Exploring Visual Attention in Musicians: Temporal, Spatial and Capacity Considerations

https://escholarship.org/uc/item/8s57282c


1069-7977

Lim, Ahnate
Sinnett, Scott

2011

Peer reviewed
Exploring Visual Attention in Musicians: Temporal, Spatial and Capacity Considerations

Ahnate Lim (ahnate@hawaii.edu)
Department of Psychology, University of Hawaii at Manoa
2530 Dole Street, Honolulu, HI 96822 USA

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

Abstract
Considerable evidence converges on how attention can be modulated through training (e.g., video game playing). While previous research suggests that musical training can modulate early perceptual and attentional processes, no single investigation to date has been conducted on the same participants to measure specific mechanisms of attention (temporal, spatial, and capacity) in musicians. In Experiment 1 we used a temporal order judgment (TOJ) task with both exogenous and endogenous cues in order to measure temporal and spatial attention. In Experiment 2, a cued-target detection task was presented with a concurrent high load task to assess capacity processing in musicians. Of the three measures, musicians performed better than controls on two, demonstrating a lower threshold for judging temporal order in addition to increased capabilities to process distracting information despite attentional resources being largely depleted. Together, these results provide novel findings on multiple aspects of attention in musicians.

Keywords: attention, musician, plasticity, temporal order judgment, spatial, capacity, exogenous, endogenous

Introduction
Attention is a fundamental cognitive mechanism that enables humans to select the most crucial information from a constant array of sensory input, thereby allowing for efficient and effective functioning. Interestingly, exposure to particular experiences has been shown to modulate various aspects of human attention, as seen from evidence at both the behavioral and neurological levels.

Recent neurological evidence suggests that brain functioning involves distributive processing and plastic characteristics (Mercado, 2008; Mesulam, 1990). For instance, this “plasticity” of the brain can be seen in common interactions across sensory modalities (Shimojo & Shams, 2001), when adapting to particular conditions such as age, disease, stress, and even addiction (Kolb, Gibb, & Robinson, 2003), or even in congenitally blind adults, as a study by Röder et al. (1999) revealed greater peripheral spatial localization abilities and more finely tuned early attentional mechanisms when compared to blindfolded controls. These and other studies suggest not only that the brain can partially compensate for losses in one modality through enhancements in another (for an example of tactile compensation, see Borsook et al., 1998), but also that attentional mechanisms can be correspondingly modulated.

Plasticity can also be seen in attentional and perceptual mechanisms of populations not suffering any sensory loss, with many recent findings indicating that modulation can occur as a side effect of particular daily activities or hobbies. For instance, a recent functional magnetic resonance imaging (fMRI) study found video-game players to have more prefrontal cortex activity during complex non-gaming tasks when compared with non-players, a change attributed to the demands on spatial attention while training with video games (Granek, Gorbet, & Sergio, 2010). Dovetailing with this finding, behavioral experiments suggest enhanced performance, including greater availability of attentional resources on various paradigms such as multiple object-tracking, enumeration, perceptual load, and the Attentional Network Test (ANT, see Posner & Rothbart, 2007; Dye, Green, & Bavelier, 2009; for review see I. Spence & Feng, 2010). One might ask then whether regular practice of other more ancient and ubiquitous activities, such as musical performance, would also result in augmentation of specific information processing capabilities in musicians?

Indeed, the topic of non-musical benefits from musical exposure has seen considerable research (not in the least due to public interest and popular notions such as the “Mozart effect”), with results suggesting enhancements in areas such as mathematics, language, spatial abilities, and memory. However, many of these ‘listening’ experiments (e.g., Rauscher, Shaw, & Ky, 1993) involved performing tasks after listening to excerpts of classical music, and have generally provided inconclusive evidence, with follow-up studies showing only limited effects, or effects accountable to mood arousal or other indirect factors (Schellenberg, 2001; Steele et al., 1999; Thompson, Schellenberg, & Husain, 2001).

Other studies take a different approach by instead comparing expert musicians to non-musicians to ascertain possible effects of long term musical training. A study by Helmbold and colleagues (2005), for example, compared 70 musicians to non-musicians on psychometric assessments of intelligence and general mental abilities, and found that musicians performed better on two tasks: flexibility of closure (detecting single elements in complex objects), and perceptual speed (finding letters amongst digits). The authors speculated that the better performance on perceptual speed tasks could be explained by the demands of musical
training in requiring quick recognition of musical symbols or structures. Another study on temporal processing and detection found musicians to be better at discriminating time change to rhythmic patterns, but only when these were simple patterns (Jones & Yee, 1997). Furthermore, recent research using a line bisection task also showed faster reaction times and fewer errors in musicians when compared to controls (Patston, Hogg, & Tippett, 2007).

Recent research has extended findings of better performance at temporal order judgment (TOJ) tasks to both video game players (West, Stevens, Pun, & Pratt, 2008) as well as musical conductors (Hodges, Hairston, & Burdette, 2005). Hodges et al. (2005) examined the effects of musical expertise by comparing conductors to age and education matched controls. Overall, conductors were found to have better pitch discrimination skills as well as shorter auditory temporal thresholds than controls. Specifically, in an auditory TOJ task, conductors required less time between two sounds to correctly discriminate which one had occurred first. Interestingly however, no such differences in performance were seen when comparing musicians and controls on an analogous visual TOJ task.

Collectively, these and other findings suggest that enhanced aspects of temporal processing in expert musicians are at the very least correlated with their extensive training. In the present investigation we extend these findings by measuring three aspects of attention; spatial, temporal, and overall capacity. While previous findings would suggest enhancements in temporal perception (see Hodges, et al., 2005), it is unknown how general musical training (non-music-conductors) might affect overall attentional resources and whether spatial attention would likewise be modulated (i.e., analogous to findings observed with expert video game players in the auditory modality; see Donohue, Woldorf, & Mitroff, 2010). Lastly, it is worth noting that both experiments in this study were conducted in the visual modality, and given that music is largely an auditory and temporal task, any potential enhancement in other sensory modalities could therefore suggest concomitant crossmodal enhancements.

**Experiment 1**

In Exp. 1 we used a visual temporal order judgment (TOJ) task with exogenous and endogenous cues to measure both temporal processing and spatial attention. The TOJ task allows the calculation of the just noticeable difference (JND) and the point of subjective simultaneity (PSS). The JND refers to the smallest amount of time needed to separate both stimuli for an observer to be able to correctly identify the order of presentation (i.e., a measure of temporal perception). The PSS reflects the degree to which a spatial cue (peripheral or central) directs attention, thereby requiring the uncued side to be presented in advance of the cued side for simultaneity to be perceived (i.e., a measure of spatial attention).

Using cues in a TOJ task also creates a ‘prior entry’ effect, which has been the subject of many experiments (Shore, Spence, & Klein, 2001; C. Spence, Shore, & Klein, 2001; Zampini, Shore, & Spence, 2005). It is premised on the idea that temporal perception is influenced by attention, and thus attended stimuli are perceived prior to unattended stimuli. Alerting participants to a particular side by using an exogenous (peripheral) cue can create this prior entry effect. Exogenous orienting can occur from any stimulus that causes a reflexive, automatic, or bottom-up orienting of attention (e.g., bright flashes, loud sounds, etc.) that immediately captures attention. By using such a cue in the TOJ task prior to the onset of the first stimuli, participants’ attention will be directed to the cued side. If both left and right stimuli are then presented simultaneously, the effect will be that the cued side is perceived as having occurred first. Thus, when the PSS is calculated for the task, it is observed as being shifted towards the cued side (Shore, et al., 2001). An analogous effect would occur for endogenous (central arrow) cues, however in this case the observer has more volitional control over orienting effects (for comparison see Schmidt, 2000). Specifically, we are interested in seeing whether patterns of perceptual effects will be different in musicians due to the possible modulation of spatial processing. For instance, if musical training does lead to enhanced perception, then the JND should be smaller for musicians when compared with non-musicians. Additionally, PSS values might change, potentially reflecting a lesser likelihood to be distracted by exogenous or endogenous cues (i.e., PSS values would be smaller for musicians).

**Methods**

**Participants** Eight trained musicians (mean age = 21±2, 4 females) were recruited from various music performance classes at the University of Hawaii at Manoa. They had 11 years of musical training and practiced 9 hours per week on average. A variety of focus instruments were reported, including woodwind, stringed, piano, and voice. Care was taken to ensure that none of the participants reported extensive experience with video games. An additional 10 control participants were recruited (mean age = 21±5, 6 females), all of which had no significant musical training or video game experience.

**Materials** Visual stimuli were presented on a 20” (60Hz) Intel Core2Duo iMac using DMDX software (Forster & Forster, 2003). Observers sat approximately 60 cm from the display. Vertical and horizontal lines subtended 0.9° within the placeholder squares (1.4° wide) 4° from fixation (see Figure 1). Exogenous cues were created by thickening placeholder squares to 4 pixels, whereas endogenous cues consisted of a central arrow (both lasting 45 ms).

**Procedure** Throughout each trial a fixation cross flanked by two placeholders would remain on the display (see Figure 1 for durations). Both left and right placeholders were equally likely to be cued, after which a target (horizontal or vertical line) would appear (equiprobably) in one of the placeholders (left or right, also equiprobably) for a specified stimulus onset asynchrony (SOA) interval, followed by the
other stimuli in the other place holder. Participants then made an unseeded forced choice response on the keyboard to indicate either “horizontal” or “vertical” first responses. An adaptation of Stelmach and Herdman’s (1991) step-function procedure was used to determine the SOAs for each trial. Each trial began with an SOA of 267 ms. Depending on whether a correct or incorrect response was made, the SOA would respectively increase or decrease (by 16.7 ms) on the next trial. The experiment terminated after a total of 14 correct/incorrect reversals occurred.

Bonferroni corrected post hoc tests revealed larger PSS scores for exogenous than endogenous cues overall (63.0 ms vs. 13.9 ms, p < .001; see Figure 3).

![Exogenous condition diagram](image1)

![Endogenous condition diagram](image2)

**Figure 1:** Time-course representation of Experiment 1.

The exogenous and endogenous conditions were presented separately and counterbalanced. Onscreen instructions and repeatable practice trials with feedback were also given to each participant before each experimental block.

**Results**

For the analyses, data from each participant were separated into horizontally or vertically cued trials (see Figure 2). A logistic model was then fitted to each cue type for each participant. Two measures were then calculated for each individual. First, the PSS was interpolated from the model for SOAs corresponding to the 50% proportion for horizontal first responses. Secondly, the JND was calculated by first interpolating the SOAs corresponding to .75 and .25 proportions, and then halving the distance between these SOAs. One musician was excluded from the analysis due to large error rates, and one control participant was excluded due to non-convergence of the fit algorithm.

Separate mixed ANOVAs were performed on PSS and JND scores. Using within subject factors (2) of PSS for exogenous and endogenous cues revealed a highly significant main effect of cue type \( F(1,14) = 21.0, p < .001 \), but no significant effects for participant type or interaction (both \( F(1,14) < 1, ns \)), indicating no substantial differences in PSS patterns across musicians and controls.

![Figure 2](image3)

**Figure 2:** Proportion of “horizontal first” responses as a function of stimulus onset asynchrony (SOA). Note the larger gap between horizontal and vertical curves for exogenous trials compared to endogenous (reflecting larger effects on PSSs), and also the steeper slopes for musicians (reflecting lower JNDs).

![Figure 3](image4)

**Figure 3:** Mean PSS and JND scores (and SE) for musicians and controls.

The same ANOVA conducted for JND scores revealed a significant main effect of participant type \( F(1,14) = 14.9, p < .01 \), but no effects of cue type \( F(1,14) < 1, ns \) or interaction \( F(1,14) = 1.03, p > .1 \), suggesting different JND patterns between musicians and controls, but similar patterns of scores across exogenous and endogenous conditions.
conditions. Planned t-tests comparing musicians and controls confirmed lower JND scores for both exogenous (36 vs. 56 ms; t(11.7) = 1.8, p < .05) and endogenous cues (36 vs. 78 ms; t(9.5) = 3.0, p < .01).

Discussion
The main finding of Experiment 1 was the improved temporal processing of musicians, evidenced by smaller JND scores which reflect that less time was needed to separate the stimuli for musicians to still be accurate, regardless of whether it was an exogenous or endogenous cue. Additionally, the PSS scores were not different across groups, suggesting that spatial attention was similarly captured, although exogenous PSS scores for both groups were significantly larger than their endogenous scores, demonstrating the automatic and stronger effects of exogenous cues. In Experiment 2 we expand on these findings by exploring the attentional capacity of musicians.

Experiment 2
Considering the larger capture by exogenous cues than endogenous cues for all participants as illustrated in Experiment 1, automatic capture of attention may be expected due to the simplicity of the required task. Although previous experiments have also shown such effects, recent findings suggest that the effects of exogenous orienting may be lessened, or eliminated, under certain circumstances.

Using a paradigm involving both a demanding central task and a cued peripheral target detection task, Santangelo and colleagues have shown that exogenous orienting does not capture attention in a mandatory fashion (see Santangelo, Olivetti Belardinelli, & Spence, 2007; Santangelo & Spence, 2007, 2008). That is, when one’s attention is engaged in performing a perceptually or attentionally demanding task, the automatic effects of exogenous cues have been shown to disappear. We adapted this task and required participants to respond to a demanding central digit detection task, while at the same time respond to orthogonally cued peripheral targets (see Santangelo, et al., 2007; Santangelo & Spence, 2007). If musicians do have increased attentional capacity, then they may continue to show cueing effects when compared with controls, despite the difficult central task.

Methods
Participants The same eight musicians and ten controls from Experiment 1 also took part in Experiment 2.

Materials A rapid serial visual presentation (RSVP) stream was constructed from randomly chosen non-repeated letters (11 selected from set of 17: B, C, D, E, F, J, K, L, M, N, P, R, S, T, Y, X, Z) each presented for 100 ms with 16.7 ms of blank screen separating each letter. For digit detection trials, numbers were selected from a set of six: 2, 3, 4, 5, 6, 9. Visual targets were black circles ( subtending 2°) and cues were black rectangles (2.5° x 1.7°; see Figure 4).

Procedure Participants were required to monitor the RSVP stream presented in the center of the display, and to respond to the occurrence of either a numerical digit within the stream or an orthogonally cued spatial target. Responses were made using one of three keys following detection of either 1) a number, 2) an upward spatial target, or 3) a downward spatial target. Numbers occurred within the stream on 67% of the trials (majority of the time), whereas on the rest of the trials no numbers occurred, and instead visual targets occurred in one of four corners concurrent to display of the letter stream.

Each trial consisted of a fixation cross (1000 ms) followed by the RSVP stream of 11 items. On digit detection trials, the numbers randomly occurred in either the third, sixth, or ninth position in the stream. A spatial cue was also presented on each trial (for 100 ms, identical to item duration), occurring in the third or sixth position on either the right or left side of the display equiprobably. When spatial targets occurred, they appeared two positions after the cue (5th or 8th position). Half of the spatial targets were cued and the other half non-cued. Each task consisted of 196 randomized trials counterbalanced with the digit, target, cue combinations, and trial repetitions. Participants were instructed to respond as soon as targets were detected.

Results
One control participant’s data was excluded from the analyses due to a high error rate exceeding 15%, all other participants’ error rates were below 10%.

A mixed ANOVA on the RT scores with task type (2) of digit detection and target discrimination revealed that responses were faster for digit detection ($F (1,15) = 7.7, p = .01$), indicating that participants correctly prioritized digit detection over target discrimination. There were no main effects of participant type ($F (1,15) = 2.4, p = .1$) or interaction ($F (1,15) < 1, ns$). A separate ANOVA for cue
types (2) also revealed overall differences in cued and non-cued trials \(F(1,15) = .57, p = .03\), but with no effects for participant type \(F(1,15) = 1.7, p > .1\) or interaction \(F(1,15) = 1.5, p > .1\). Planned comparisons between cued trials and non-cued trials revealed that attentional capture from cues occurred only for musicians (495 vs. 510 ms; \(t(7) = 2.1, p < .05\), Cohen's \(d = 0.80\)), and not for controls (551 vs. 556 ms; \(t(8) = 1.1, p > .1\); see Figure 5). Furthermore, independent t-tests revealed that musicians responded faster than controls for both digit detection (473 vs. 519 ms; \(t(14) = 1.6, p = .06\), Cohen's \(d = 0.78\)) and cued target discrimination trials (495 vs. 551 ms; \(t(13) = 1.4, p = .08\), Cohen's \(d = 0.68\)), with these differences approaching significance.

![Figure 5: Mean scores (and SE) for the three types of tasks in Experiment 2.](image)

### Discussion

Results from control participants thus replicates Santangelo and colleagues’ findings that under high load conditions, exogenous orienting has been shown to disappear (Santangelo & Spence, 2008). Crucial to this study, however, exogenous cuing effects remained only for musicians, suggesting a possible increase in attentional resources that may spill over to process cues even under conditions of high load (Lavie 1995; for example with VGPs see Green & Bavelier, 2003).

### General Discussion

This preliminary study has both theoretical and practical relevance. To begin with, this is the first study of its kind exploring temporal and spatial mechanisms of visual attention and perception, as well as attentional capacity, in the same group of trained musicians. There are two main findings. First, the lower just noticeable differences scores for musicians in Exp. 1 suggests that temporal discrimination in musicians was significantly better than controls, regardless of cue type. Combined with the faster reaction times observed in Exp. 2, these findings are in line with other research showing increased perceptual speed and detection in trained musicians (Helmbold, et al., 2005; and also extending results beyond conductors used in Hodges, et al., 2005). Secondly, attentional capacity in musicians appeared to be larger than that of controls, to the extent that there was a significant difference between the processing of cued and non-cued items despite attentional resources being arguably exhausted by a concurrent task (i.e., as evidenced by no significant cuing effect in the control group).

Lastly, it should be noted that similar point of subjective simultaneity scores between musicians and controls in Exp. 1 suggests that musical experience did not significantly modulate spatial attention. While these results may differ from other expert populations (e.g., video game players), this may be due to the fact that musical training places a heavier emphasis on temporal processing for synchronicity of performance, rather than on the processing of rapidly presented peripheral events (i.e., as seen in video game play).

These behavioral findings are also supplemented by considerable evidence converging on greater neuroplasticity in musicians resulting in both functional and anatomical differences. These include for instance, increases of grey and white matter volume in specific sites of the left cerebellum, more pronounced cortical reorganizations for musically related motor activity, as well as larger evoked potentials for instrumental tones when compared to controls (Gaser & Schlaug, 2003; Münte, Altenmüller, & Jäncke, 2002). Whether or not musically related brain plasticity may be related to attentional mechanisms is speculative, but may suggest that structural changes due to repeated exposure and training with particular stimuli can indeed influence other areas of performance.

Another important point to note is the exclusive use of visual stimuli. Incidentally, our results may be seen to support a supramodal account of attention where a single attentional reservoir is used by multiple sensory modalities (e.g., Farah, Wong, Monheit, & Morrow, 1989; for example of audio enhancements with VGPs see Donohue, et al., 2010). This account would suggest that training in an environment dealing largely with auditory stimuli (such as music) may actually lead to accompanying enhancements in visual attention (although see Hodges, et al., 2005).

As a preliminary study, this research raises several questions relevant to future investigations. What are the specific training parameters that lead to attentional plasticity and can these parameters be used in music education or to enhance training in other domains? It should be noted that musical training might not be the sole reason for enhanced performance. It is possible that individuals with an already superior attentional system are drawn to and predisposed to expertise in activities such as music. The objective of future studies should be to incorporate training conditions to more directly ascertain the cognitive effects of musical training.

### References


