Title
Multimodal Temporal Perception in Musicians: Evidence for Both Segregated and Supramodal Attentional Systems?

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 34(34)

ISSN
1069-7977

Authors
Lim, Ahnate
Sinnett, Scott

Publication Date
2012

Peer reviewed

Multimodal Temporal Perception in Musicians: Evidence for Both Segregated and Supramodal Attentional Systems?

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

Although musical training has been correlated with modulations of early perceptual and attentional processes, the majority of investigations neglect the possibility of cross modality enhancements. We investigated the effects of musical training by measuring spatial and temporal attention in a temporal order judgment task in auditory, visual, and crossmodal conditions with and without non-predictive cues. In Experiment 1, musicians had lower detection thresholds when compared to controls in all conditions (marginal in auditory). Experiment 2 showed mixed findings, with musicians demonstrating reduced capture from visual cues on the visual task compared to controls, and lower detection thresholds on the auditory task with visual cues. Adding spatial cues to the temporal order judgment tasks increased temporal thresholds for both groups, but only when they occurred within the same modality as the task, and not when presented in a different modality. The findings support both supramodal and segregated accounts of attentional resources.

Keywords: attention; perception; musicians; temporal order judgment; multisensory; visual; auditory; crossmodal

Introduction

The human attentional system is impressively competent at processing information, considering how the efficiency and selectivity of attention facilitates perception and goal directed behavior amidst a constant plethora of stimuli. Interestingly, the neurological underpinnings of attention may change under certain conditions. This "plasticity" has been associated with compensations for losses in one sensory modality with enhancements in another modality (Röder et al., 1999). Furthermore, improved behavioral performance may also occur as a side effect of specific activities or hobbies such as video game playing (Granek, Gorbet, & Sergio, 2010; Green, Li, & Bavelier, 2010, but see Boot, Blakely, & Simons, 2011) and musical training (Hodges, Hairston, & Burdette, 2005; Lim & Sinnett, 2011).

Although the topic of non-musical cognitive benefits (e.g., mathematics, spatial-reasoning and linguistics) occurring as a result of musical training has been the focus of much research (for a summary, see Rauscher, 2003), there has been less emphasis on the effects of precise mechanisms of attention and perception. There is evidence from numerous studies conducted with musicians suggesting greater neuroplasticity when compared with non-musicians (see for example, Gaser & Schlaug, 2003; Münte, Altenmüller, & Jäncke, 2002). Although it should be noted that these brain differences could equally be attributed to a predisposition that leads people to become musicians, rather than any specific training related enhancement. Even so, behavioral evidence from studies comparing musicians to non-musicians demonstrates improved perceptual abilities on various tasks in different modalities. These have included visual perceptual speed and discrimination (Helmbold, Rammsayer, & Altenmüller, 2005; Patston, Hogg, & Tippett, 2007) as well as auditory temporal discrimination (Hodges et al., 2005; Jones & Yee, 1997).

These studies also highlight an interesting possibility of training effects on attention: crossmodal enhancements (e.g., visual enhancements after auditory training). Some authors have suggested that the attentional system operates in a supramodal fashion, with all senses having access to a single reservoir of attentional resources (see Farah, Wong, Monheit, & Morrow, 1989; Pavani, Husain, Ládavas, & Driver, 2004; but see also Sinnett, Costa, & Soto-Faraco, 2006; C Spence & Driver, 1996; Wickens, 1984, for examples of a segregated attentional system). Thus, by testing musicians for enhanced attentional and perceptual capabilities in the visual modality, we can indirectly assess whether training in one sense (i.e., auditory musical training) leads to performance enhancements in another. This would provide support for a supramodal attentional system, and would closely align with recent investigations involving video game players, where auditory enhancements were observed despite the training being mostly visual based (Donohue, Woldorff, & Mitroff, 2010; Green, Pouget, & Bavelier, 2010).

The temporal order judgment (TOJ) task is an ideal tool to assess temporal processing differences between musicians and non-musicians. More importantly, the TOJ task can be presented under both unimodal and crossmodal conditions. The task requires participants to determine the correct order of subsequently presented stimuli, and allows for two measures of perceptual processing to be calculated: the just noticeable difference (JND), and the point of subjective simultaneity (PSS). The JND is a measure of the resolution or threshold of temporal discrimination, while the PSS is the time in which one stimulus can be presented before the other such that they are still perceived as occurring simultaneously (e.g., in a crossmodal task, it can indicate whether auditory or visual stimuli must be presented first for them to be perceived as simultaneous).

Humans are generally proficient at temporal discrimination. In studies examining within and cross-modal (visual, auditory, tactile) TOJs, Hirsh and Sherrick Jr. (1961) found that participants could discriminate temporal order between stimuli (JND) when presented as quickly as 20ms apart. Crucially, in crossmodal tasks (i.e., audiovisual presentations) the visual stimuli had to lead auditory stimuli by approximately 40-80ms for participants to perceive them as being presented simultaneously (PSS; see also Zampini, Shore, & Spence, 2003). Furthermore, research suggests that the resolution of temporal acuity is better in the auditory modality than in vision or touch (Chen & Yeh, 2009).

Given the efficacy at which humans can discriminate temporally, it is worth noting that significant gains or losses in TOJ performance can occur as a result of brain injury (Sinnett, Juncadella, Rafal, Azanon, & Soto-Faraco, 2007) or training (Donohue et al., 2010). This suggests that temporal perception is perhaps dependent on attentional mechanisms, and not purely a sensory-based process. Aside from studies showing *enhancements* on TOJs in video game players (e.g., Donohue et al., 2010), recent research has extended findings of better performance (lower JNDs) in the auditory modality to musical-conductors (Hodges et al., 2005) and in the visual modality to performing-musicians (Lim & Sinnett, 2011). Nevertheless, it is worth noting that research with musicians has not yet looked at crossmodal TOJs and the possibility that musical training, mostly auditory in nature, might have effects on visual TOJs.

Expert musicians were compared with non-musicians on a series of TOJ tasks that were presented under unimodal (visual or auditory), or crossmodal conditions. Given evidence from previous research, we expected to see lower JND scores for musicians when compared with controls in all conditions, but did not expect any differences in PSS scores.

Experiment 1

Participants

Twenty musicians (age = 28 ± 12; 5 females) were recruited from the music department at the University of Hawaii at Manoa, local music studios, and through flyers. Musicians were required to have at least three years of formal training in music, and to have a regular practice schedule of at least six hours/week over the past six months. Control participants (n = 20; age = 22 ± 5, 16 females. Note, pooled t-test comparisons showed no differences between males and females for any of the conditions, all p > .1) were recruited from undergraduate courses, and had little or no training in music. All participants received either $10 or course credit for their participation. Ethical approval was obtained from the University's Committee on Human Subjects.

Stimuli and Apparatus

The basic TOJ task involves presenting participants with two stimuli separated by variable time intervals, referred to as the stimulus onset asynchrony (SOA). The SOA length is manipulated to increasing or decreasing intervals that correspondingly makes the task easier or harder. A staircase approach was used in this experiment to adjust SOAs (see Stelmach & Herdman, 1991). The SOA started at 167ms and, for each successive trial, either decreased or increased (by 16.7ms) in a stepwise manner dependent on whether the participant answered the previous trial correctly. As the experiment progressed, each trial's SOA decreased making the order of occurrence difficult to determine. It can be inferred then, that as time progresses, changes in stepwise direction (up and down) will increase, reflecting increasing uncertainty in the participant. The task terminates once a total of twelve turning points have occurred.

Stimuli were presented on a 21" iMac using Bootcamp and DMDX software. Participants were seated at an eye to monitor distance of approximately 60cm. Prior to each trial a fixation-cross (0.5°) flanked by two square placeholders (1.4°) on the left and right was presented (see Figure 1). Stimuli for the visual task were horizontal and vertical lines (0.9°) and occurred centrally within the placeholders. For the auditory stimuli, processed samples of a dog and crow sound (350ms) were used (played at approximately 75db). In the crossmodal condition, the visual stimulus consisted of a black square (0.9°) within the placeholder, whereas the auditory stimulus was 50ms of white noise.

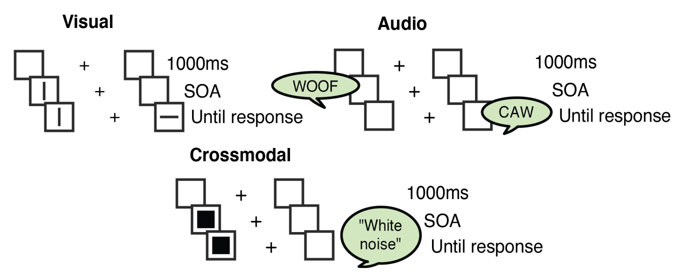

Figure 1. Stimuli for the three TOJ tasks in Experiment 1.

Procedure

In all three conditions participants made unspeeded responses signaling which stimulus they believed had appeared first using one of two keyboard buttons. Onscreen instructions were presented first, followed by a short sequence of practice trials that included accuracy feedback. Presentation side (i.e., left or right) and stimuli order (e.g., horizontal or vertical line first) were randomized, as was the order of experimental conditions (e.g., audio, visual, crossmodal) for each participant.

Results

Calculations of the JND and PSS were based on approaches used by previous studies (C. Spence, Baddeley, Zampini, James, & Shore, 2003; Stelmach & Herdman, 1991). Data from musicians and controls were pooled into separate groups. The average ratio of responses "horizontal line first" (e.g., for visual condition; for auditory condition ratio of crow sounds were used, etc.) was then plotted as a function of the time in which the horizontal line preceded the vertical line. Data was then fit using a logistic function:

$$f(x, a, b) = \frac{1}{1 + \exp(-(x - a)/b)}$$

which was then used to obtain the JND and PSS estimates (similar to C. Spence et al., 2003). The PSS corresponds to parameter a, and is usually expected to fall at 0ms (or close) in unimodal conditions, as there is no reason to assume that a particular visual (or auditory) stimulus would be preferred over the other. It is more informative in the crossmodal condition as any shift would indicate whether auditory or visual events must precede the other for subjective simultaneity to be perceived. The JND relates to parameter b, which is adjusted to obtain the 75% JND as follow:

$$\mathrm{JND}_{75} = \ln 3 \cdot b$$

Given that data was pooled within each group (musicians and controls), confidence intervals (95%) for each group's estimates and comparison p-values were calculated using a parametric bootstrap method with 999 replications (Efron & Tibshirani, 1993; for similar use of the bootstrap, see Azañón & Soto-Faraco, 2007).

Auditory condition Differences between musicians and controls were non-significant for PSS (2ms, CI = -8 to 12ms; vs. 9ms, CI = 1 to 16ms; p = 0.31; respectively) and approaching significance for JND scores (43ms, CI = 34 to 53ms; and 56ms, CI = 45 to 68ms; p = 0.07).

Visual condition The average PSS score for musicians' was significantly lower than controls by 10ms (-4ms, CI = -9 to 2ms; vs. -14ms, CI = -22 to 5ms; p = 0.037; respectively), with negative PSS values indicating a possible bias in responses towards horizontal lines. The average JND score for musicians' was also significantly lower than controls by 18ms (29ms, CI = 23 to 35ms; vs. 47ms, CI = 37 to 56ms; p = 0.006).

Crossmodal condition Differences between musicians and controls were non-significant for PSS (-43ms, CI = -60 to -25ms; and -63ms, CI = -93 to -30ms; p = 0.261; respectively). It is worth noting that the negative PSS results indicate a bias in response towards the auditory modality (visual stimuli needed to be presented prior to auditory stimuli for simultaneity to be perceived). Musicians' average JND score was significantly lower than controls by 59ms (104ms, CI = 80 to 127ms; and 163ms, CI = 112 to 207ms; p = 0.021).

Discussion

There are two important findings that merit discussion. First, with the exception of the visual condition, no differences were observed for PSS between musicians and non-musicians. In the visual condition it is possible that there was a small bias towards horizontal lines for musicians, but note that performance hovered around zero as expected. The largest PSS differences were seen in the crossmodal condition. Specifically, visual stimuli had to precede auditory stimuli for both musicians and controls (by 43 and 63ms, see Figure 2) for them to be perceived as occurring simultaneously (Hirsh & Sherrick Jr, 1961; Zampini et al., 2003).

Secondly, the temporal threshold for musicians was significantly lower in all conditions (visual, auditory, and crossmodal), although it should be noted that only marginal significance was observed in the auditory condition (p = 0.07). Given that musical training is largely auditory in nature, a more robust difference in auditory JND scores was expected. It is possible that the "realistic" auditory stimuli used in our experiment may be more difficult than simpler tones, and therefore any effect might be somewhat masked. Furthermore, it might be possible that as the sounds were non-tonal, musicians may not have had a distinct advantage. Lastly, given that humans discriminate temporal events better in the auditory modality when compared to the visual modality, it is possible that performance was similar due to a ceiling effect.

Lastly, and also supported by Zampini et al. (2003), JND scores for the crossmodal condition increased nearly three-fold when compared to unimodal conditions, demonstrating that the task was more difficult.

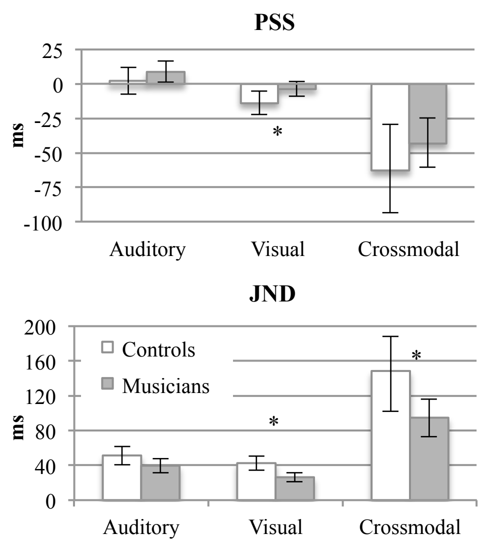

Figure 2. PSS and JND scores for Experiment 1. For PSS scores, positive indicates stimuli crow/horizontal/audio, and negative indicates dog/vertical/visual stimuli appearing first. Error bars indicate 95% CIs.

Experiment 2

Spatial cues can also be incorporated into the TOJ tasks, allowing for a measure of how attention is oriented and captured. The presentation of exogenous cues prior to stimuli onset in a TOJ task creates a 'prior entry' effect, where attention is directed towards the cued side and subsequently affects performance on the task, regardless of whether or not the cue is predictive of location (see Shore, Spence, & Klein, 2001)

Exogenous orienting can occur from any stimulus that causes a reflexive or automatic capture of attention (e.g., bright flashes, loud sounds, etc.). By presenting an exogenous cue in the TOJ task prior to the onset of stimuli, the cued side will be perceived as having occurred first. The PSS score then indicates how much in advance the uncued side must be presented before the cued for simultaneity to be perceived (see Shore et al., 2001). Thus, Experiment 2 included all of the unimodal conditions of Experiment 1 with the addition of within and crossmodal cues to determine whether spatial attention would differ between musicians and non-musicians. If musical training can improve spatial and temporal processing, it would be expected that musicians should have a smaller orienting effect, which would be manifested in lower PSS and JND scores than controls across all conditions. This would be indicative of improved temporal processing (JND) and less influence from peripheral distraction (smaller PSS). Furthermore, to our awareness this would be the first time multimodal cued TOJ tasks were conducted on musicians.

Participants, Apparatus, and Procedure

The same participants from Experiment 1 also took part in Exp. 2 (all conditions from both experiments were interleaved and fully randomized). The discussion of the experiments is separated here for ease of understanding.

Stimuli and procedure were identical to those in Exp. 1, except for the addition of exogenous non-predictive cues in all conditions. In the visual condition, the cue was created by thickening the placeholder box of the respective side to a thickness of 4 pixels for 45ms. In the auditory condition, the cue was a laterally presented 500Hz sine wave lasting 45ms. The crossmodal condition consisted of two tasks: the first was an auditory TOJ task with visual cues, while the second was a visual TOJ task with auditory cues. All cues were randomly determined and had an equal chance of validly or invalidly cuing the target stimuli.

Results

The JND and PSS scores were calculated using similar methods as in Exp. 1, by pooling musicians and control participants into two separate groups. For each of the four conditions, data from the two groups were fit to a weighted logistic function according to which stimuli was cued (e.g., horizontal/vertical bar, dog/crow sound, etc). The overall PSS value for each condition was computed as half the distance between each of the PSS values for the two curves. The average of the two JND values for each curve was used as the overall JND score. This approach essentially calculates the PSS for each type of stimulus cued, and averages the effect (see Shore et al., 2001). In order to gauge the influence of the cue, the two fitted curves were compared against one another. Logically, if the two curves were to map out on top of one another then the average would be 0 (PSS), as would be expected if the cue did not have any effect (assuming no bias for one stimulus type or the other). Thus, the larger the difference between the logistic fits for each cue, the larger the PSS, and by extension the greater effect that the cues had in general. This can similarly be applied to the calculation of JND, although as the slope, the JND scores are expected to be similar for each stimulus type.

Unimodal Cues *Auditory condition:* The magnitude of the PSS shifts was not significantly different between musicians and controls (23ms, CI = 12 to 37ms; vs. 29ms, CI = 17 to 42ms; p = 0.259; respectively, see Figure 3). Similarly, there were no differences in JND scores between the two groups (92ms, CI = 77 to 110ms; vs. 109ms, CI = 93 to 125ms; p = 0.106). *Visual condition:* The magnitude of PSS shifts was significantly lower for musicians than controls by 29ms (30ms, CI = 10 to 46ms; vs. 59ms, CI = 48 to 71ms; p = 0.023). On the other hand, JND scores for both groups were not significantly different (80ms, CI = 63 to 93ms; vs. 84ms, CI = 75 to 96ms; p = 0.29).

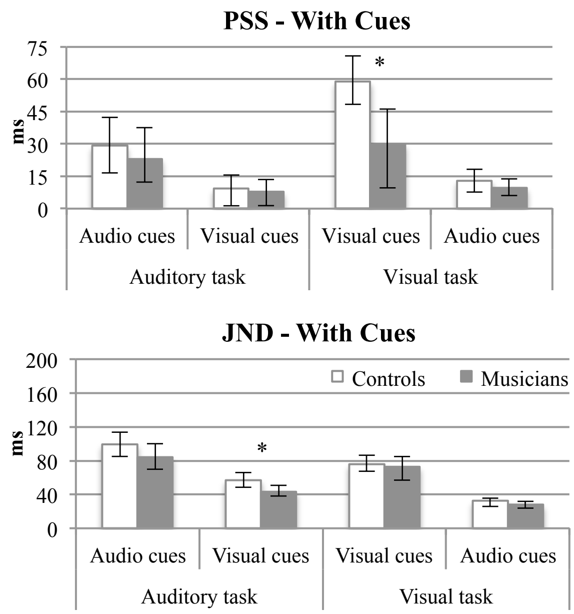

Figure 3. PSS and JND scores for Experiment 2. Asterisks indicate significant (*p* < .05) between group differences. Error bars indicate 95% CIs.

Crossmodal Cues *Auditory TOJ with visual cues:* The magnitude of the PSS shifts did not significantly differ between musicians and controls (8ms, CI = 1 to 14ms; vs. 9ms, CI = 1 to 15ms; p = 0.39; respectively). On the other hand, JND scores were significantly lower by 16ms for musicians compared to controls (47ms, CI = 42 to 56ms; vs. 63ms, CI = 53 to 73ms; p = 0.014). *Visual TOJ with auditory cues:* The magnitude of the PSS shifts did not significantly differ between musicians and controls (10ms, CI = 6 to 14ms; vs. 13ms, CI = 8 to 18ms; p = 0.19). Similarly, there were no differences in JND scores between the two groups (31ms, CI = 26 to 35ms; vs. 35ms, CI = 29 to 38ms; p = 0.088).

Cross experiment comparisons: Further understanding of the cuing effects can be determined by comparing the results from the cued tasks in Experiment 2 to the no-cue unimodal tasks (auditory and visual) of Experiment 1. When doing so, JND differed for unimodal conditions but not for crossmodal conditions. That is, the additional cues in Experiment 2 made the unimodal tasks harder for both musicians and non-musicians, as evidenced by longer temporal thresholds (JND) in both the auditory and visual modalities (all $p < .01$). However, when the cues were presented in a separate modality (i.e., the crossmodal conditions of Exp. 2), JND scores were indistinguishable from the unimodal no cue conditions (Exp. 1) for both musicians and non-musicians (all $p > .05$). Collectively, this may suggest that a difficult unimodal task can be made easier when presented as a crossmodal task (Sinnett et al., 2006; Sinnett et al., 2007; C Spence & Driver, 1996).

Discussion

Robust findings from cross-experiment analyses broadly suggest that unimodal cues have detrimental effects on JND scores, whereas crossmodal cues do not. These results were similar for both musicians and controls. Excluding these cross experiment analyses, the only observed significant differences in Experiment 2 between musicians and controls were the lower PSS scores in the visual unimodal condition for musicians, and the lower JND score in the auditory-task/visual-cues condition for musicians. The lower PSS score indicates that musicians were captured less by the unimodal visual cues than non-musicians, while the lone JND difference seemingly suggests that crossmodal processing was easier for musicians, but only when judging temporal order for auditory targets that were cued visually.

General Discussion

There are a number of important findings. To begin with, performance differences between musicians and controls were mixed in the auditory condition (musicians did have significantly lower JND scores in the auditory-task/visual-cues condition of Experiment 2, as well as marginally lower JNDs in the auditory condition in Experiment 1 [$p = .07$], while the unimodal auditory condition of Experiment 2 was not significant). Thus, we do not see as strong a trend as Hodges et al. (2005), where auditory JND scores were significantly lower for musical conductors when compared to controls. This may be due to the use of different stimuli and experimental conditions. In the present experiment realistic sounds (dog and crow) were used, while auditory tones were used in Hodges et al.'s studies. Thus it is possible that pitch discrimination skills would not aid musicians in the auditory task used here. Furthermore, it is also possible that differences in auditory temporal processing may exist between conductors and performing musicians. It is worth noting however, that across all task types, JND scores for musicians were numerically lower than those for controls, although these differences were statistically significant in only four out of the seven conditions (Exp 1: auditory (marginal), visual, and crossmodal; Exp 2: Audio-task/visual cues).

Pertinent to the discussion is the tentative support for a supramodal account of attentional resources, supported by the fact that musicians outperformed controls on several non-auditory related tasks, including smaller capture from visual cues, and lower JNDs for visual and crossmodal conditions (without cues). That is, it appears that musical training might have lead to improved visual processing. Having said that, as musical training involves much exposure to auditory stimuli, it was reasonable to expect enhancements in the auditory modality, although this was not consistently observed. Enhancements in the visual modality however, could be attributed to 1) better attentional resources, and/or 2) concomitant training in the visual modality from reading music while at the same time listening to and playing music, etc. Since we cannot rule out the second possibility however, these results can only be seen as tentative support for a supramodal account, pending further investigation with specific training conditions. Interestingly however, the robust findings of Experiment 2 where crossmodal PSS and JND scores were in fact lower than their unimodal counterparts (all $p < .05$ and $p < .001$; respectively), may provide stronger evidence for the exact opposite viewpoint: that is, a segregation of attentional systems (e.g., Sinnett et al., 2006; Wickens, 1984). Nevertheless, the current set of data makes it difficult to arrive at a decisive claim on either side of the debate, and may suggest a two-part attentional system that operates with both segregated and supramodal capacities. Indeed, it is likely that many previous findings supporting one theoretical account or the other may indeed be constrained by the varying methodologies used.

The segregated account is supported by the novel finding that was observed across both musician and control groups regarding the selective deficits in JND for only unimodal cues and not crossmodal cues. That is, when a within modality cue was added to the task, JND scores increased significantly for both musicians and control participants. However, when the cues were presented across modalities (i.e., a visual cue and an auditory TOJ task, or vice versa), performance was significantly better, and in fact did not differ from the no-cue conditions. This possibly suggests that the threshold of temporal detection may be robust to crossmodal distraction, while at the same time be vulnerable to distractions within the same modality.

As the between group differences in the auditory task in Experiment 1 were only marginally significant, this may suggest that auditory temporal acuity is less amenable to improvement through training (at least for the stimuli and task conditions used here), and that concomitant training effects are perhaps more robust in the visual domain. Importantly, the visual enhancements observed in JND in Experiment 1 lend support to the idea that attentional allocation, and therefore the improvement through training, may not be constrained within particular sensory modalities, but instead distributed to multiple modalities. Nevertheless, an important criticism of studies that used "trained" populations such as musicians and video game players, is the extent to which observed differences in experimental settings can actually be attributed to prior training. Boot et al. (2011), for instance, claimed that participants are often aware of the purpose of the study as they are specifically recruited for their expertise, and that this awareness and potential motivational factor may very well influence performance. Unfortunately, our recruitment strategy for musicians did not allow us to keep them blind to the purpose of the study, and they may have been influenced by such knowledge. To this extent, our between group conclusions are largely speculative. Moreover, the nature of musical training in sighted individuals is in itself a multimodal experience, and further training studies would be better equipped to draw conclusions by controlling for the type of training each participant receives.

References

Azañón, E., & Soto-Faraco, S. (2007). Alleviating the crossed-hands deficit by seeing uncrossed rubber hands. *Experimental brain research, 182*(4), 537-548.

Boot, W. R., Blakely, D. P., & Simons, D. J. (2011). Do Action Video Games Improve Perception and Cognition? *Frontiers in Psychology, 2*.

Chen, K., & Yeh, S. (2009). Asymmetric cross-modal effects in time perception. *Acta psychologica, 130*(3), 225-234.

Donohue, S., Woldorff, M., & Mitroff, S. (2010). Video game players show more precise multisensory temporal processing abilities. *Attention, Perception, & Psychophysics, 72*(4), 1120.

Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap* (Vol. 57): Chapman & Hall/CRC.

Farah, M., Wong, A., Monheit, M., & Morrow, L. (1989). Parietal lobe mechanisms of spatial attention: modality-specific or supramodal? *Neuropsychologia, 27*(4), 461-470.

Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *Journal of Neuroscience, 23*(27), 9240.

Granek, J., Gorbet, D., & Sergio, L. (2010). Extensive video-game experience alters cortical networks for complex visuomotor transformations. *Cortex, 46*(9).

Green, C. S., Li, R., & Bavelier, D. (2010). Perceptual learning during action video game playing. *Topics in Cognitive Science, 2*(2), 202-216.

Green, C. S., Pouget, A., & Bavelier, D. (2010). Improved Probabilistic Inference as a General Learning Mechanism with Action Video Games. *Current Biology, 20*(17), 1573-1579.

Helmbold, N., Rammsayer, T., & Altenmüller, E. (2005). Differences in primary mental abilities between musicians and nonmusicians. *Journal of Individual Differences, 26*(2), 74-85.

Hirsh, I., & Sherrick Jr, C. (1961). Perceived order in different sense modalities. *Journal of Experimental Psychology, 62*(5), 423-432.

Hodges, D. A., Hairston, W. D., & Burdette, J. H. (2005). Aspects of multisensory perception: the integration of visual and auditory information in musical experiences. *Annals of the New York Academy of Sciences, 1060*, 175-185.

Jones, M., & Yee, W. (1997). Sensitivity to time change: The role of context and skill. *Journal of Experimental Psychology: Human Perception and Performance, 23*(3), 693.

Lim, A., & Sinnett, S. (2011). *Exploring Visual Attention in Musicians: Temporal, Spatial and Capacity Considerations*. Paper presented at the 33rd Annual Conference of the Cognitive Science Society.

Münte, T., Altenmüller, E., & Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. *Nature Reviews Neuroscience, 3*(6), 473-478.

Patston, L., Hogg, S., & Tippett, L. (2007). Attention in musicians is more bilateral than in non-musicians. *Laterality, 12*(3), 262.

Pavani, F., Husain, M., Ládavas, E., & Driver, J. (2004). Auditory Deficits in Visuospatial Neglect Patients. *Cortex, 40*(2), 347-365.

Rauscher, F. (2003). Can Music Instruction Affect Children's Cognitive Development? *ERIC Digests* Retrieved 10 October, 2011, from http://www.ericdigests.org/2004-3/cognitive.html

Röder, B., Teder-Sälejärvi, W., Sterr, A., Rösler, F., Hillyard, S., & Neville, H. (1999). Improved auditory spatial tuning in blind humans. *Nature, 400*(6740), 162-166.

Shore, D. I., Spence, C., & Klein, R. M. (2001). Visual prior entry. *Psychological Science, 12*(3), 205-212.

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

Sinnett, S., Juncadella, M., Rafal, R., Azanon, E., & Soto-Faraco, S. (2007). A dissociation between visual and auditory hemi-inattention: Evidence from temporal order judgements. *Neuropsychologia, 45*(3), 552-560.

Spence, C., Baddeley, R., Zampini, M., James, R., & Shore, D. I. (2003). Multisensory temporal order judgments: When two locations are better than one. *Perception & psychophysics, 65*(2), 318.

Spence, C., & Driver, J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology: Human Perception and Performance, 22*(4), 1005-1030.

Stelmach, L. B., & Herdman, C. M. (1991). Directed attention and perception of temporal order. *Journal of Experimental Psychology: Human Perception and Performance, 17*(2), 539-550.

Wickens, C. D. (1984). Processing Resources in Attention. In R. Parasuraman & R. Davies (Eds.), *Varieties of attention* (pp. 63). New York: Academic Press.

Zampini, M., Shore, D., & Spence, C. (2003). Audiovisual temporal order judgments. *Experimental brain research, 152*(2), 198-210.