Long-Term Memory Biases Auditory Spatial Attention

3 authors:

Jacqueline Zimmermann
University of Toronto
3 PUBLICATIONS   **42** CITATIONS

SEE PROFILE

Morris Moscovitch
University of Toronto
433 PUBLICATIONS   **35,909** CITATIONS

SEE PROFILE

Claude Alain
University of Toronto
244 PUBLICATIONS   **11,186** CITATIONS

SEE PROFILE

Journal of Experimental Psychology: Learning, Memory, and Cognition

Long-Term Memory Biases Auditory Spatial Attention

Jacqueline F. Zimmermann, Morris Moscovitch, and Claude Alain

Long-Term Memory Biases Auditory Spatial Attention

Jacqueline F. Zimmermann, Morris Moscovitch, and Claude Alain
University of Toronto and Baycrest Centre for Geriatric Care, Toronto, Ontario, Canada

Long-term memory (LTM) has been shown to bias attention to a previously learned visual target location. Here, we examined whether memory-predicted spatial location can facilitate the detection of a faint pure tone target embedded in real world audio clips (e.g., soundtrack of a restaurant). During an initial familiarization task, participants heard audio clips, some of which included a lateralized target ($p = 50\%$). On each trial participants indicated whether the target was presented from the left, right, or was absent. Following a 1 hr retention interval, participants were presented with the same audio clips, which now all included a target. In Experiment 1, participants showed memory-based gains in response time and d'. Experiment 2 showed that temporal expectations modulate attention, with greater memory-guided attention effects on performance when temporal context was reinstated from learning (i.e., when timing of the target within audio clips was not changed from initially learned timing). Experiment 3 showed that while conscious recall of target locations was modulated by exposure to target-context associations during learning (i.e., better recall with higher number of learning blocks), the influence of LTM associations on spatial attention was not reduced (i.e., number of learning blocks did not affect memory-guided attention). Both Experiments 2 and 3 showed gains in performance related to target-context associations, even for associations that were not explicitly remembered. Together, these findings indicate that memory for audio clips is acquired quickly and is surprisingly robust; both implicit and explicit LTM for the location of a faint target tone modulated auditory spatial attention.

Keywords: long-term memory, spatial cognition, auditory attention, target detection and localization

Supplemental materials: http://dx.doi.org/10.1037/xlm0000398.supp

In psychology and cognitive neuroscience, attention and memory are two important phenomena, typically studied in parallel. Because the two can influence each other, it is important, also, to understand how memory and attention interact with one another. In that regard, much of the research and related theoretical frameworks focus on how attention influences what is effectively encoded, stored, and retrieved from memory (Ciaramelli, Grady, Levine, Ween, & Moscovitch, 2010; Dulas & Duarte, 2013; Fernandes & Moscovitch, 2000; Silva, Groeger, & Bradshaw, 2006). Comparatively few studies have been dedicated to the reverse relationship between the two—that is, how memory influences attention; those that do, have focused exclusively on the visual modality. Related auditory research has been limited to studying the effects of information in short-term memory (STM), in particular how spatial and temporal pattern expectations guide attention (Rimmele, Jolsvai, & Sussman, 2011; Sanders & Astheimer, 2008). Yet, in our daily lives, we often encounter situations where the voice of a friend (i.e., a familiar sound object in long-term memory [LTM]) guides our search in an acoustic environment where we have encountered it before (among the sounds in a coffee shop). The current study and paradigm provide evidence generalizable to these situations, illustrating how efficient orienting behavior in these environments is modulated by auditory scenes strongly encoded in memory.

Visual Memory Effects on Perception and Attentional Orientation

Using a visual search task, Chun and Jiang (1998) were the first, to our knowledge, to show that LTM can facilitate visuospatial attention. When the spatial arrangement of distracter stimuli was repeated from a previous trial, participants were faster at detecting a target stimulus within the array, even though they were not aware of its being repeated, suggesting that implicit memory acquired earlier in the experiment influenced visuospatial attention. This orienting effect toward remembered target locations during visual search has been labeled "contextual cueing." Since then, numerous studies have produced evidence for both improved behavioral performance (e.g., Jiang & Leung, 2005) as well as modulation of brain activity associated with memory-guided visual attention in stimulus arrays as well as ecologically relevant scenes (Chaumon, Hasboun, Baulac, Adam, & Tallon-Baudry, 2009; Summerfield, Lepsien, Gitelman, Mesulam, & Nobre, 2006). For example, a

group of researchers showed that participants were able to memorize the location of targets within photographs (e.g., a picture of a building), which increased perceptual sensitivity to remembered locations up to 24 hours later, as evidenced by both faster response times (reaction time [RT]) and enhanced neural activity for targets presented in learned spatial contexts (Patai, Doallo, & Nobre, 2012; Summerfield et al., 2006; Summerfield, Rao, Garside, & Nobre, 2011). These studies provided converging evidence that LTM biases visuospatial attention, which in turn facilitates target detection.

Auditory Memory Effects on Attention

Cherry (1953) first produced evidence that a highly familiar auditory stimulus can bias attention by showing that individuals attend to their own name within a stream of irrelevant or unattended information, a phenomenon now referred to as the "cocktail party effect." The long-standing interpretation of this phenomenon has been one of LTM effects on attention; highly meaningful auditory objects that are stored in LTM have the capability of capturing and holding attention (Treisman, 1964). Recently, a number of studies have challenged this interpretation (Conway, Cowan, & Bunting, 2001; Wood & Cowan, 1995), showing that only a small portion of participants actually report hearing their own name within the stream of unattended speech in dichotic listening tasks. Moreover, those participants who do show the cocktail party effect are those who have difficulty blocking out the irrelevant (name) stream, raising the possibility that the effect could be explained simply in terms of lapses of attention (Conway et al., 2001).

The influence of auditory LTM on the deployment of attention remains equivocal. Even if classic cocktail party phenomena could be explained as a result of LTM activating attention, the degree to which self-relevance and bottom-up attentional capture contributes to the effect is debated. The majority of studies where LTM may guide auditory attention use highly familiar stimuli, such as music (Bey & McAdams, 2002; Dowling, Lung, & Herrbold, 1987), or the participant's own name (Nakane, Miyakoshi, Nakai, & Naganawa, 2016; Cherry, 1953), that contain an element of self-relevance in terms of repeated exposure and emotional associations throughout the life span. In such cases, automatic capture mechanisms may facilitate attention, rather than higher-order top-down processing by LTM (Röer, Bell, & Buchner, 2013; Roye, Schröger, Jacobsen, & Gruber, 2010). In addition, the link between auditory memory and attention has not been tested experimentally in these studies, but rather inferred from dichotic listening results. In the current study, we investigated memory-guided attention where LTM was acquired in a controlled lab setting such that memory acquisition was experimentally manipulated. To avoid confounds imposed by self-relevance, the acquired LTM was not personally significant.

Although much is known about memory processes in visual attention, far less is known about audition. There is increasing evidence that memory and attention processes differ substantially for auditory and visual stimuli (Bigelow & Poremba, 2012; Berry, Li, Lin, & Lustig, 2014; Günther et al., 2014; Snyder & Gregg, 2011). Qualitative differences between visual and auditory domains have been identified at sensory (Zimmermann, Moscovitch, & Alain, 2016), short-term (Bigelow & Poremba, 2012), and

long-term memory (Cohen, Horowitz, & Wolfe, 2009; Snyder & Gregg, 2011) levels. The basis for these differences lies in the dynamic nature and longer time-course of auditory processing compared to visual processing (Zimmermann et al., 2016). Separate processes have also been identified for auditory and visual attention (Berry et al., 2014; Günther et al., 2014). For example, allocation of auditory attention is generally slower than visual attention (Günther et al., 2014; Salmi, Rinne, Degerman, & Alho, 2007), but the auditory modality is less affected by declines in sustained attention over time (Berry et al., 2014). Spatial selectivity is stronger in vision, while temporal selectivity is stronger in audition (Rimmele et al., 2011). Moreover, the interaction between memory and attention may differ for auditory stimuli compared to visual stimuli because auditory memory is largely verbal in nature and we have less experience with remembering nonverbal auditory stimuli (Snyder & Gregg, 2011). In addition, memory-guided auditory search may be much more demanding and rely on different mechanisms from visual search because auditory stimuli typically contain fewer identifiable details and features that could be used to guide attention.

These differences highlight the importance of investigating the interplay between attention and memory in audition because the findings and conclusions from studies using visual material may not apply to the auditory domain. The current study aimed to fill this gap and determine whether LTM biases attention to auditory targets presented in learned spatial locations. Our findings reveal memory-guided attention as a general psychological construct that is not specific to one sensory system.

In order to test the effects of auditory memory on attention, we first needed to determine whether complex auditory scenes could be bound with simple targets, which was established through an initial learning task and confirmed with a memory test. Only then could we assess whether the presentation of an audio clip, previously associated with a specific target location, can subsequently bias attention and speed up the processing of the target at the expected location.

Explicit and Implicit Memory-Guided Attention

Our second aim was to investigate whether auditory memory that is not consciously accessible can also facilitate spatial attention. In visual search paradigms, both explicit (Summerfield et al., 2011; Patai et al., 2012) and implicit memory (Ciaramelli, Lin, & Moscovitch, 2009; Chun & Jiang, 2003) facilitates visual spatial attention. For example, Chun and Jiang (2003) showed that participants located targets more successfully when cued with previously encountered contextual displays even when they were not able to explicitly recall the target's general location (i.e., display quadrant) at above-chance levels. Notably, implicit memory modulated brain activity associated with the deployment of visual attention after delay periods of one week (Chun & Jiang, 2003). Prior work, however, has predominantly examined the effects of implicit and explicit memory separately, and only in vision. Moreover, implicit memory is predominantly studied using array-based contextual cueing stimuli, while explicit memory is studied using scene-based stimuli (Rosenbaum & Jiang, 2013). Thus, it is difficult to dissociate the influence of paradigm/stimuli and memory accessibility (implicit vs. explicit memory).

How might the effect of explicit and implicit memory on attention differ? In contrast to explicit memory, implicit memory-guided attention lacks the intentional component of top-down control (Rosenbaum & Jiang, 2013). Moreover, the implicit process is not guided by semantic knowledge in the same manner as explicit memory-guided attention (Brockmole & Henderson, 2006), which may be particularly detrimental in audition where semantic knowledge plays an important role. Semantic knowledge may be more important in guiding auditory processing because nonverbal auditory stimuli are generally more ambiguous than visual stimuli. For example, an audio clip of a noisy restaurant may be difficult to identify, store, and access because it can easily be confused with other audio clips from the same or different restaurants.

Since implicit and explicit memory are mediated, to some extent, by separate systems, it is likely that their effects on attention differ substantially. That is, different memory sources may contribute to differences in attentional guidance strategies (Patai et al., 2012). For example, Kunar, Flusberg, and Wolfe (2006) suggested that attentional guidance becomes much more effective when participants become aware of an association between a visual context and target location.

At the same time, it is likely that explicit and implicit memory-guided attention rely on similar brain mechanisms in the medial temporal lobe. For example, Summerfield et al. (2006) showed that the hippocampus was the only structure in which activity levels correlated with the magnitude of explicit memory-guided behavioral benefits. Similarly, hippocampus-dependent processing is observed in implicit contextual cueing paradigms.

Here, it is important to distinguish between two kinds of implicit memory-guided orienting studied in the visual modality: contextual cueing paradigms (i.e., in which participants perform a visual search to locate a target among distracters) and reinforcement memory (i.e., a type of reinforcement learning, where trials convey probabilistic stimulus-response relationships). Though the two tasks often appear the same to the participant, reinforcement tasks contain underlying probabilities associated with participant responses, whereas contextual cueing tasks contain probabilities associated with location. While contextual cueing is thought to involve hippocampus-dependent memory processing, attention guided by stimulus-response relationships appears striatum-dependent (Goldfarb, Chun, & Phelps, 2016). The two processes are representative of different scenarios encountered in the real world where LTM biases attention and increases perceptual sensitivity, either when we use knowledge about our visual environments to bias attention toward specific locations in space (e.g., using memory cues, such as the location of a garage wall or lamp post, when searching for our car in the parking lot), or when we use knowledge about causal relationships between environmental stimuli (e.g., green compared to red pedestrian crossing light usually means we do not have to be as attentive to approaching cars).

The current study finds support for explicit memory-guided attention as well as the contextual cueing form of implicit memory-guided attention in the auditory domain. We showed that memory for the location (Experiment 1) and timing (Experiment 2) of an auditory target within an auditory context modulates selective attention, even with little opportunity to consolidate the memory in the first place (Experiment 3). Memory cue-based gains in performance are large even when cues activate only implicit memory, and are comparable to the attention gains based on explicitly remembered trials.

Experiment 1

Experiment 1 examined whether LTM facilitates the detection of a lateralized target tone embedded in an audio clip. First, in the encoding phase participants were presented several times with audio clips that comprised a lateralized target near thresholds (memory condition). They were also presented with audio clips that did not include a target (neutral condition). After a 1-hr delay, the same audio clips were presented, but now all of them comprised a lateralized target. We hypothesized that participants would be more accurate and faster on those trials where the audio clip and the target were combined during the learning phase (i.e., memory condition), compared to trials where no memory association exists (i.e., neutral condition). Explicit and implicit memory effects were examined by comparing performance on those trials where participants did and did not explicitly remember the location of the target.

Method

Participants. Sixteen healthy young adults ($M = 23.5$ years; range 19–33 years; six males), with normal hearing and normal or corrected-to-normal vision were recruited for the experiment. Hearing was assessed using pure tone thresholds for octave frequencies ranging from 250 to 8,000 Hz, with the criterion for normal hearing being thresholds lower than or equal to 25 decibels (dB) of sound pressure level (SPL) and less than or equal to a 15 dB SPL difference between the two ears at each octave frequency. In addition, all participants were right-handed, fluent in English to ensure understanding of the experimental process, and had no history of psychiatric, neurological, or other major illness. Participants were recruited from the Rotman Research Institute participant database, and received monetary compensation for their participation. All participants provided informed consent prior to taking part in the experiment, according to the Research Ethics Board at the Baycrest Center.

Stimuli. Eighty audio clips were used in the experiment, all retrieved from "http://www.freesounds.org/." We used a relatively small stimulus set due to known limitations of memory for nonverbal auditory stimuli, and also because of the constraint for audio clips to be recognizably distinct from one another (Cohen, Evans, Horowitz, & Wolfe, 2011; Snyder & Gregg, 2011). The clips were chosen to maintain considerable semantic relevance (e.g., soundtrack of dogs barking, amusement park) in order to increase the likelihood that an appropriate association could be formed and labeled in LTM (See supplemental material for the list of audio clips). The same 80 audio clips were used in the learning task, explicit spatial memory recall task, and memory-guided attention task. All audio clips were prepared using Adobe Audition 1.5 (Adobe Systems Inc., San Jose, CA); clips were cut to a length of 2,500 ms with a 100 ms rise and fall time, down-sampled to a standard sampling rate of 44,100 Hz, and batch normalized to a total root mean square (RMS) of −23 dB SPL (expressed as the difference between the volume at each ms interval and peak volume).

The auditory target was a 500 Hz pure tone with duration of 200 ms, with 20 ms rise and fall time. Its volume was adjusted for each participant to allow 80% detectability.

All stimuli were presented through insert earphones EAR-TONE 3a), at a listening volume of approximately 60 dB SPL, with some sounds peaking at about 80 dB SPL, as measured using a Larson-Davis (Depew, NY) System 824 Sound Level Meter. Acoustic stimuli and visual cues were presented using Presentation software (version 13, Neurobehavioral Systems, Albany, CA).

Experimental procedure. Prior to the experimental tasks, an audiometric test was administered to confirm normal hearing thresholds. Participants completed a total of four experimental tasks, which were performed in a sound attenuating testing booth. Participants were seated in a comfortable chair approximately 1 m from a computer screen.

Determining individual signal-to-noise (SNR) thresholds. The purpose of the first task was to establish individual thresholds that would be used in the experiment, which was important to ensure that all participants were engaging in effortful listening. A not-so-easily detectible target allowed us to better examine modulation of perceptual sensitivity by LTM (Patai et al., 2012). For each participant, we identified the SNR needed to detect the pure tone target embedded within a subset of four audio clips with a ~80% detection accuracy level. The audio clips used during the calibration task were excluded from the other phases of the experiment in order to prevent potential memory contamination effects. Each audio clip was presented twice during each trial for a length of 500 ms, with a 500-ms interval between the first and second presentation. The target (500 ms with 50-ms rise and fall time) was embedded in the first or second presentation at random. Using a two-alternative forced choice procedure, participants indicated which stimulus held the target tone by pressing *1* or *2* on a keyboard. Participants were given 2,000 ms to respond following the offset of the second repetition of each clip, after which visual feedback was presented. The 79% detectability SNR was estimated using a three-down one-up algorithm (Levitt, 1971). Beginning with a 0.7 SNR, following three consecutive correct responses, the SNR was decreased by 5% of its original volume, while one incorrect response increased SNR by a factor of 5. The threshold was calculated by taking an average of the last 8 of 12 reversals.

The volume of the pure tone target was raised from the originally calculated 79% detectability SNR by 10% in the following tasks, since many of the soundtracks used during the learning task and memory-guided attention task masked the embedded target when the original SNR was used in a pilot study. The chosen volume allowed for approximately 80–90% correct detection of the target within audio clips during the main experiment.

Learning task. A total of 80 audio clips, presented binaurally, were divided into memory and neutral cue trials (40 each). In the memory trials, a pure tone target was paired with the audio clip, presented in the left (20 trials) or right (20 trials) ear at random. In the neutral trials, no target was presented. Each participant was presented with the same 80 trials over 8 learning blocks (640 trials in total) to promote a strong association between audio clips and location of the target when present. In addition, to strengthen the target-audio clip associations, the trials for which participants made an incorrect response were repeated until a correct response was made within each block. The order of trials was random within each block. Within the memory trials, the target tone was played

2,000 ms after sound onset, and lasted for 200 ms. Participants were instructed to listen for and memorize the location of the target within each audio clip, and pressed the left, right, or down arrow key on a keyboard when the target was played from the left side, right side, or if no target was present, respectively. Participants were given 2,000 ms to respond following the offset of the audio clip, and subsequently received visual feedback for 500 ms (indicating hit, miss or incorrect response). Participants were asked to respond as quickly and as accurately as possible.

Explicit spatial memory recall task. Immediately following the learning task, a cued recall memory task was administered to determine whether participants formed explicit associations between audio clips and target location. Participants were presented with the same audio clips as in the learning phase, but without targets. For each audio clip, participants indicated whether the target had been presented from the left or right ear or whether no target had been present. Participants were given as much time as needed to make their response. Subsequently, confidence in responses was rated using a 3-step scale keypress response (*1* indicating not confident at all, *2* indicating fairly confident, *3* indicating very confident). There was a 500-ms time window between the button press and the next trial.

Memory-guided attention task. Following an hour retention interval, participants were presented with the same 80 audio clips from the learning task, each repeated twice in order to ensure that they had sufficient time to access learned target-context associations (i.e., to ensure sufficient cueing). Figure 1 provides a trial overview. A 1,000-ms interstimulus interval separated the first (S1) and second (S2) presentation of the audio clip. While S1 did not include a target tone and only served as a retrieval cue to guide attention toward remembered target location, S2 always comprised an embedded pure tone target.

For memory trials, the target was always presented at the learned location. For neutral trials, which did not include a pure tone target during the learning phase, a target was presented from either the left or right side at random. The audio clips in the neutral condition were equally familiar to those in the memory condition, but were not associated with a specific target location. Memory and neutral trials were randomly intermixed throughout the memory-guided attention task.

An initial pilot test indicated that lengthy exposure to cues is needed to activate auditory memory for target location—context associations. Therefore, we chose to insert the target toward the end of the audio clips (2,000 ms after onset of the audio clip) in order to increase the informativeness of the memory cue.

On each trial, participants pressed a button as quickly and as accurately as possible when they heard the pure tone target. A 2,500-ms time window was given for responses starting from the onset of the target (i.e., 2,000 ms after offset of the audio clip), and 1,000 ms preceded the onset of the next trial. Participants completed two blocks of trials to examine how memory-guided modulation of attention changes over time.

Analyses

Response time (RT) and perceptual sensitivity were analyzed using a series of repeated measures (RM) ANOVAs for the learning and memory-guided attention tasks. RT was measured in milliseconds (ms) from target onset to keypress response. Percep-
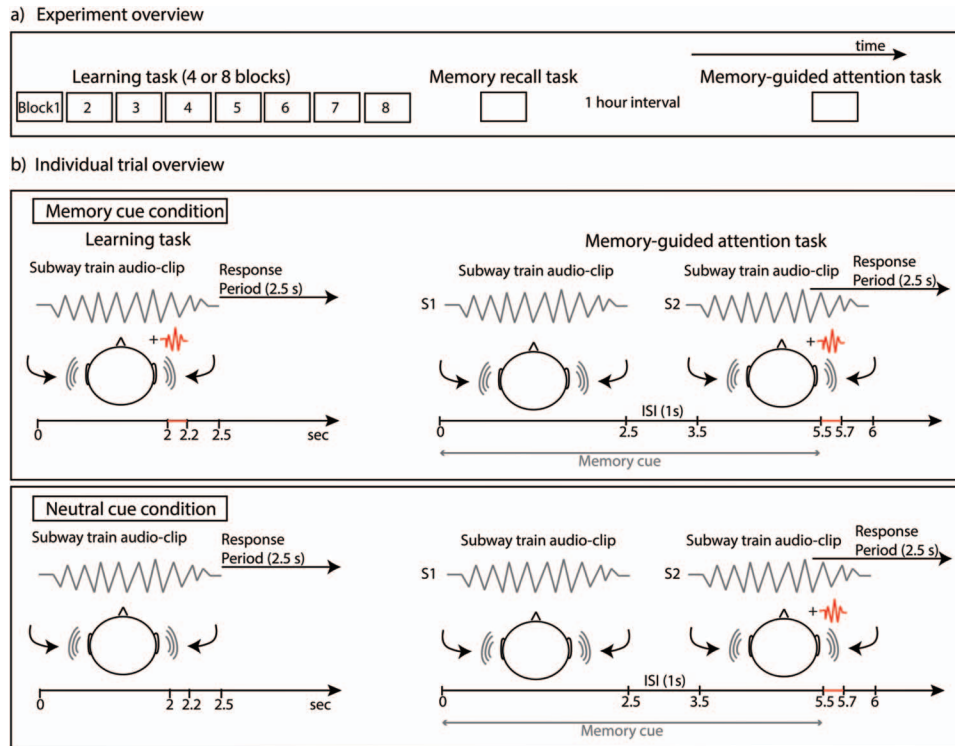
Figure 1. (a) Overview of the three main experimental tasks: learning task (purpose: to create target-context associations), memory recall task (purpose: to determine whether and for which audio clips the target location was consciously accessible), and memory-guided attention task (purpose: to examine whether memory formed in the learning task guides attention). (b) Overview of one trial for each experimental condition, for the learning task and memory-guided attention task. In the memory cue condition, participants learned the location of the target, located in left or right hemispace, and the target was then presented at the learned location in the memory-guided attention task. RT and d' was compared to neutral cue condition trials, where participants did not learn the location of a target (i.e., no target-context association formed in the learning task), and the target was then presented at a novel location in the memory-guided attention task. Note: S1 and S2 in the memory-guided attention task represent the first and second repetition of the audio clip, separated by an 1,000 ms ISI. Only S2 contained an embedded pure tone target, 2,000 ms after S2 onset. See the online article for the color version of this figure.

tual sensitivity was quantified as the rate of correct detections minus false alarms [$d' = z$(hits) – z(false alarms)]. In the learning task, d' was calculated based on only first-presentation trials (the first presentation of each audio-clip). For the learning task, learning block (block 1–8) was input as a factor to measure learning effects. For the memory-guided attention task, trial type (memory vs. neutral trials) was input as a factor to measure memory-guided attention. To understand memory-guided attention, the primary analysis was performed using only those stimuli for which a target-context association was formed (i.e., correctly recalled trials). We also performed a second analysis using all trials. We excluded from the analysis incorrect trials, trials where RT was faster than 100 ms, as well as outliers (RT >2 *SD* from the mean).

Recall of target-context associations was compared to chance levels using a one-sample *t* test. Recall was analyzed for all trials (to examine memory for target presence and target location), as well as for target-present trials only (to examine memory for target location only). RMANOVAs were also used to compare recall of the location of the target tone across reported confidence levels.

In addition, to examine the relationship between memory and memory-guided attention, we performed a simple bivariate correlation between participants' accuracy in recalling target locations during the memory task and memory-guided attention (operationalized as standardized difference scores in RT to detect the target within memory and neutral trials).

Results and Comment

Creating associations between audio clips and the location of an auditory target. Figure 2 shows group mean RT and perceptual sensitivity ([*d*]) as a function of learning block. Participants improved in detecting and localizing auditory targets embedded in audio clips with increasing exposure. The RMANOVA on RT with learning block (block 1–8) and target (present, absent) factors yielded main effects of learning block, $F(7, 105) = 7.49$, $p < .001$, target, $F(1, 15) = 32.82$, $p < .001$ as well as a significant interaction, $F(7, 105) = 5.98$, $p < .001$. There was a steep decrease in RT from the first to the second learning block (differ-
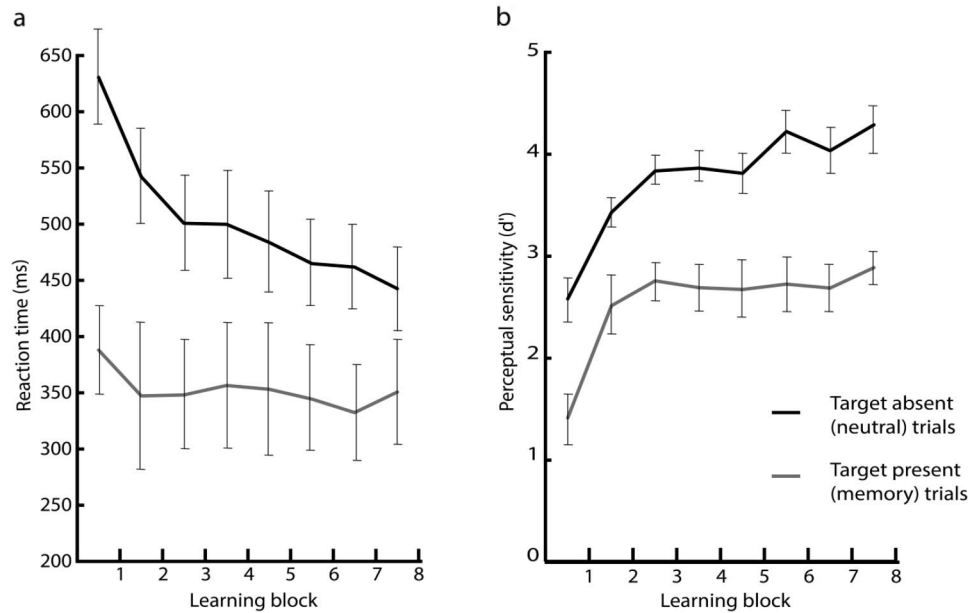
Figure 2. Formation of auditory target-context memory associations over learning blocks. Participants detected targets embedded within audio clips (a) more quickly and (b) with increasing perceptual sensitivity over learning blocks. Error bars represent standard error of the mean.

ence between block 1 and 2 = 64.58 ms, $d = 2.90$, 95% CI for d [1.87, 3.94]; mean difference between other consecutive blocks = 8.02 ms; $d = 0.04$, 95% CI for d [−0.68, 0.77]), indicating that auditory target-context associations are formed quickly and additional exposure may not strengthen these associations comparably. Perceptual sensitivity gains were observed over the course of the experiment, $F (7, 105) = 9.47$, $p < .001$; with the largest gains occurring over the first two blocks (Figure 2b).

The main effect of target presence revealed longer RTs for target-absent trials ($M = 503.31$ ms, $SEM = 38.32$) compared to target-present trials ($M = 352.55$ ms, $SEM = 47.86$). The interaction effect showed that participants learned more quickly that a target tone was present within audio clips than its location, that is, left or right ear (see Figure 2a). This was expected based on previous auditory memory studies, which suggest that binding auditory targets to specific spatial locations may be more difficult in comparison to similar visual tasks (Cohen et al., 2009).

We performed a separate analysis using only target-present trials to determine whether localization of the target improved over the course of the learning task. In terms of RT, the main effect of learning block did not reach significance due to the small sample size and the large variability in search times across participants, although the means did reflect a decrease in search time over the course of the task (block 1: $M = 388.03$ ms, $SE = 156.95$; block 4: $M = 356.53$ ms, $SE = 223.02$; block 8: $M = 350.67$ ms, $SE = 185.83$; Figure 2a—Target-present trials). d', however, improved significantly over learning blocks (block 1: $d' = 1.40$, $SE = 0.24$; block 4: $d' = 2.68$, $SE = 0.23$; block 8: $d' = 2.872$, $SE = .18$), $F (7, 105) = 5.94$, $p < .001$. In fact, the gain in accuracy over time was larger for target-present trials (17.19% increase in correct detection over learning blocks) than the gain calculated using all trials (target-present and target-absent trials; 13.52% increase),

indicating that individuals can form strong associations between audio clips and a specific auditory target location (Figure 2b).

Overall, performance gains to detect the auditory target were quite large, with gains in RT being greater than 100 ms over the course of the learning task, and correct target detection increasing from an 81% level within the first learning block to near perfect detection by the end of the learning task (95% for block 8). The performance improvement suggests that memory for auditory target-context associations is formed quickly and may improve perceptual sensitivity over time. We concluded that the learning task successfully facilitated the formation of associations between audio clips and target location, which could later be used to orient attention to the expected hemispace.

Explicit spatial recall task: Memory for auditory target-context associations. Participants correctly recalled 75% of previously learned auditory target-context associations ($M = 60.3/$80), significantly more than the proportion expected by chance ($M = 26.67/80$), $t (15) = 10.74$, $p < .001$, $d = 2.69$, 95% CI for d [1.66, 3.86]. They also correctly responded to a large proportion of trials ($M = 25/40$; 62.5%), even when target-absent (neutral) trials were excluded from the analysis, $t (15) = 4.93$, $p < .001$, $d = 1.23$, 95% CI for d [0.58, 1,93] (compared to chance), indicating that they not only remembered whether a pure tone target was paired with a particular audio clip, but also remembered its specific spatial location. The interaction between correct recall and confidence level suggests that participants were able to gauge whether they had correctly recalled the location of the target, $F (2, 30) = 14.41$, $p < .001$; there were more correct responses and fewer incorrect responses as confidence level increased.

Memory-guided auditory spatial attention. In line with the main hypothesis, we confirmed that auditory spatial attention can be modulated by LTM. The presentation of a memory cue (audio

clip) facilitated preparation of attention toward remembered target location. Thus, attention was allocated toward the location of the target before its onset, whereas in neutral trials, attention was likely divided between left and right hemispace (e.g., no memory-guided bias of attention).

RMANOVAs yielded a main effect of memory cue, with targets embedded within memory trials being detected faster, $F (1, 15) = 37.28$, $p < .001$, $d = 0.96$, 95% CI for d [0.19, 1.72], and with greater perceptual sensitivity, $F (1, 15) = 43.64$, $p < .001$; $d = 2.61$; 95% CI for d [1.63, 3.59], than target tones within neutral trials (see Table 1 for RT and d' data). Memory cues were also associated with 100-ms gains in RT and over 12% increase in correct detection accuracy. Our findings are consistent with those from the visual modality, but with memory-related gains greater than those observed in visual parallels (e.g., Patai et al., 2012; Stokes, Atherton, Patai, & Nobre, 2012).

Unexpectedly, there was also a main effect of target side, with higher d' for right ([d'] = 3.31, $SE = .14$) than left side targets ([d'] = 2.79, $SE = .21$), $F (1, 15) = 5.07$, $p < .05$, $d = 0.82$; 95% CI for d [0.07, 1.57]. We also performed an analysis using all trials (i.e., correct and incorrect at the memory task). The pattern of findings was essentially identical.[1]

To further assess memory-guided attention, we calculated a bivariate correlation between the number of correctly recalled target-context associations for each participant and the gain in RT based on memory cue (normalized difference calculation: Mneutral $- M$memory/Mneutral $+ M$memory). We hypothesized that higher recall should be associated with greater differences in RTs between memory and neutral trials. In other words, participants with a better memory for target-context associations should also demonstrate stronger memory-guided modulation of spatial attention. Results confirmed the hypothesis, as there was a positive relationship, $r = .57$, $p < .05$ between memory recall accuracy and cueing effects, indicating that auditory memory is closely related to spatial attention (Figure 3a).

The memory-guided attention task was repeated twice to increase the number of trials to assess memory-guided modulation effect over time. For the neutral trials, it was possible that participants created an association between the audio clip and the target, which could carry over to the second test. If this was the case, then the difference between memory and neutral trials should be smaller for the second block of trials than for the first block. We excluded one participant from this analysis due to an incomplete dataset ($N = 15$). RMANOVAs (conducted using memory-correct trials) revealed that the memory-guided attentional orientation effect persisted during the second block of trials, with faster RTs ($F [1, 14] = 13.32$, $p < .005$; $d = .8568$; 95% CI for d [0.08, 1.64]), and higher d' ($F [1, 14] = 14.37$, $p < .005$; $d = 1.33$; 95% CI for d [0.50, 2.16]) for memory than neutral trials. Moreover, the effect of memory cue on RT and d' did not differ significantly between the two blocks of trials.

In Experiment 1, we only examined the effect of memory cue on auditory spatial attention, yet temporal factors may have a critical influence on the memory-guided attention effect. Prior research has shown that participants are more accurate and faster in responding when they know in advance when the auditory target will occur, even if this knowledge is implicit (Rimmele et al., 2011; Sanders & Astheimer, 2008; Shen & Alain, 2011, 2012). Similarly, in Experiment 1 participants may have successfully learned when

the target would occur, as it was always presented at the same time within the audio clips.

Experiment 2

We conducted a second experiment to examine the effects of target-context binding by varying the time of target occurrence within audio clips during the memory-guided attention task. The main goals of Experiment 2 were to determine (a) whether temporal expectations guide attention and (b) whether the effects of memory on spatial attention persist even when temporal expectations are disrupted. Given that temporal expectations play an important role in auditory processing (Rimmele et al., 2011), we hypothesized that memory-guided attention effect on auditory spatial attention would be weaker when the target occurrence varies from the originally encoded temporal position. That is, memory-guided orienting, reflected by the difference in RT and d' to detect targets embedded within memory versus neutral trials, should be attenuated in Experiment 2.

The secondary goal of Experiment 2 was to examine whether implicit auditory memory for target locations will bias deployment of spatial attention in a similar fashion as explicit memory. The reported findings for Experiment 1 pertained only to the effects of explicit auditory memory on deployment of spatial attention. Due to strong recall across participants, and therefore a low number of memory-incorrect trials in either experiment alone[2], we combined data from Experiment 1 and 2 to analyze attention driven by associations that are not consciously accessible.

Based on prior research reporting implicit memory-guided effects on visuospatial attention (Chun & Jiang, 2003; Johnson, Woodman, Braun, & Luck, 2007), we predicted that trials in which the location of the auditory target was not consciously accessible may also bias auditory spatial attention. Specifically, we expected that target-context associations for memory-incorrect trials would show a similar, albeit attenuated, advantage in RT and d' to detect targets embedded within memory cue clips that were not explicitly recalled as that for clips where target location was explicitly remembered.

Method

Sixteen young adults ($M = 24.3$ years; range 20–31 years; six males) recruited from the Rotman Research Institute participated in Experiment 2 for monetary compensation. None of the participants took part in Experiment 1. The materials and procedures were the same as those of Experiment 1, except that the target occurrence within audio clips was different from the temporal

---

[1] The effect of memory cue on RTs was significant, $F (1, 15) = 34.19$, $p < .001$, $d = 0.91$, 95% CI for d [0.1548, 1.6723], with a similar gain in RT as the memory-correct trial analysis for memory cue trials ($M = 520.14$ ms, $SE = 26.13$) compared to neutral trials ($M = 610.48$ ms, $SE = 23.10$). The effect of cue on d' was also significant, $F (1, 15) = 37.55$, $p < .001$, $d = 1.92$, 95% CI for d [1.05, 2.79], with a similar gain in d' to locate targets within memory trials when only memory-correct trials were analyzed.

[2] We used an exclusion criterion such that only data from participants who had more than six incorrect trials per condition were analyzed, resulting in a total of 11 participants across the two experiments being included in the analysis.

Table 1
Memory-Guided Deployment of Spatial Attention

Experiment	Included trials	Measure	Condition		
			Memory Cue	**Neutral Cue**	
Experiment 1 (eight learning blocks)	All trials	RT (ms)	520.14 ± 26.44	610.48 ± 22.99	*
		Accuracy (d')	4.15 ± .24	2.03 ± .16	*
	Memory-correct trials	RT (ms)	489.19 ± 25.02	583.04 ± 23.41	*
		Accuracy (d')	3.66 ± .24	2.10 ± .15	*
Experiment 2 (disrupted learned temporal associations)	All trials	RT (ms)	707.07 ± 57.44	783.69 ± 70.81	*
		Accuracy (d')	2.93 ± .17	2.86 ± .22	—
	Memory-correct trials	RT (ms)	682.89 ± 58.78	739.22 ± 67.10	—
		Accuracy (d')	3.90 ± .29	1.98 ± .15	*
	Memory-incorrect trials	RT (ms)	592.62 ± 47.83	671.80 ± 31.98	*
		Accuracy (d')	1.77 ± .27	1.28 ± .15	*
Experiment 3 (four learning blocks)	All trials	RT (ms)	619.15 ± 37.76	678.46 ± 34.50	*
		Accuracy (d')	3.27 ± .17	2.41 ± .14	*
	Memory-correct trials	RT (ms)	589.217 ± 37.062	649.509 ± 31.161	*
		Accuracy (d')	3.89 ± .30	2.58 ± .17	*
	Memory-incorrect trials	RT (ms)	698.76 ± 53.90	743.57 ± 54.48	—
		Accuracy (d')	3.36 ± .35	2.68 ± .46	—

Note. RT and perceptual sensitivity to locate targets preceded by memory and neutral cues are reported ($M \pm SEM$). Data from all trials, only those trials for which participants correctly recalled target location (i.e., memory cue is consciously accessible), and incorrectly recalled trials are reported. Note that memory-incorrect trials were combined from Experiments 1 and 2 to increase data set size. Significant differences between conditions are highlighted (*).

position that was originally encoded (association formed during learning). In the learning task, participants consistently heard the target tone 2,000 ms following the onset of the audio clip; in the memory-guided attention task, we randomized the target tone to onset 1,300, 1,800 or 2,300 ms following onset of the second audio clip presentation.

Results and Comment

As in Experiment 1, participants were more accurate and quicker to locate the target with increased practice, with 9% correct detection accuracy increases and 100-ms gains over the eight learning blocks; $F(7, 105) = 5.10$, $p < .01$ and $F(7, 105) = 5.03$, $p < .001$, respectively. 21% of variance in RTs was attributable to effects of learning block (block 1: RT $M = 545.95$ ms, $SE = 31.45$; block 4: $M = 471.25$ ms, $SE = 33.51$; block 8: $M = 446.53$ ms, $SE = 36.67$). Perceptual sensitivity for target-present trials only also increased significantly as a function of exposure (block 1: $d' = 1.27$, $SE = 0.26$; block 4: $d' = 1.91$, $SE = 0.23$; block 8: $d' = 2.07$, $SE = 0.35$), $F(7, 105) = 2.95$, $p < .01$, confirming that individuals are good at binding targets to specific spatial locations within audio clips.

Temporal expectations modulate attention. There was a main effect of target onset (i.e., regardless of cue condition), such that varying the target occurrence within audio clips from the originally learned temporal position induced uncertainty. Variable target onsets[3] generated longer RTs ($M = 745.38$, $SE = 45.38$) than fixed target onset[4] ($M = 565.31$, $SE = 19.05$), $F(1, 30) = 7.30$, $p < .05$, $d = 0.96$, 95% CI for d [0.19, 1.72]. This is consistent with previous research showing that knowing when a target would occur speeds up perceptual decisions (Rimmele et al., 2011; Sanders & Astheimer, 2008).

Memory-guided attention persists when temporal target-context associations are disrupted.[3] Experiment 2 provided further evidence that LTM biases auditory spatial attention. Al-

though the target occurrence was different in the memory-guided attention task than during the learning phase, the orienting effect persisted; we observed gains in RT and d' for memory trials compared to neutral trials, $F(1, 15) = 5.47$, $p < .05$, $d = 0.30$, 95% CI for d [−0.43, 1.02] and $F(1, 15) = 28.07$, $p < .001$, $d = 2.06$, 95% CI for d [1.17, 2.95], respectively (see Table 1 for RT and d' data). The results indicate that memory cueing is not limited to a specific time point within the audio clip and appears to generalize to other temporal positions.

A mixed ANOVA was conducted to test whether the benefit of memory cue was comparable in Experiment 1 (constant onset of target) and Experiment 2 (onset of the target is changed from learning). The effect of memory cue on RT was comparable in Experiment 1 (memory-based gain in RT: $M = 90.33$ ms, $SE = 15.34$) and Experiment 2 (memory-based gain in RT: $M = 76.62$ ms, $SE = 32.76$; Experiment x cue type, $F[1, 30] = 0.14$, $p = .71$). However, the analysis of d' revealed an interaction effect between experiment and cue type, $F(1, 30) = 9.62$, $p < .005$. Cue-related gains in d' were larger when temporal target-context associations were not disrupted.

The higher d' in Experiment 1 suggests that memory representations of target occurrence within the audio clip is important in guiding auditory attention. Auditory scene analysis appears more dependent on the precise temporal structure than visual processing (Bizley & Cohen, 2013). Auditory events are ephemeral, fleeting, and changing over time, hence they can only be revisited by orienting attention to auditory representations in short-term memories (Backer & Alain, 2012). This is markedly different from visual scene analysis, which often involves more static stimuli (e.g., photographs) that typically remain available for further reexamination.

---

[3] Data for this analysis was taken from Experiment 2.
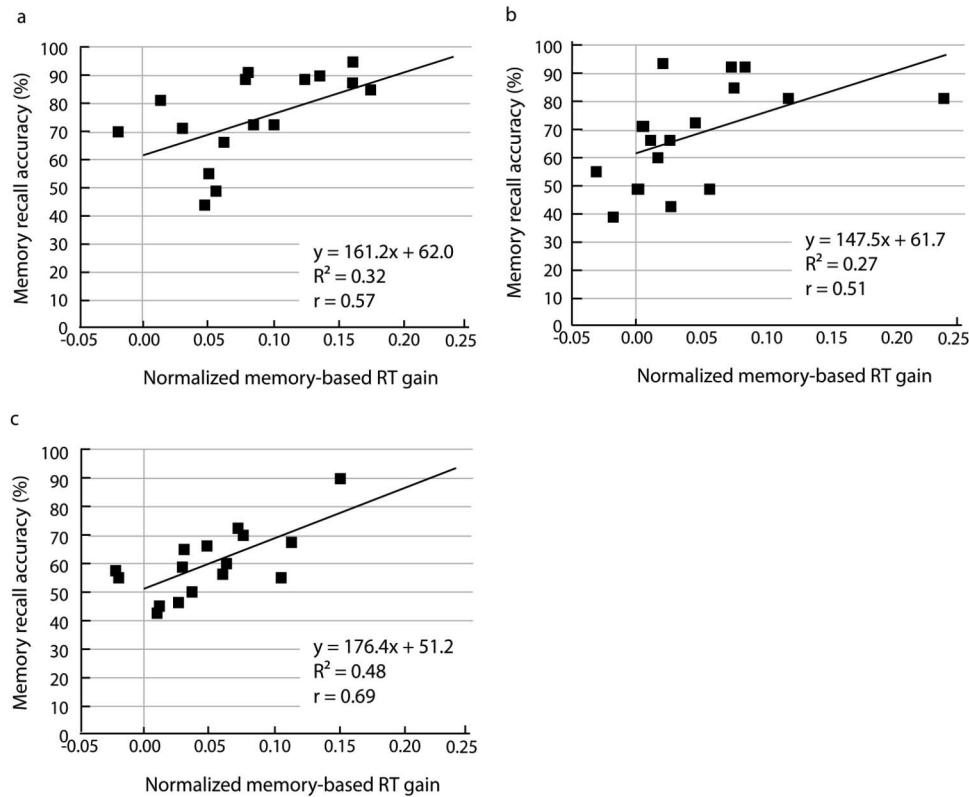[4] Data from this analysis was taken from Experiment 1.

Figure 3. Relationship between memory for target location and attentional orientation: Individuals that remembered target-context associations with greater recall accuracy also demonstrated stronger memory-based attentional change. Memory-based attentional change was calculated as the gain in RT to detect targets embedded within memory cue audio clips as compared to within neutral audio clips. Difference scores were normalized (Mneutral – Mmemory/Mneutral +Mmemory). (a) Memory-based attentional change in Experiment 1, where participants were exposed to eight learning blocks. (b) Memory-based attentional change in Experiment 2, where temporal context of the target was changed from originally learned context. (c) Memory-based attentional change in Experiment 3, where participants were exposed to four learning blocks.

Importantly, our results also showed that spatial attention is modulated by auditory LTM despite interruptions in expected timing. To assess the relationship between memory for the target location, and the ensuing modulation of spatial attention, we calculated a Pearson's correlation between participants' recall performance and normalized memory-cue based RT difference scores. The correlation was significant ($r = .51$, $p < .05$; see Figure 3b), and comparable to that observed in Experiment 1 ($r = .57$, $p < .05$; see Figure 3a) where temporal context of the target was held constant between learning and test (Fisher's exact test $p = .83$). This confirms that the impact of spatial memory cues on perception is robust and that memory-guided attention persists even when the timing of that spatial information is changed.

Implicit memory-guided auditory spatial attention.[5] Participants showed enhanced performance for audio clips in which target location had been learned (i.e., memory cue trials) even when the association between the audio clip and target location was not explicitly recalled. For memory-incorrect trials, targets preceded by memory cues were located more quickly ($F [1, 10] = 6.58$, $p < .05$; $d = 0.59$; 95% CI for d [-0.32, 1.50]) and with greater perceptual sensitivity ($F [1, 10] = 8.23$, $p < .05$; $d = 0.67$;

95% CI for d [-0.24, 1.59,]) than those preceded by neutral cues (see Table 1). Memory cue accounted for 40% ($\eta^2 = .40$) of variance in RTs in the model for memory-incorrect trials, which was nearly as strong as the effect size reported based on explicitly remembered trials ($\eta^2 = .49$).

Moreover, cue-based gains in RT ($M = 79.18$ ms, $SE = 15.85$ ms) and d' ($M = 0.49$, $SE = 0.69$) were large, and comparable to the gains reported for explicitly remembered trials. A two-way RMANOVA testing the effect of cue (memory, neutral) and recall (explicitly remembered trials, memory-incorrect trials) on RT revealed no significant interaction, $F (1, 10) = 1.66$, $p = .23$. This finding suggests that implicit memory modulates auditory spatial attention, equivalent to the influence of consciously accessible associations. Our results complement research on visual implicit contextual cueing (Chun & Jiang, 2003), and suggest that LTM might engage a common supramodal spatial attention-control network.

___

[5] This analysis comprised data from Experiment 1 and Experiment 2 because of inadequate number of trials in each experiment alone.

Experiment 3

In both Experiments 1 and 2, perceptual sensitivity reached asymptote after four blocks of trials (see Figure 2). This suggests that participants quickly associated specific audio clips with a lateralized target. In Experiment 3, we reduced the number of learning trials from eight to four blocks of trials to assess whether a more limited exposure could yield LTM target-context associations that would facilitate spatial attention. We hypothesized that presenting fewer repetitions of target-audio clip pairings (Experiment 3) may impair memory-guided attention compared to when a greater number of learning trials are performed (Experiment 1). Based on results of Experiment 2, where explicit and implicit memory had similar effects on memory-guided attention, we anticipated this effect of exposure for both explicitly accessible and inaccessible associations.

Method

Sixteen young adults ($M = 24.06$ years; range 20–28 years; 8 males) recruited from the Rotman Research Institute participated in Experiment 3 for monetary compensation. None of the participants took part in Experiment 1 or Experiment 2. The materials and procedures were the same as those of Experiment 1, except that participants completed only four learning blocks instead of eight. We compared data in this experiment with those of Experiment 1 because all the methodology was held constant, with the only modification to the paradigm being the amount of exposure to target-context associations during learning.

Results and Comment

As in Experiment 1 and Experiment 2, participants were quicker and more accurate ([d']) to locate the target with increased practice. Over four learning blocks, participants showed 96-ms gains in RT and over 8% gains in correct detection accuracy. These gains were comparable to 113-ms gains in RT and over 13% gains in accuracy in Experiment 1, which had eight learning blocks instead of four. The RMANOVA on RT with learning block (block 1–4) and target (present, absent) factors yielded main effects of learning block, $F(3, 45) = 6.57$, $p < .005$, target, $F(1, 15) = 67.33$, $p < .001$, as well as a significant interaction, $F(3, 45) = 9.03$, $p < .001$. Again, RT gains were strongest over the first two learning blocks, with 30% of the variance in RTs attributable to effects of learning block (block 1: RT $M = 571.95$, $SE = 40.00$; block 2: $M = 498.13$, $SE = 28.53$; block 3: $M = 481.48$, $SE = 34.20$; block 4: $M = 475.71$, $SE = 32.71$). Perceptual sensitivity increased significantly over learning blocks for all trials, $F(3, 45) = 19.89$, $p < .001$, as well as for target-present trials ($F[3, 45] = 6.98$, $p < .005$; block 1: $d' = 1.78$, $SE = 0.22$; block 2: $d' = 2.63$, $SE = 0.30$; block 3: $d' = 2.25$, $SE = 0.28$; block 4: $M = 2.84$, $SE = 0.28$).

Robust memory for auditory target-context associations is formed following limited exposure. Participants remembered 60% of target-context associations (i.e., the presence/location of the target within audio clips) on average, which was significantly more than expected by chance, $t(15) = 8.82$, $p < .001$, $d = 2.20$, 95% CI for d [1.30, 3.21]. Above-chance recall was also observed when target-absent trials were excluded from the analysis (i.e.,

when analyzing target-present trials only), $t(15) = 2.20$ $p < .05$, $d = 0.55$, 95% CI for d [0.01, 1.10]. However, recall of specific location of auditory targets was reduced by more than 15% when participants heard audio clips four times during learning (Experiment 3: $M = 59.84\%$, $SE = 3.01$) compared to eight times (Experiment 1: $M = 75.47\%$, $SE = 3.95\%$), $F(1, 30) = 9.90$, $p < .005$, $d = 1.11$, 95% CI for d [0.34, 1.89].

As in Experiment 1, we tested whether participants' confidence in their response reflected response accuracy. A two factor RMANOVA revealed that the interaction of confidence level and correct recall did not reach significance, indicating that participants were unable to gauge the correctness of their recall when exposure to the recalled items was limited during the learning phase.

Memory-guided auditory spatial attention persists even when associations in memory are less salient. While conscious recall of the auditory target's location within audio clips is reduced with fewer repetitions (i.e., number of learning trials), the influence of LTM associations on perceptual processes remains unaffected. RMANOVAs revealed significant cue-based gains in RT, $F(1, 15) = 17.25$, $p < .005$ ($d = 0.41$, 95% CI for d [−0.32, 1.14]), and d', $F(1, 15) = 45.05$, $p < .001$ ($d = 1.36$, 95% CI for d [0.57, 2.17]; see Table 1). Similar gains in RT, $F(1, 15) = 6.62$, $p < .05$ ($d = 0.44$, 95% CI for d [−0.29, 1.17]), and d', $F(1, 15) = 15.55$, $p = .001$ ($d = 1.35$, 95% CI for d [0.55, 2.15]), were observed when only those trials available for conscious recall were analyzed (i.e., memory-correct trials; see Table 1).

ANOVAs revealed that there was no significant change in memory-guided attention based on the amount of exposure during learning (i.e., eight learning blocks in Experiment 1 vs. four in Experiment 3), in terms of RT, $F(1, 30) = 2.19$, $p = .15$. However, an effect of exposure on d' was found ($F[1, 30] = 5.89$, $p < .05$). This was the case when we included all trials in the analysis, as well as for memory-correct trials only. Our results show that four blocks of learning trials are sufficient to create long-term auditory target-context associations capable of modulating spatial attention, and that adding additional training does not make much difference.

To assess the relationship between memory, and the ensuing modulation of spatial attention, we calculated a Pearson's correlation between participants' recall performance and normalized memory-cue based RT difference scores. The correlation was significant ($r = .69$, $p < .005$; see Figure 3c), and comparable to that observed in Experiment 1 ($r = .57$, $p < .05$; see Figure 3a) where more training was administered (Fisher's exact test $p = .61$). This provides further evidence that explicit memory-guided orienting effects occur even with limited opportunity to acquire memory during learning.

Learning effects on implicit memory-guided attention. This analysis was performed using 11 participants, which fulfilled the inclusion criteria used in Experiment 1 and 2 for memory-incorrect trials. The RTs to memory and neutral cue trials were comparable when exposure was limited (i.e., four learning blocks instead of eight), $F(1, 10) = 1.55$, $p = .24$. Perceptual sensitivity to memory and neutral cue trials was also comparable, $F(1, 10) = 1.53$, $p = .24$. It appears that implicit memory-guided attention requires more exposure than explicit memory-guided attention. It may be the case that there was not enough statistical power to observe the expected implicit cueing effects due to the small

number of participants (see "General Discussion" for further discussion).

Overall, the results of Experiment 3 suggest that limited exposure to auditory targets within audio clips of everyday situations yield robust LTM associations that optimize attentional orientation. While decreasing exposure does not have an effect on explicit memory-guided attention, it does impede implicit auditory memory-guided attention effects.

General Discussion

Formation of Long-term Associations Between an Auditory Target and Complex Audio-Clips

In the present study, participants were very efficient at learning the location of a target stimulus within an audio clip of a real life event. Learning effects were replicated across three different experiments. The learning curve is consistent with prior work using speech stimuli, with steep learning taking place within few repetitions followed by more gradual benefits with subsequent presentations (Alain, Campeanu, & Tremblay, 2010; Alain, Snyder, He, & Reinke, 2007; Ben-David, Campeanu, Tremblay, & Alain, 2011; Du et al., 2015; Shen & Alain, 2011, 2012). A similar learning trend is observed in visual contextual cueing paradigms, where object location is cued by its containing contextual array (Chun, 2000).

Up to now, our knowledge on auditory memory is based, to a great extent, on speech sounds (Baddeley, 1990) and simple sound stimuli such as pure tones (see Cowan, 1984; Jääskeläinen, Ahveninen, Belliveau, Raij, & Sams, 2007). Some have suggested that we may be remarkably poor at memorizing natural sounds (Cohen et al., 2009), as opposed to visual memory which can store a great amount of detail of visual scenes following a single exposure (Brady, Konkle, Alvarez, & Oliva, 2008). However, the current results indicate that memory for complex audio clips, and associations with specific spatial locations, is robust even after a few exposures (i.e., Experiment 3). Spatial memory formation (e.g., learning to localize targets within target-present trials) follows a time course similar to that for simple target detection learning (e.g., learning to associate an audio clip with target presence/absence).

Spatial Orientation in Remembered Audio-Clips

The results of all three experiments provide converging evidence that auditory LTM biases auditory spatial attention. "Memory" audio clips successfully cued spatial attention toward expected target location, whereas "neutral cue" audio clips did not facilitate target detection because attention was likely divided between the left and right auditory field. Memory-guided attention is robust and persists under a variety of listening conditions. For example, strong modulation of attention by memory was observed even when temporal target-context associations were different from those at learning (Experiment 2), or with limited exposure to the target-context associations during learning (Experiment 3). Overall, search performance for targets within an auditory contextual cueing paradigm, like visual contextual cueing, is fast and automatic, and not easily disrupted (Woodman, Carlisle, & Reinhart, 2013). Memory-guided attention is likely a higher-order

process that is independent of sensory modality; memory traces for auditory and visual information affect spatial attention in a similar manner.

We showed that auditory attention is modulated by both spatial and temporal knowledge (i.e., expectation); participants attended to a particular spatial location and temporal position within audio clips based on learnt target-context associations. These findings challenge previous work that suggests that only temporal expectations, but not spatial expectations alone, facilitate orientation. For example, Rimmele et al. (2011) conducted a study in which participants were made to anticipate the reappearance of a target tone based on a preceding sequence of tones, with spatial expectations created by presenting tones in a spatially constant motion along a trajectory, and temporal expectations created by holding the timing of the stimulus constant. Only temporal expectations created RT and accuracy benefits.

In the present study, we showed that spatial expectation also plays a significant role in driving attention, occurring when temporal context is held constant as well as when it is modified from learning. It is important to note that Rimmele et al. (2011) created expectations for the target (both spatial and temporal) based on an immediately preceding sequence of auditory events, thereby examining how attention is guided by STM. We created expectations by cueing participants to the location or timing of a target based on information in LTM. Therefore, we reasoned that temporal expectations predominantly bias attention when information is in an active, readily available state (Rimmele et al., 2011; Sanders & Astheimer, 2008; Shen & Alain, 2011), but when spatial and temporal information are encoded in LTM, then attention can be modulated by both kinds of information. This conclusion is consistent with evidence that both spatial and temporal cues that activate visual LTM have strong effects on visuospatial attention (spatial cueing: Patai et al., 2012; temporal cueing: Cravo, Rohenkohl, Santos, & Nobre, in press).

As with vision, the effect of spatial cueing in the current auditory study also depends on the nature of the spatial memory cue itself, which was different from the cueing used by Rimmele et al. (2011; and previously by Doherty, Rao, Mesulam, & Nobre, 2005 in visual parallels). Spatial expectations were established by left/right localization, and were perhaps more coarse than the motion trajectory used by Rimmele et al. (2011) to create expectations in STM. In addition, as we did not employ direct temporal cueing, as opposed to spatial cueing, temporal expectations were not reactivated to the same degree as spatial expectations. Together, these two factors may have emphasized the importance of spatial expectations and reduced the importance of temporal expectations in target processing in our study compared to Rimmele et al. (2011).

Few studies have examined effects of number of exposures directly, though visual work generally corroborates fast target-context association learning, and the effects of limited learning trials on visual memory-guided attention. Most parallel visual paradigms use five to six learning blocks to establish strong contextual memories (Patai et al., 2012; Summerfield et al., 2011), though as little as a single exposure to visual targets within scenes can guide spatial attention (Rosen, Stern, Michalka, Devaney, & Somers, 2016). Four learning blocks in the current study were sufficient to create robust auditory memory that guided attention.

We were also interested in examining whether memory-guided orientation persists after some time. Since no audio clips within the second session of the memory-guided attention task were entirely neutral cues (i.e., participants may have created target-context associations from the first session for neutral trials), these results also gauge interference (i.e., interference of the first session on previously learned associations from the learning task). The results from the second session showed that auditory associations in memory have a strong effect on auditory attention, in terms of increased speed and perceptual sensitivity, and that memory-guided attention is quite resistant to interference and/or fading over the interval we tested. Evidence that the effect is robust and persists over testing sessions lays the foundation for studying the influence of LTM on attention using longer delays (e.g., days, or week(s)).

Implicit Associations Guiding Orientation of Attention Within Audio-Clips

The results demonstrate that attentional systems in the auditory domain also rely largely on implicit top-down processes. That is, even those associations that are not consciously accessible can bias auditory spatial attention. Participants were able to use audio clips as implicit memory cues to guide attention toward targets when they were allowed sufficient exposure to the audio clip—target pairs during learning (in Experiment 1 and 2). Our results complement findings in vision where explicit contextual memories are not necessary to drive memory-guided attention. For example, Salvato, Patai, & Nobre (2016) showed that memory-guided orienting of attention in scene-based contexts is preserved even when explicit memory is impaired.

However, the effects of consciously accessible and inaccessible memory on attentional guidance have previously not been directly compared. Implicit and explicit memory differ from one another because implicit processing lacks intentional control. Moreover, particularly in audition, the semantic knowledge present in explicit memory may be important in guiding attention, which is not the case in implicit memory-guided attention (see Introduction). Nonetheless, Experiment 2 showed that consciously inaccessible memory had the same effect on attention as accessible memory. This result suggests that conceptual processing and semantic elaboration associated with explicit memory is not necessary to facilitate memory-guided attention in audition. Data-driven or perceptual-based processes provide sufficient and seemingly equivalent forces that can direct attention to location during auditory scene analysis. The ability for implicit memory traces to modulate auditory perception occurs not only in simple recognition priming paradigms (Agus, Thorpe, & Pressnitzer, 2010), but also in the more complex spatial cueing paradigms as in the current study.

Attentional guidance by consciously inaccessible memory is highly adaptive and evolutionarily significant. Explicit memory reactivation is costly in terms of time and cognitive resources. The biological significance of implicit memory-guided attention is further supported by studies showing that rewarding context accelerates implicit attentional guidance in visual search (Tseng & Lleras, 2013). Moreover, implicit contextual cueing effects are robust in young children (Darby, Burling, & Yoshida, 2014) as well as animals (Wasserman, Teng, & Brooks, 2014). Highly familiar auditory information is used to modulate attention in everyday listening situations, for example when attention is selectively focused on familiar voices (Johnsrude et al., 2013).

It may be the case, however, that auditory information that is not as strongly coded (i.e., due to lack of practice and familiarity) does not have this capability. With fewer repetitions of the audio clip–target pairings (Experiment 3), the attentional bias was not reproduced for consciously inaccessible trials, possibly because (a) more repetitions are needed for consolidation of target-context associations; (b) the context itself was not coded implicitly, let alone the target-context association; and (c) implicit associations were formed but were not strong enough to facilitate a modulation of attention toward targets. Since we did not have a direct measure of implicit memory in the current paradigm, it was difficult to assess which of these alternatives was supported.

Future Directions

In addition to quantitative differences between sensory memories (better memory, better identification, etc.), there may also be qualitative distinctions between auditory and visual memory. For instance, while all sensory systems may share similar characteristics and rely on comparable mechanisms as revealed by previous research (Snyder & Gregg, 2011) as well as our findings, each system may be unique in different ways. For example, it could be the case that auditory stimuli are predominant in some circumstances, and visual ones, in another. It may be interesting to examine whether attention is better allocated when both audio clips and target are presented in the same modality as compared to different modalities (e.g., visual scene and auditory target tone). Moreover, we could examine coordination between audition and vision using dynamic stimuli (e.g., a visual movie clip with a soundtrack).

To develop a better understanding of memory-based deployment of auditory attention, neural correlates of the behavioral orienting effect should be studied. In the visual domain, Summerfield et al. (2011) and Patai et al. (2012) showed that LTM for the location of visual targets within scene contexts generated spatial biases in neural activity in anticipation of the target stimulus. Event-related potentials were also modulated by memory cues during target selection. For example, the N2pc component, which is related to sensitivity to and selection of visual stimuli (Kuo, Rao, Lepsien, & Nobre, 2009), was significantly attenuated for targets preceded by valid memory cues as compared to targets appearing within neutral cue scenes (e.g., Stokes et al., 2012). Examining changes in brain activity related to memory-guided modulation of auditory spatial attention will provide greater insight into the mechanisms of the auditory memory and attention interaction, and will also allow us to study the effects of implicit versus explicit memory for sound in more depth. Moreover, studying related neural processes will enable disassociation of biases triggered by spatial memory for targets within clips from those arising from perceptual modulation of target processing (Summerfield et al., 2011).

In the current study, memory cues only provided information about the spatial location of the target stimulus. Although expectations for timing were generated in the learning phase and biased responses, participants were cued spatially and asked to respond to spatial location. In addition to spatial cues, providing temporal cues to directly guide attention would allow us to further assess the

effects of LTM for timing on attention, and determine whether temporal and spatial cueing in audition activates separate brain structures and processes. Based on what is known about implicit expectations formed by regularities in space and time, both in vision (Doherty et al., 2005) and audition (Rimmele et al., 2011), we might expect that temporal cueing will have effects at earlier processing stages than spatial cueing. This may be the case particularly in audition where temporal processing often dominates spatial processing.

Concluding Remarks

The present study aimed to develop a preliminary understanding of the effects of auditory memory on spatial attention, which has remained a vastly uncharted area until now. Indeed, to our knowledge this study was the first to demonstrate that memory for auditory stimuli can influence deployment of auditory spatial attention sometime later (after 1 hr), laying the foundation for a host of related investigations. For example, once auditory memory-guided attention mechanisms are better understood, the theory can be applied to the optimization of hearing and communication in ageing populations or acoustically affected stroke patients. We provide evidence for the role of both consciously accessible (explicit) and inaccessible (implicit) long-term contextual memories in biasing auditory spatial attention. Individuals are able to create robust associations for target location within cluttered naturalistic audio clips that have a lasting influence on spatial attention under a variety of conditions.

References

Agus, T. R., Thorpe, S. J., & Pressnitzer, D. (2010). Rapid formation of robust auditory memories: Insights from noise. *Neuron, 66,* 610–618. http://dx.doi.org/10.1016/j.neuron.2010.04.014

Alain, C., Campeanu, S., & Tremblay, K. (2010). Changes in sensory evoked responses coincide with rapid improvement in speech identification performance. *Journal of Cognitive Neuroscience, 22,* 392–403. http://dx.doi.org/10.1162/jocn.2009.21279

Alain, C., Snyder, J. S., He, Y., & Reinke, K. S. (2007). Changes in auditory cortex parallel rapid perceptual learning. *Cerebral Cortex, 17,* 1074–1084. http://dx.doi.org/10.1093/cercor/bhl018

Backer, K. C., & Alain, C. (2012). Orienting attention to sound object representations attenuates change deafness. *Journal of Experimental Psychology-Human Perception and Performance, 38,* 1554–1566. http://dx.doi.org/10.1037/a0027858

Baddeley, A. D. (1990). *Human memory: Theory and practice*. Needham Heights, MA: Allyn & Bacon.

Ben-David, B. M., Campeanu, S., Tremblay, K. L., & Alain, C. (2011). Auditory evoked potentials dissociate rapid perceptual learning from task repetition without learning. *Psychophysiology, 48,* 797–807. http://dx.doi.org/10.1111/j.1469-8986.2010.01139.x

Berry, A. S., Li, X., Lin, Z., & Lustig, C. (2014). Shared and distinct factors driving attention and temporal processing across modalities. *Acta Psychologica, 147,* 42–50. http://dx.doi.org/10.1016/j.actpsy.2013.07.020

Bey, C., & McAdams, S. (2002). Schema-based processing in auditory scene analysis. *Perception & Psychophysics, 64,* 844–854. http://dx.doi.org/10.3758/BF03194750

Bigelow, J., & Poremba, A. (2012). Comparing short-term memory among sensory modalities. *Visual Cognition, 20,* 1012–1016. http://dx.doi.org/10.1080/13506285.2012.726452

Bizley, J. K., & Cohen, Y. E. (2013). The what, where and how of auditory-object perception. *Nature Reviews Neuroscience, 14,* 693–707. http://dx.doi.org/10.1038/nrn3565

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences of the United States of America, 105,* 14325–14329. http://dx.doi.org/10.1073/pnas.0803390105

Brockmole, J., & Henderson, J. (2006). Using real-world scenes as contextual cues for search. *Visual Cognition, 13,* 99–108. http://dx.doi.org/10.1080/13506280500165188

Chaumon, M., Hasboun, D., Baulac, M., Adam, C., & Tallon-Baudry, C. (2009). Unconscious contextual memory affects early responses in the anterior temporal lobe. *Brain Research, 1285,* 77–87. http://dx.doi.org/10.1016/j.brainres.2009.05.087

Cherry, C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the Acoustical Society of America, 25,* 975–979. http://dx.doi.org/10.1121/1.1907229.ISSN0001-4966

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences, 4,* 170–178. http://dx.doi.org/10.1016/S1364-6613(00)01476-5

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36,* 28–71. http://dx.doi.org/10.1006/cogp.1998.0681

Chun, M. M., & Jiang, Y. (2003). Implicit, long-term spatial contextual memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 224–234. http://dx.doi.org/10.1037/0278-7393.29.2.224

Ciaramelli, E., Grady, C., Levine, B., Ween, J., & Moscovitch, M. (2010). Top-down and bottom-up attention to memory are dissociated in posterior parietal cortex: Neuroimaging and neuropsychological evidence. *The Journal of Neuroscience, 30,* 4943–4956. http://dx.doi.org/10.1523/JNEUROSCI.1209-09.2010

Ciaramelli, E., Lin, O., & Moscovitch, M. (2009). Episodic memory for spatial context biases spatial attention. *Experimental Brain Research, 192,* 511–520. http://dx.doi.org/10.1007/s00221-008-1548-9

Cohen, M. A., Evans, K. K., Horowitz, T. S., & Wolfe, J. M. (2011). Auditory and visual memory in musicians and nonmusicians. *Psychonomic Bulletin & Review, 18,* 586–591. http://dx.doi.org/10.3758/s13423-011-0074-0

Cohen, M. A., Horowitz, T. S., & Wolfe, J. M. (2009). Auditory recognition memory is inferior to visual recognition memory. *Proceedings of the National Academy of Sciences of the United States of America, 106,* 6008–6010. http://dx.doi.org/10.1073/pnas.0811884106

Conway, A. R., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin & Review, 8,* 331–335. http://dx.doi.org/10.3758/BF03196169

Cowan, N. (1984). On short and long auditory stores. *Psychological Bulletin, 96,* 341–370. http://dx.doi.org/10.1037/0033-2909.96.2.341

Cravo, A., Rohenkohl, G., Santos, K., & Nobre, A. (in press). *Dynamic anticipation based on temporal associations in memory*.

Darby, K., Burling, J., & Yoshida, H. (2014). The role of search speed in the contextual cueing of children's attention. *Cognitive Development, 29,* 17–29. http://dx.doi.org/10.1016/j.cogdev.2013.10.001

Doherty, J., Rao, A., Mesulam, M., & Nobre, A. (2005). Synergistic effect of combined temporal and spatial expectations on visual attention. *Journal of Neuroscience, 25,* 8259–8266. http://dx.doi.org/10.1523/JNEUROSCI.1821-05.2005

Dowling, W. J., Lung, K. M., & Herrbold, S. (1987). Aiming attention in pitch and time in the perception of interleaved melodies. *Perception & Psychophysics, 41,* 642–656. http://dx.doi.org/10.3758/BF03210496

Du, Y., He, Y., Arnott, S. R., Ross, B., Wu, X., Li, L., & Alain, C. (2015). Rapid tuning of auditory "what" and "where" pathways by training. *Cerebral Cortex, 25,* 496–506. http://dx.doi.org/10.1093/cercor/bht251

Dulas, M. R., & Duarte, A. (2013). The influence of directed attention at encoding on source memory retrieval in the young and old: An ERP study. *Brain Research, 1500,* 55–71. http://dx.doi.org/10.1016/j.brainres.2013.01.018

Fernandes, M. A., & Moscovitch, M. (2000). Divided attention and memory: Evidence of substantial interference effects at retrieval and encoding. *Journal of Experimental Psychology: General, 129,* 155–176. http://dx.doi.org/10.1037/0096-3445.129.2.155

Goldfarb, E. V., Chun, M. M., & Phelps, E. A. (2016). Memory-guided attention: Independent contributions of the hippocampus and striatum. *Neuron, 89,* 317–324. http://dx.doi.org/10.1016/j.neuron.2015.12.014

Günther, T., Konrad, K., Häusler, J., Saghraoui, H., Willmes, K., & Sturm, W. (2014). Developmental differences in visual and auditory attention: A cross-sectional study. *Zeitschrift für Neuropsychologie, 25,* 143–152. http://dx.doi.org/10.1024/1016-264X/a000126

Jääskeläinen, I. P., Ahveninen, J., Belliveau, J. W., Raij, T., & Sams, M. (2007). Short-term plasticity in auditory cognition. *Trends in Neurosciences, 30,* 653–661. http://dx.doi.org/10.1016/j.tins.2007.09.003

Jiang, Y., & Leung, A. W. (2005). Implicit learning of ignored visual context. *Psychonomic Bulletin & Review, 12,* 100–106. http://dx.doi.org/10.3758/BF03196353

Johnson, J. S., Woodman, G. F., Braun, E., & Luck, S. J. (2007). Implicit memory influences the allocation of attention in visual cortex. *Psychonomic Bulletin & Review, 14,* 834–839. http://dx.doi.org/10.3758/BF03194108

Johnsrude, I. S., Mackey, A., Hakyemez, H., Alexander, E., Trang, H. P., & Carlyon, R. P. (2013). Swinging at a cocktail party: Voice familiarity aids speech perception in the presence of a competing voice. *Psychological Science, 24,* 1995–2004. http://dx.doi.org/10.1177/0956797613482467

Kunar, M. A., Flusberg, S. J., & Wolfe, J. M. (2006). Contextual cuing by global features. *Perception & Psychophysics, 68,* 1204–1216. http://dx.doi.org/10.3758/BF03193721

Kuo, B. C., Rao, A., Lepsien, J., & Nobre, A. C. (2009). Searching for targets within the spatial layout of visual short-term memory. *The Journal of Neuroscience, 29,* 8032–8038. http://dx.doi.org/10.1523/JNEUROSCI.0952-09.2009

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America, 49*(Suppl. 2), 467.

Nakane, T., Miyakoshi, M., Nakai, T., & Naganawa, S. (2016). How the non-attending brain hears its owner's name. *Cerebral Cortex, 26,* 3889–3904. http://dx.doi.org/10.1093/cercor/bhv184

Patai, E. Z., Doallo, S., & Nobre, A. C. (2012). Long-term memories bias sensitivity and target selection in complex scenes. *Journal of Cognitive Neuroscience, 24,* 2281–2291. http://dx.doi.org/10.1162/jocn_a_00294

Rimmele, J., Jolsvai, H., & Sussman, E. (2011). Auditory target detection is affected by implicit temporal and spatial expectations. *Journal of Cognitive Neuroscience, 23,* 1136–1147. http://dx.doi.org/10.1162/jocn.2010.21437

Röer, J. P., Bell, R., & Buchner, A. (2013). Self-relevance increases the irrelevant sound effect: Attentional disruption by one's own name. *Journal of Cognitive Psychology, 25,* 925–931. http://dx.doi.org/10.1080/20445911.2013.828063

Rosen, M. L., Stern, C. E., Michalka, S. W., Devaney, K. J., & Somers, D. C. (2016). Cognitive control network contributions to memory-guided visual attention. *Cerebral Cortex, 26,* 2059–2073. http://dx.doi.org/10.1093/cercor/bhv028

Rosenbaum, G. M., & Jiang, Y. V. (2013). Interaction between scene-based and array-based contextual cueing. *Attention, Perception, & Psychophysics, 75,* 888–899. http://dx.doi.org/10.3758/s13414-013-0446-9

Roye, A., Schröger, E., Jacobsen, T., & Gruber, T. (2010). Is my mobile ringing? Evidence for rapid processing of a personally significant sound

in humans. *The Journal of Neuroscience, 30,* 7310–7313. http://dx.doi.org/10.1523/JNEUROSCI.1113-10.2010

Salmi, J., Rinne, T., Degerman, A., & Alho, K. (2007). Orienting and maintenance of spatial attention in audition and vision: An event-related brain potential study. *The European Journal of Neuroscience, 25,* 3725–3733. http://dx.doi.org/10.1111/j.1460-9568.2007.05616.x

Salvato, G., Patai, E. Z., & Nobre, A. C. (2016). Preserved memory-based orienting of attention with impaired explicit memory in healthy ageing. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior, 74,* 67–78. http://dx.doi.org/10.1016/j.cortex.2015.10.019

Sanders, L. D., & Astheimer, L. B. (2008). Temporally selective attention modulates early perceptual processing: Event-related potential evidence. *Perception & Psychophysics, 70,* 732–742. http://dx.doi.org/10.3758/PP.70.4.732

Shen, D., & Alain, C. (2011). Temporal attention facilitates short-term consolidation during a rapid serial auditory presentation task. *Experimental Brain Research, 215*(3–4), 285–292. http://dx.doi.org/10.1007/s00221-011-2897-3

Shen, D., & Alain, C. (2012). Implicit temporal expectation attenuates auditory attentional blink. *PLoS ONE, 7*(4), e36031. http://dx.doi.org/10.1371/journal.pone.0036031

Silva, M. M., Groeger, J. A., & Bradshaw, M. F. (2006). Attention-memory interactions in scene perception. *Spatial Vision, 19,* 9–19. http://dx.doi.org/10.1163/156856806775009223

Snyder, J. S., & Gregg, M. K. (2011). Memory for sound, with an ear toward hearing in complex auditory scenes. *Attention, Perception, & Psychophysics, 73,* 1993–2007. http://dx.doi.org/10.3758/s13414-011-0189-4

Stokes, M. G., Atherton, K., Patai, E. Z., & Nobre, A. C. (2012). Long-term memory prepares neural activity for perception. *Proceedings of the National Academy of Sciences of the United States of America, 109*(6), E360–E367. http://dx.doi.org/10.1073/pnas.1108555108

Summerfield, J. J., Lepsien, J., Gitelman, D. R., Mesulam, M. M., & Nobre, A. C. (2006). Orienting attention based on long-term memory experience. *Neuron, 49,* 905–916. http://dx.doi.org/10.1016/j.neuron.2006.01.021

Summerfield, J. J., Rao, A., Garside, N., & Nobre, A. C. (2011). Biasing perception by spatial long-term memory. *The Journal of Neuroscience, 31,* 14952–14960. http://dx.doi.org/10.1523/JNEUROSCI.5541-10.2011

Treisman, A. M. (1964). Selective attention in man. *British Medical Bulletin, 20,* 12–16. http://dx.doi.org/10.1093/oxfordjournals.bmb.a070274

Tseng, Y. C., & Lleras, A. (2013). Rewarding context accelerates implicit guidance in visual search. *Attention, Perception, & Psychophysics, 75,* 287–298. http://dx.doi.org/10.3758/s13414-012-0400-2

Wasserman, E. A., Teng, Y., & Brooks, D. I. (2014). Scene-based contextual cueing in pigeons. *Journal of Experimental Psychology: Animal Learning and Cognition, 40,* 401–418. http://dx.doi.org/10.1037/xan0000028

Wood, N., & Cowan, N. (1995). The cocktail party phenomenon revisited: How frequent are attention shifts to one's name in an irrelevant auditory channel? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 255–260. http://dx.doi.org/10.1037/0278-7393.21.1.255

Woodman, G. F., Carlisle, N. B., & Reinhart, R. M. G. (2013). Where do we store the memory representations that guide attention? *Journal of Vision, 13,* 1. http://dx.doi.org/10.1167/13.3.1

Zimmermann, J. F., Moscovitch, M., & Alain, C. (2016). Attending to auditory memory. *Brain Research, 1640,* 208–221. http://dx.doi.org/10.1016/j.brainres.2015.11.032