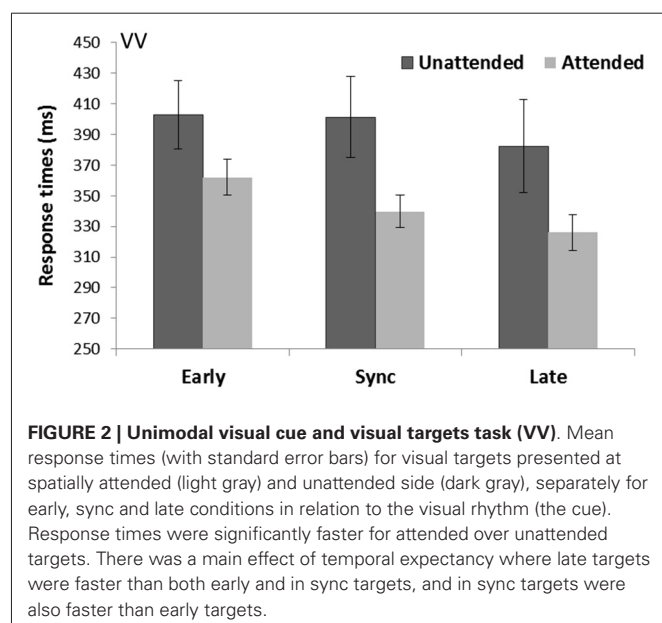
(340.0 ms). There was a main effect of Task ($F_{(3,60)} = 6.24$, $p = 0.001$, $\eta_p^2 = 0.24$) and Bonferroni *post hoc* test showed that the AV task was significantly faster (297.8 ms) compared to AA (380.5 ms) and VV task (369.1 ms) ($p = 0.002$ and $p = 0.005$ respectively).

There was a Temporal expectancy*Task interaction ($F_{(6,120)} = 3.30$, $p = 0.005$, $\eta_p^2 = 0.14$) and planned analysis for each task is presented below. There was no Task*Spatial attention interaction ($p = 0.59$, $\eta_p^2 = 0.04$). Important to note is there was no Spatial attention*Temporal expectancy interaction ($p = 0.37$, $\eta_p^2 = 0.02$). There was a Rhythm count*Temporal expectancy interaction ($F_{(2,120)} = 10.27$, $p < 0.001$, $\eta_p^2 = 0.15$) suggesting the effect of Temporal expectancy was different according to the number of stimuli in the cue. This interaction will also be explored in the analysis of each task. No other main effects or interaction were significant.

VISUAL CUE—VISUAL TARGET (VV)

Overall participants missed less than 1% of targets and responded to 1.9% of catch trials in the VV task.

There was a significant effect of Spatial attention ($F_{(1,15)} = 5.74$, $p = 0.03$, $\eta_p^2 = 0.28$) with attended trials being faster (342.6 ms) than unattended trials (395.5 ms). The purely visual task showed a main effect of Temporal expectancy ($F_{(2,30)} = 17.60$, $p < 0.001$, $\eta_p^2 = 0.54$). Pairwise comparisons (Bonferroni corrected) demonstrated late targets (339.7 ms) to be faster ($p = 0.002$) than in sync targets (370.6 ms), and late targets were faster ($p = 0.001$) than early targets (382.4 ms), and in sync targets were also faster compared to early targets ($p = 0.049$; see **Figure 2**). There was a main effect of Rhythm count ($F_{(1,15)} = 24.03$, $p < 0.001$, $\eta_p^2 = 0.62$) with on average faster RTs for visual targets preceded by five stimuli (356.9 ms) compared to four stimuli (381.2 ms). There was no Spatial attention*Temporal expectancy interaction ($p = 0.18$, $\eta_p^2 = 0.11$).

FIGURE 3 | Unimodal auditory cue and auditory targets task (AA). Mean response times (with standard error bars) for auditory targets presented at spatially attended (light gray) and unattended side (dark gray), separately for early, sync and late conditions in relation to the auditory rhythm (the cue). Response times were significantly faster for attended over unattended targets.

AUDITORY CUE—AUDITORY TARGET (AA)

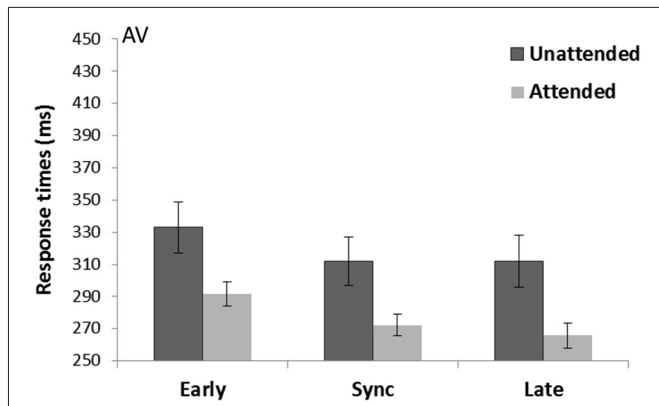Overall participants missed 1.4% of targets and responded to 4.4% of catch trials in the AA task.

There was a significant main effect of Spatial attention ($F_{(1,15)} = 8.99$, $p = 0.009$, $\eta_p^2 = 0.38$) with faster RTs for attended (354.7 ms) compared to unattended trials (406.3 ms) (see **Figure 3**). There was also a main effect of Rhythm count ($F_{(1,15)} = 8.15$, $p = 0.012$, $\eta_p^2 = 0.35$) with overall faster RTs for when the cue consisted of five (376.7 ms) compared to four tones (384.3 ms). There was no main effect of Temporal expectancy ($p = 0.59$, $\eta_p^2 = 0.03$) or Spatial attention*Temporal expectancy interaction ($p = 0.36$, $\eta_p^2 = 0.07$).
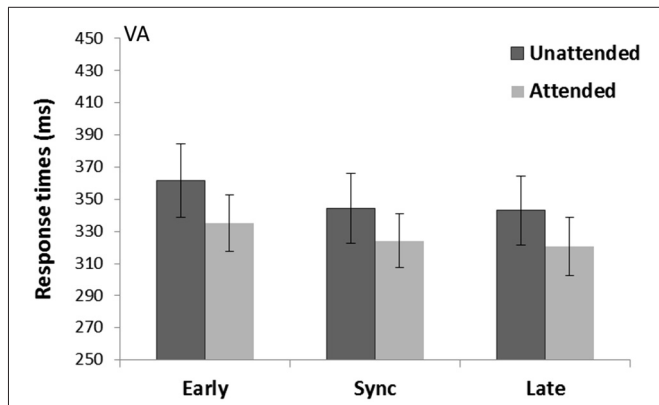
AUDITORY CUE—VISUAL TARGET (AV)

Participants missed 2% of targets and responded to 4.4% of catch trials in the AV task.

A main effect of Spatial attention ($F_{(1,15)} = 26.07$, $p < 0.001$, $\eta_p^2 = 0.64$) showed attended trials were faster (276.6 s) compared to unattended trials (319.0 ms). There was also a main effect of Temporal expectancy ($F_{(2,30)} = 14.38$, $p < 0.001$, $\eta_p^2 = 0.49$). Pairwise-comparisons (Bonferroni corrected) showed both sync (292.2 ms) and late targets (288.9 ms) were significantly faster than early targets (312.4 ms; $p < 0.001$ and $p = 0.001$ respectively; see **Figure 4**). There was no Spatial attention*Temporal expectancy interaction ($p = 0.27$, $\eta_p^2 = 0.08$).

There was a Rhythm count*Temporal expectancy interaction ($F_{(2,30)} = 10.74$, $p = 0.002$, $\eta_p^2 = 0.42$). Follow-up analysis of trials with a rhythm of four tones preceding the visual target showed an effect of Temporal expectancy ($F_{(1,15)} = 20.56$, $p < 0.001$, $\eta_p^2 = 0.58$). Pairwise-comparisons (Bonferroni corrected) showed a difference between early (327.9 ms) and in sync (296.4 ms) and early and late targets (288.8 ms) (both p's < 0.001). When there were five tones in the rhythm, no effect of target Temporal expectancy was present ($p = 0.14$).

FIGURE 2 | Unimodal visual cue and visual targets task (VV). Mean response times (with standard error bars) for visual targets presented at spatially attended (light gray) and unattended side (dark gray), separately for early, sync and late conditions in relation to the visual rhythm (the cue). Response times were significantly faster for attended over unattended targets. There was a main effect of temporal expectancy where late targets were faster than both early and in sync targets, and in sync targets were also faster than early targets.

FIGURE 4 | Crossmodal auditory cue and visual targets (AV) task. Mean response times (with standard error bars) for targets presented at spatially attended (light gray) and unattended side (dark gray), separately for early, sync and late conditions in relation to the auditory rhythm (the cue). Response times were significantly faster for attended over unattended visual targets. There was a main effect of Temporal expectancy where in sync and late targets were faster than early targets.

FIGURE 5 | Crossmodal visual cue and auditory targets task (VA). Mean response times (with standard error bars) for tones presented at spatially attended (light gray) and unattended side (dark gray), separately for early, sync and late conditions in relation to the visually presented rhythm (the cue). Response times were significantly faster for spatially attended over unattended targets and a main effect of Temporal expectancy showed in sync and late targets were faster than early targets.

VISUAL CUE—AUDITORY TARGET (VA)

Overall participants missed less than 1% of targets and responded to 1.6% of catch trials in the VA task.

A main effect of Spatial attention ($F_{(1,15)} = 6.83$, $p = 0.02$, $\eta_p^2 = 0.31$) revealed attended trials were faster (326.6 ms) compared to unattended trials (349.6 ms) (see **Figure 5**). There was also a main effect of Temporal expectancy ($F_{(2,30)} = 9.75$, $p = 0.001$, $\eta_p^2 = 0.39$) and pairwise-comparisons (Bonferroni corrected) showed in sync (370.6 ms) and late targets (354.3 ms) were faster than early target (382.4 ms) ($p = 0.02$ and $p < 0.001$ respectively). There was no Spatial attention*Temporal expectancy interaction ($p = 0.72$, $\eta_p^2 = 0.02$).

DISCUSSION

In four separate tasks, this study investigated the effects of voluntary spatial attention and bottom-up temporal expectancy in both unimodal and crossmodal conditions. In all tasks response times to targets were faster when they appeared at the attended compared to unattended location. This indicated that participants followed instructions and the results replicated what has previously been observed in unimodal visual (e.g., Posner, 1980; Wright and Ward, 1994) and auditory spatial attention tasks (Spence and Driver, 1994), as well as audiovisual crossmodal tasks (Spence and Driver, 1996; or Spence, 2010 for a review). The present study also demonstrated a main effect of temporal expectancy in both crossmodal tasks indicating that rhythmic stimuli in one modality automatically influenced the temporal expectancy in the other modality. An effect of rhythmic cueing was also observed in the unimodal visual task but not unimodal auditory task. Interestingly, there was no observed interaction between top-down spatial attention and temporal expectancy effects in any task suggesting temporal and spatial processing independently affected target response times.

That selectively attending to a spatial location enhances perceptual processing and facilitates behavior at the attended locations has been well documented (for a recent review see Carrasco, 2014). Although relatively less researched, voluntary temporal expectation has also been shown to influence perception and drive behavior (Nobre et al., 2011). Moreover, animal studies have demonstrated that temporal expectations can modulate neural processing in early sensory areas such as primary visual cortex (Lima et al., 2011) and primary auditory cortex (Jaramillo and Zador, 2011). As with spatial selective attention, temporal predictability can be divided into voluntary and stimulus-driven processes. Rhythmic cueing has been used to create temporal predictability, and the rhythm can be used to induce both voluntary or stimulus driven effects, or a combination of both, particularly depending on the instructions. In the present study the rhythmic structure of the cue was not task relevant, in order to investigate whether or not stimulus driven expectancies affected target processing. In other words, whether rhythmic cueing automatically influenced response times even when this temporal characteristic was not relevant or beneficial to the task.

In the unimodal visual task and the two crossmodal conditions the preceding rhythm influenced target detection times. In the visual task, responses to targets in sync with the rhythm were faster compared to early targets, and late targets were also faster compared to when the target was in sync with the rhythm. Similar effects of visual rhythmic cueing have been observed in a study by Rohenkohl et al. (2011) where participants attended to either the color or speed (rhythm) of a moving disc across the screen, to predict an upcoming target. They found both types of cue speeded up response times to targets. However, they found rhythmic cues facilitated response times, compared to an arrhythmic condition, regardless if the participant was instructed to use the rhythmic information or not, whilst the symbolic color cue was only effective when participants explicitly used this information. However, when using rhythmic stimuli to induce temporal expectancies,

it may still be difficult to tear apart the contribution of purely bottom-up effects caused by the rhythm, and any top-down influence such as directing attention to a specific point in time. In other words, are temporal expectancies created passively and purely unintentionally when we are exposed to rhythmic patterns? In some studies using a rhythmic cue, the target was always in sync with the rhythm in the rhythmic condition, as compared to a non-rhythmic condition (e.g., Doherty et al., 2005; Rohenkohl et al., 2012) and in others, the target was more likely to occur in sync compared to out of sync with a rhythm (e.g., Praamstra et al., 2006). In other words, crating conditions where it would be advantageous to use the rhythm to anticipate targets even though not explicitly instructed to do so. Breska and Deouell (2014) specifically investigated whether a rhythm automatically creates temporal expectancies. They included a condition where it was detrimental to use the rhythm to perform the task but still found that the rhythm affected target detection, concluding rhythms automatically exerts an effect on target processing. In the present study the rhythm was not task relevant, but, participants were not explicitly discouraged from using the rhythm. However, the probability of the target appearing at early, sync or late intervals was equally likely and therefore any strategy of expecting the target at a particular time point would not be advantageous. This together with participants concurrently performing another task, directing spatial attention, suggests the rhythmic cueing effects were mainly bottom-up and not involving higher level of processing. However, future studies may wish to specifically address the automaticity of rhythmic cueing and how systematically varying the automaticity is affected by top-down spatial attention. In any case, it can be concluded from the present study that rhythmic cueing effects were observed even though participants performed a concurrent spatial attention task.

In rhythmic cueing studies the perception of targets is typically best when the stimulus coincides with the rhythm and perceptual benefits decrease if the stimulus is presented too early or too late (e.g., Jones et al., 2002; Mathewson et al., 2010). In the current study the detrimental effect of asynchronous stimuli was only present for early, but not for late stimuli. The target was equally likely to appear at early, sync or late intervals meaning there was no strategic benefit in expecting the target at any particular time. However, the pattern of results can in part be explained by the hazard or foreperiod effect. The "hazard function" is an effect whereby an event is more likely to occur at a specific point in time if it has not yet occurred (Luce, 1986). In other words, if the stimulus has not appeared at the early time point it is then more likely to occur at the sync and subsequently the late time interval. This in turn can increase the anticipation and enhance motor readiness. Several steps were taken to account for and to minimize this potential bias. First, catch trials were used whereby no target was presented. This introduces the possibility that if a target has not occurred in sync with the rhythm, it will not necessarily occur at the late time interval. Moreover, to further reduce the hazard function effects and to increase temporal uncertainty of when the target may occur, the cue randomly consisted of either four or five stimuli. Finally, the participants were not informed about the

temporal manipulation of the study. Nevertheless, the hazard function fits well with the pattern observed in the purely visual task (VV) whereby response times decreased orderly from early, sync and then late conditions. The expected pattern of results in terms of a model of entrainment would be that in sync targets would be faster than both early and late targets. In the two cross modal tasks there was no difference between late and in sync targets which may suggest both hazard function and entrainment effects influenced RTs. That is, in the early condition both entrainment and hazard effects predict slower RTs compared to the in sync condition. However, when the target is late, the hazard function predicts faster RTs compared to in sync targets, whilst an entrainment model would predict slower RTs. It is therefore possible that both entrainment and hazard function effects were present in this study. Future research may wish to use target discrimination tasks or detection of targets at perceptual threshold to further isolate entrainment effects from hazard functions.

Both crossmodal tasks showed similar effects of rhythmic cueing with a facilitation of response times for targets coinciding with the rhythm compared to when they appeared early. Importantly, this shows that stimulus driven rhythmic cueing is not limited to within a specific modality but effects can span across modalities. This is in line with a recent study by Miller et al. (2013) who found saccades to a visual target were faster when the target was preceded by a synchronous compared to an asynchronous auditory rhythm (see also Bolger et al., 2013 for similar results). The present study extends their findings by showing that audiovisual effects of rhythmic cueing are also found when the modalities are reversed, that is, a visual rhythm entrains auditory targets. This may suggest for a common mechanisms of temporal expectancy created by rhythms which is not modality specific. In line with this, Besle et al. (2011) observed large scaled entrainment of brain areas using intracranial electrocortical recordings in patients with epilepsy. They specifically found that the entrainment of visual stimuli was not confined to the primary visual areas but was observed over a larger brain area. That is, they observed effects in line with a centralized rather than purely modality specific entrainment mechanism.

The one spurious result in the present study was the lack of an temporal expectancy effect in the unimodal auditory task. Auditory entrainment of rhythms has shown to affect target discrimination of tones (e.g., Jones et al., 2002), as well as response times (Sanabria et al., 2011). In contrast, finding an effect of spatial attention in an auditory detection task, as was observed here, has proven more difficult with many studies reporting a null result (e.g., Posner, 1978; Scharf et al., 1987; Buchtel and Butter, 1988; Hugdahl and Nordby, 1994; although see Spence and Driver, 1994 for a positive effect of auditory spatial attention). Whether the introduction of an auditory spatial task diminished any auditory temporal effects remains unclear, however it seems unlikely as the auditory rhythm in the crossmodal task led to temporal expectancy effects of visual stimuli.

The study aimed to investigate how two processes which have independently shown to influence perception and modulate behavior interacted. The results showed that in no task was there

an interaction between spatial attention and rhythmic cueing ($p = 0.37$, $\eta_p^2 = 0.02$). In other words, any effects of rhythmic cueing were similar regardless if the target appeared at the spatially attended or unattended location. Whilst the results here show both unimodal and crossmodal effects of spatial attention, and unimodal and crossmodal effects of rhythmic cueing, spatial attention and temporal expectancy themselves did not interact, neither at a unimodal nor crossmodal level. This suggests temporal and spatial processes can operate independently in driving behavior, at least as measured with response times. Doherty et al. (2005) found similar independent effects of spatial attention and rhythmic cueing on response times, even though their participants were instructed to use the rhythmic structure to predict an upcoming target and thus introducing a voluntary aspect of rhythmic cueing. In contrast, Rohenkohl et al. (2014) recently showed temporal expectation improved perception when the target appeared at a spatially attended location. However, at unattended locations, temporal expectancy did not affect target processing. Rohenkohl et al. did not use rhythmic cueing but the temporal expectation was top-down. Moreover, they investigated perceptual sensitivity rather than response times which may also account for differences. Cravo et al. (2013) measured response times and perceptual accuracy and found both measured to be improved in a rhythmic compared to arrhythmic condition, but, the two measures showed independent effects. There is thus evidence to suggest perceptual modulation following rhythmic stimuli may be different to response time effects. Future research may wish to explore whether perceptual sensitivity effects of automatic entrainment are also independent from spatial attention effects in unimodal and crossmodal conditions. Moreover, the current study used a target detection task which was relatively easy. Within spatial attention research, endogenous and exogenous effects are typically independent when task demands are low. However, when the attentional and cognitive load increases, the two processes have shown to interact when competing for shared resources (Berger et al., 2005). It is conceivable that top-down spatial attention and bottom-up temporal expectancy effects show a similar pattern. In other words, future research could increase the difficulty of the task to investigate whether endogenous spatial attention and stimulus-driven temporal expectancies are independent even when demands on attentional resources are high.

The automatic effect of presenting rhythmic stimuli demonstrated in the present study is partly in line with research on neural oscillations which have seen a recent increase in popularity in the last decade. Evidence is mounting that, not only does our brain self-generate rhythmic oscillations which drives perception and action (e.g., Buzsaki, 2006; Thut and Miniussi, 2009), but these neural oscillations can also be re-set and driven by rhythms and events in our environment (Lakatos et al., 2008; Arnal and Giraud, 2012). Investigating the function and underlying mechanisms of entrainment will not only further our understanding of what drives our behavior and influences our perception, but recent findings have suggested that certain psychiatric and developmental disorders show abnormal neural oscillation patterns (see Calderone et al., 2014 for a recent review).

REFERENCES

Arnal, L. H., and Giraud, A. L. (2012). Cortical oscillations and sensory predictions. *Trends Cogn. Sci.* 16, 390–398. doi: 10.1016/j.tics.2012.05.003

Berger, A., Henik, A., and Rafal, R. (2005). Competition between endogenous and exogenous orienting of visual attention. *J. Exp. Psychol. Gen.* 134, 207–221. doi: 10.1037/0096-3445.134.2.207

Besle, J., Schevon, C. A., Mehta, A. D., Lakatos, P., Goodman, R. R., McKhann, G. M., et al. (2011). Tuning of the human neocortex to the temporal dynamics of attended events. *J. Neurosci.* 31, 3176–3185. doi: 10.1523/JNEUROSCI.4518-10.2011

Bolger, D., Trost, W., and Schön, D. (2013). Rhythm implicitly affects temporal orienting of attention across modalities. *Acta Psychol. (Amst)* 142, 238–244. doi: 10.1016/j.actpsy.2012.11.012

Breska, A., and Deouell, L. Y. (2014). Automatic bias of temporal expectations following temporally regular input independently of high-level temporal expectation. *J. Cogn. Neurosci.* 26, 1555–1571. doi: 10.1162/jocn_a_00564

Buchtel, H. A., and Butter, C. M. (1988). Spatial attention shifts: implications for the role of polysensory mechanisms. *Neuropsychologia* 26, 499–509. doi: 10.1016/0028-3932(88)90107-8

Busch, N. A., Dubois, J., and VanRullen, R. (2009). The phase of ongoing EEG oscillations predicts visual perception. *J. Neurosci.* 29, 7869–7876. doi: 10.1523/JNEUROSCI.0113-09.2009

Buzsaki, G. (2006). *Rhythms of the Brain.* Oxford: Oxford University Press.

Calderone, D. J., Lakatos, P., Butler, P. D., and Castellanos, F. X. (2014). Entrainment of neural oscillations as a modifiable substrate of attention. *Trends Cogn. Sci.* 18, 300–309. doi: 10.1016/j.tics.2014.02.005

Carrasco, M. (2014). "Spatial covert attention: perceptual modulation," in *The Oxford Handbook of Attention*, eds K. Nobre and S. Kastner (Oxford: Oxford University Press), 183–230.

Correa, A., Lupianez, J., and Tudela, P. (2005). Attentional preparation based on temporal expectancy modulates processing at the perceptual level. *Psychon. Bull. Rev.* 12, 328–334. doi: 10.3758/bf03196380

Coull, J. T., and Nobre, A. C. (1998). Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *J. Neurosci.* 18, 7426–7435.

Cravo, A. M., Rohenkohl, G., Wyart, V., and Nobre, A. C. (2013). Temporal expectation enhances contrast sensitivity by phase entrainment of low-frequency oscillations in visual cortex. *J. Neurosci.* 33, 4002–4010. doi: 10.1523/jneurosci.4675-12.2013

Davranche, K., Nazarian, B., Vidal, F., and Coull, J. (2011). Orienting attention in time activates left intraparietal sulcus for both perceptual and motor task goals. *J. Cogn. Neurosci.* 23, 3318–3330. doi: 10.1162/jocn_a_00030

Doherty, J. R., Rao, A., Mesulam, M. M., and Nobre, A. C. (2005). Synergistic effect of combined temporal and spatial expectations on visual attention. *J. Neurosci.* 25, 8259–8266. doi: 10.1523/jneurosci.1821-05.2005

Hugdahl, K., and Nordby, H. (1994). Electrophysiological correlates to cued attentional shifts in the visual and auditory modalities. *Behav. Neural Biol.* 62, 21–32. doi: 10.1016/s0163-1047(05)80055-x

Jaramillo, S., and Zador, A. M. (2011). The auditory cortex mediates the perceptual effects of acoustic temporal expectation. *Nat. Neurosci.* 14, 246–251. doi: 10.1038/nn.2688

Jones, M. R. (2010). "Attending to sound patterns and the role of entrainment," in *Attention and Time*, eds A. C. Nobre and J. T. Coull (Oxford: Oxford University Press), 137–330.

Jones, M. R., Johnston, H. M., and Puente, J. (2006). Effects of auditory structure on anticipatory and reactive attending. *Cogn. Psychol.* 53, 59–96. doi: 10.1016/j.cogpsych.2006.01.003

Jones, M. R., Moynihan, H., MacKenzie, N., and Puente, J. (2002). Temporal aspects of stimulusdriven attending in dynamic arrays. *Psychol. Sci.* 13, 313–319. doi: 10.1111/1467-9280.00458

Kayser, C., Petkov, C. I., and Logothetis, N. K. (2008). Visual modulation of neurons in auditory cortex. *Cereb. Cortex* 18, 1560–1574. doi: 10.1093/cercor/bhm187

Lakatos, P., Chen, C. M., O'Connell, M. N., Mills, A., and Schroeder, C. E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron* 53, 279–292. doi: 10.1016/j.neuron.2006.12.011

Lakatos, P., Karmos, G., Mehta, A., Ulbert, I., and Schroeder, C. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320, 110–113. doi: 10.1126/science.1154735

Lakatos, P., Musacchia, G., O'Connel, M. N., Falchier, A. Y., Javitt, D. C., and Schroeder, C. E. (2013). The spectrotemporal filter mechanism of auditory selective attention. *Neuron* 77, 750–761. doi: 10.1016/j.neuron.2012.11.034

Lampar, A., and Lange, K. (2011). Effects of temporal trial-by-trial cuing on early and late stages of auditory processing: evidence from event-related potentials. *Atten. Percept. Psychophys.* 73, 1916–1933. doi: 10.3758/s13414-011-0149-z

Lange, K., Krämer, U. M., and Röder, B. (2006). Attending points in time and space. *Exp. Brain Res.* 173, 130–140. doi: 10.1007/s00221-006-0372-3

Lange, K., and Röder, B. (2006). Orienting attention to points in time improves stimulus processing both within and across modalities. *J. Cogn. Neurosci.* 18, 715–729. doi: 10.1162/jocn.2006.18.5.715

Lawrence, M. A., and Klein, R. M. (2013). Isolating exogenous and endogenous modes of temporal attention. *J. Exp. Psychol. Gen.* 142, 560–572. doi: 10.1037/a0029023

Lima, B., Singer, W., and Neuenschwander, S. (2011). Gamma responses correlate with temporal expectation in monkey primary visual cortex. *J. Neurosci.* 31, 15919–15931. doi: 10.1523/JNEUROSCI.0957-11.2011

Luce, R. D. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization.* Oxford: Oxford University Press.

Mangun, G. R., and Hillyard, S. A. (1990). Allocation of visual attention to spatial locations: tradeoff functions for event-related brain potentials and detection performance. *Percept. Psychophys.* 47, 532–550. doi: 10.3758/bf03203106

Mathewson, K. E., Fabiani, M., Gratton, G., Beck, D. M., and Lleras, A. (2010). Rescuing stimuli from invisibility: inducing a momentary release from visual masking with pre-target entrainment. *Cognition* 115, 186–191. doi: 10.1016/j.cognition.2009.11.010

Miller, J. E., Carlson, L. A., and McAuley, J. D. (2013). When what you hear influences when you see: listening to an auditory rhythm influences the temporal allocation of visual attention. *Psychol. Sci.* 24, 11–18. doi: 10.1177/0956797612446707

Naccache, L., Blandin, E., and Dehaene, S. (2002). Unconscious masked priming depends on temporal attention. *Psychol. Sci.* 13, 416–424. doi: 10.1111/1467-9280.00474

Nobre, A. C. (2010). "How can temporal expectations bias perception and action," in *Attention and Time*, eds A. C. Nobre and J. T. Coull (Oxford, UK: Oxford University Press), 137–330.

Nobre, A. C., and Rohenkohl, G. (2014). "Time for the fourth dimension in attention," in *The Oxford Handbook of Attention*, eds K. Nobre and S. Kastner (Oxford: Oxford University Press), 676–724.

Nobre, A. C., Rohenkohl, G., and Stokes, M. G. (2011). "Nervous anticipation -Top down biasing across space and time," in *Cognitive Neuroscience of Attention*. 2nd Edn. ed M. I. Posner (New York: Guilford Publications), 159–186.

Posner, M. I. (1978). *Chronometric Explorations of Mind.* Hillsdale, NJ: Erlbaum.

Posner, M. I. (1980). Orienting of attention. *Q. J. Exp. Psychol.* 32, 3–25. doi: 10.1080/00335558008248231

Posner, M. I., Snyder, C. R., and Davidson, B. J. (1980). Attention and the detection of signals. *J. Exp. Psychol.* 109, 160–174. doi: 10.1037/0096-3445.109.2.160

Praamstra, P., Kourtis, D., Kwok, H. F., and Oostenveld, R. (2006). Neurophysiology of implicit timing in serial choice reaction time performance. *J. Neurosci.* 26, 5448–5455. doi: 10.1523/jneurosci.0440-06.2006

Raichle, M. E. (2010). Two views of brain function. *Trends Cogn. Sci.* 14, 180–190. doi: 10.1016/j.tics.2010.01.008

Rohenkohl, G., Coull, J. T., and Nobre, A. C. (2011). Behavioural dissociation between exogenous and endogenous temporal orienting of attention. *PLoS One* 6:e14620. doi: 10.1371/journal.pone.0014620

Rohenkohl, G., Cravo, A. M., Wyart, V., and Nobre, A. C. (2012). Temporal expectation improves the quality of sensory information. *J. Neurosci.* 32, 8424–8428. doi: 10.1523/jneurosci.0804-12.2012

Rohenkohl, G., Gould, I. C., Pessoa, J., and Nobre, A. C. (2014). Combining spatial and temporal expectations to improve visual perception. *J. Vis.* 14:8. doi: 10.1167/14.4.8

Sanabria, D., Capizzi, M., and Correa, A. (2011). Rhythms that speed you up. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 236–244. doi: 10.1037/a0019956

Santangelo, V., and Spence, C. (2008). Is the exogenous orienting of spatial attention truly automatic? Evidence from unimodal and multisensory studies. *Conscious. Cogn.* 17, 989–1015. doi: 10.1016/j.concog.2008.02.006

Scharf, B., Quigley, S., Aoki, C., Peachey, N., and Reeves, A. (1987). Focused auditory attention and frequency selectivity. *Percept. Psychophys.* 42, 215–223. doi: 10.3758/bf03203073

Spence, C. (2010). Crossmodal spatial attention. *Ann. N Y Acad. Sci.* 1191, 182–200. doi: 10.1111/j.1749-6632.2010.05440.x

Spence, C. J., and Driver, J. (1994). Covert spatial orienting in audition: exogenous and endogenous mechanisms. *J. Exp. Psychol. Hum. Percept. Perform.* 20, 555–574. doi: 10.1037//0096-1523.20.3.555

Spence, C., and Driver, J. (1996). Audiovisual links in endogenous covert spatial attention. *J. Exp. Psychol. Hum. Percept. Perform.* 22, 1005–1030. doi: 10.1037//0096-1523.22.4.1005

Spence, C., and Driver, J. (2004). *Crossmodal Space and Crossmodal Attention.* England: Oxford University Press.

Tang, X., Li, C., Li, Q., Gao, Y., Yang, W., Yang, J., et al. (2013). Modulation of auditory stimulus processing by visual spatial or temporal cue: an event-related potentials study. *Neurosci. Lett.* 553, 40–45. doi: 10.1016/j.neulet.2013.07.022

Thut, G., and Miniussi, C. (2009). New insights into rhythmic brain activity from TMS-EEG studies. *Trends Cogn. Sci.* 13, 182–189. doi: 10.1016/j.tics.2009.01.004

Vroomen, J., and de Gelder, B. (2000). Sound enhances visual perception: crossmodal effects of auditory organization of vision. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 1583–1590. doi: 10.1037//0096-1523.26.5.1583

Wright, R. D., and Ward, L. M. (1994). Shifts of visual attention: an historical and methodological overview. *Can. J. Exp. Psychol.* 48, 151–166. doi: 10.1037/1196-1961.48.2.151

Yeshurun, Y., and Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature* 396, 72–75. doi: 10.1038/23936

Zanto, T. P., Pan, P., Liu, H., Bollinger, J., Nobre, A. C., and Gazzaley, A. (2011). Age-related changes in orienting attention in time. *J. Neurosci.* 31, 12461–12470. doi: 10.1523/jneurosci.1149-11.2011