

A MULTIMODAL COMPARISON OF TEMPORAL ORDER JUDGMENTS IN MUSICIANS
AND NON-MUSICIANS

A THESIS SUBMITTED TO THE GRADUATE DIVISION OF THE
UNIVERSITY OF HAWAII AT MANOA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS
IN
PSYCHOLOGY

DECEMBER 2011

By
Ahnate Lim

Thesis Committee:
Scott Sinnett, Chairperson
Leonidas A.A. Doulas
Patricia Couvillon

Abstract

Considerable evidence converges on the plasticity of attention and the possibility that it can be modulated through regular training. Musical training, for instance, has been correlated with modulations of early perceptual and attentional processes. However, the extent to which musical training can modulate mechanisms involved in processing information (i.e., perception and attention) is still widely unknown, with the majority of investigations neglecting the possibility of within and across sensory modality enhancements. This is theoretically important, as a multisensory approach to understanding how training can modulate not only performance within a sense, but also in other senses (i.e., audition and crossmodal interactions), aligns better with real world processing. Furthermore, if training in one sensory modality can lead to concomitant enhancements in separate modalities, then this could be taken as evidence of a supramodal attentional system. Accordingly, the research conducted here further investigated the effects of musical training using tasks designed to measure spatial and temporal attention in auditory, visual, and crossmodal conditions. This is the first investigation to assess both spatial and temporal attention within and across sensory modalities using the same expert participants. Interestingly, results were mixed with expert musicians showing only marginal enhancements when compared with controls in auditory tasks, but with more significant enhancements for temporal and spatial processing in the visual modality, as well as overall lower detection thresholds for visual and crossmodal tasks. Furthermore, adding spatial cues to a temporal order judgment task increased temporal thresholds (i.e., made the task harder), but only when they occurred within the same modality as the task, as opposed to when presented in a different modality. Implications for design and future training studies are discussed.

Table of Contents

Abstract	i
List of Figures	iv
List of Abbreviations	v
Introduction.....	1
<i>Temporal perception</i>	6
Experiment 1: Measuring temporal perception.....	11
<i>Participants</i>	11
<i>Stimuli and apparatus</i>	12
<i>Procedure</i>	14
<i>Results</i>	14
<i>Discussion</i>	19
Experiment 2: Measuring temporal and spatial perception	23
<i>Participants</i>	23
<i>Stimuli and apparatus</i>	23
<i>Procedure</i>	24
<i>Results</i>	24
Unimodal cues.....	26
Crossmodal cues	28
Cross-experiment comparisons	30
<i>Discussion</i>	32
General Discussion	33

References.....	38
Appendix 1.....	43

List of Figures

Figure 1. Stimuli for the three TOJ tasks in Experiment 1	13
Figure 2. Pooled data for the auditory condition in Experiment 1.....	17
Figure 3. Pooled data for the visual condition in Experiment 1	18
Figure 4. Pooled data for the crossmodal condition in Experiment 1.....	19
Figure 5. PSS and JND scores for Experiment 1	21
Figure 6. Stimuli for the four TOJ tasks in Experiment 2	24
Figure 7. An example of a typical response curve from Experiment 2	26
Figure 8. Pooled data for the auditory-task/auditory-cues condition in Experiment 2.....	27
Figure 9. Pooled data for the visual-task/visual-cues condition in Experiment 2	28
Figure 10. Pooled data for the auditory-task/visual-cues condition in Experiment 2.....	29
Figure 11. Pooled data for the visual-task/auditory-cues condition in Experiment 2.....	30
Figure 12. PSS and JND scores for Experiment 2.....	33
Figure 13. Combined table of findings for Experiment 1 and 2	34

List of Abbreviations

ANT	Attentional Network Test
CI	Confidence Interval
ERP	Event-Related Potential
fMRI	Functional Magnetic Resonance Imaging
JND	Just Noticeable Difference
ms	Milliseconds
PSS	Point of Subjective Simultaneity
RSVP	Rapid Serial Visual Presentation
SOA	Stimulus Onset Asynchrony
TOJ	Temporal Order Judgment
VGP	Video Game Players

Introduction

When considering the scope and requirements of day-to-day activities, the human attentional system is an impressively competent information processing system even by modern standards of computational technology. That is, at any given point in time our sensory system is inundated with a plethora of stimuli, whereupon the efficiency and selectivity of attention allows for effective navigation and decision making from moment to moment, permitting goal-oriented behavior to be accomplished.

Interestingly, the neurological foundations of attention can sometimes change under certain conditions. Indeed, a more plastic understanding of brain function has become widely accepted, as opposed to a modular notion of brain functioning (Kolb, Gibb, & Robinson, 2003). For instance, this “plasticity” of the brain has been associated with slight compensations when a sensory modality is lost (Shimojo & Shams, 2001), or in adapting to conditions such as age, disease, stress, and even addiction (Kolb, et al., 2003). In a behavioral and neurophysiological study Röder et al. (1999), for example, compared congenitally blind adults with blindfolded controls on a central and peripheral sound localization task. Behavioral results showed blind participants to possess greater peripheral spatial localization abilities, dovetailing with electrophysiological (ERP) results that revealed more finely tuned early spatial attentional mechanisms. This study, in conjunction with others conducted with deaf participants (see for example, Rettenbach, Diller, & Sireteanu, 1999), suggests that the brain can partially make up for loss in one modality with compensation in another (for a tactile example, see also Borsook et al., 1998).

In specific relation to this proposal, the plasticity can also be seen in attentional and perceptual mechanisms of populations not suffering any sensory loss. Indeed, many recent

findings indicate that attentional mechanisms can be significantly modulated as a side effect of specific daily activities or hobbies. For instance, a recent functional magnetic resonance imaging (fMRI) study found video-game players to have more prefrontal cortex activity than non-players during complex non-gaming tasks (e.g., one task required participants to monitor a target while moving their hands to control a cursor in the opposite direction), a change that the authors attribute to the constant demands on spatial attention while training with video games (Granek, Gorbet, & Sergio, 2010).

Such neurological differences as mentioned above are similarly manifested in behavioral performance on a number of tasks. For example, research using expert video game players (VGPs) has recently sparked interest in investigating the effects of habitual game play and human cognition, leading to findings suggesting greater availability of attentional resources and also flexibility of the deployment of attention on multiple object-tracking tasks, enumeration tasks, perceptual load tasks, and the Attentional Network Test (ANT, see Posner & Rothbart, 2007; for a review of the above tasks, as well as others, see: Dye, Green, & Bavelier, 2009; Green, Li, & Bavelier, 2010; I. Spence & Feng, 2010). While the popularity of video games has grown in recent years, it is a fairly new pastime only witnessing widespread play since the 1990s. Musical performance, on the other hand, is as old as civilization itself (Sachs, 1943). Furthermore, similar to video game play, it might be possible that regular practice of this ancient and ubiquitous activity could result in augmented attentional capacities.

The topic of non-musical benefits (i.e., outside of their domain of expertise) in musicians has been the focus of much research, not in the least due to public interest and popular notions such as the “Mozart effect.” This commercialized phenomenon, which spans a number of products claiming to enhance development and increase intelligence in infants (e.g., Baby Bach,

Baby Mozart, etc.), often misrepresents a popular study by Rauscher, Shaw, and Ky (1993), that in fact only demonstrated a temporary (15 minutes) enhancement of spatial-temporal reasoning after listening to Mozart. Many similar investigations claiming performance enhancements after listening to classical music have demonstrated limited effects, or effects possibly accountable to mood arousal or other indirect factors (Steele et al., 1999; Thompson, Schellenberg, & Husain, 2001). Moreover, a wide range of research has examined how musical training can influence performance in areas such as mathematics (proportional reasoning scores), language (reading test scores), spatial-temporal abilities, and verbal memory (for a concise summary of findings, see Rauscher, 2003). Despite these latter efforts, it should nevertheless be noted that there is considerable skepticism and controversy on the existence of actual cognitive enhancements due to musical training, with many arguing that the present evidence is inconclusive (Črnčec, Wilson, & Prior, 2006; Schellenberg, 2001).

In examining possible training related attentional enhancements, it should be noted that specific training parameters are often necessary. For example, the purported improvements for video game players are found exclusively when training with “action” video games, which by their very nature demand quick responses to peripheral cues, multi-tasking, and efficient perceptual processing (Cohen, Green, & Bavelier, 2008). For instance, the regular use of action video games has been found to enhance visual spatial resolution (Green & Bavelier, 2007), as well as temporal resolution (Donohue, Woldorff, & Mitroff, 2010). However, one could easily claim that musical training does not have such similar characteristics.

Specifically, the demands of video game and musical activities are different in at least three aspects: modality of emphasis, vigilance requirements, and the degree of symbolic manipulation. Although both activities involve multisensory information, video games focus on

visual stimuli whereas music is more auditory based. Secondly, where most action video games require quick responses and vigilance for dealing with unexpected occurrences (e.g., impending foes), music arguably does not require this same extent of alertness, and certainly does not involve the same level of unpredictability (even musical improvisation often occurs within structures). Thirdly, varying with the type of performance, music can demand spontaneous and continuous manipulations of symbolic representations (e.g., improvisation) that may consume significant cognitive resources, where this may be less likely to occur during action video game play. Thus, there are clear contrasts between the two activities, and it is presently open to investigation whether this will translate to differences in information processing in other domains.

Despite the differences between training with action video games and music, experimentation has nevertheless also shown attentional enhancements for the latter sample. A study by Helmbold and colleagues (2005), for example, compared 70 musicians to non-musicians on psychometric assessments of intelligence and general mental abilities (verbal comprehension, word fluency, mental rotation, perceptual speed, reasoning, number computations, various memory tasks, etc.) using five sub-tests from the Leistungsprüfsystem (a German intelligence test). No differences in performance were found except on two tests: flexibility of closure, and perceptual speed. The flexibility of closure task (identifying hidden visual patterns) involved detecting single elements in complex objects, while the measure of perceptual speed involved finding visually presented letters amongst digit distractors. Relevant to our interest of musicians' attentional resources, the authors speculated that the better performance on tasks measuring perceptual speed could be explained by the demands of musical training in requiring quick recognition of musical symbols or structures. In a related study, Jones

and Yee (1997) found musicians to be better at discriminating time change to auditory rhythmic patterns, but only when these were simple patterns. Furthermore, recent research using a line bisection task also showed faster reaction times and fewer errors in musicians, suggesting better visual perceptual processing (Patston, Hogg, & Tippett, 2007). Collectively, these findings suggest that temporal and spatial processing might indeed be enhanced in the expert musician when compared to non-musicians.

Reflecting perhaps the interest in ascertaining the true effects and structural differences caused by musical training, there is a wealth of neuropsychological studies on the musician's brain. A considerable amount of evidence seems to suggest that musicians have greater neuroplasticity resulting in both functional and anatomical differences, including increases of grey and white matter volume in particular locations within the left cerebellum. Other evidence from neurophysiological studies suggests more pronounced cortical reorganizations for musically related motor activity, and larger evoked potential fields (25%) in response to instrumental tones when compared to controls (Gaser & Schlaug, 2003; Münte, Altenmüller, & Jäncke, 2002). Whether or not these musically related brain differences also lead to benefits in temporal processing is speculative (i.e., perhaps these differences already existed, and lead people to become musicians). However, the aforementioned behavioral evidence with musicians does point towards increased attentional and perceptual abilities on various tasks in both the auditory and visual modalities (for visual enhancements, see also Lim & Sinnott, 2011).

Using a sample of expert musicians enables a closer look at specific theories of attention. For instance, Farah, Wong, Monheit, and Morrow (1989; see also Pavani, Husain, Ládavas, & Driver, 2004) have claimed that all sensory modalities have access to a single reservoir of attentional resources, while other authors have argued for a segregated system including

individualized resources for specific modalities (Sinnott, Costa, & Soto-Faraco, 2006; Sinnott, Juncadella, Rafal, Azanon, & Soto-Faraco, 2007; C Spence & Driver, 1996; Wickens, 1984). Thus, by testing musicians for enhanced attentional and perceptual capabilities in the visual modality, we can indirectly assess whether training in one sense leads to performance enhancements in another. Such a finding would closely align with recent investigations involving VGPs, where auditory enhancements were observed despite the training being mostly visual based (Donohue, et al., 2010; Green, Pouget, & Bavelier, 2010).¹

Accordingly, the research described here has both theoretical and practical relevance. Firstly, extending the investigation of enhanced attentional mechanisms to novel populations (e.g., musicians) will support many of the findings pertaining to the plasticity of the human attentional system, specifically in regards to the speed at which information can be processed. Secondly, comparing performance across different sensory modalities enables us to take an initial step in determining the extent to which attention may be best described by segregated or supramodal theoretical frameworks. Lastly, it should be noted that this is the first study with musicians that directly compares performance across identical tasks in visual, auditory, and crossmodal conditions. However, before describing the approach taken to address these questions, and the findings, it is important to first discuss in detail temporal perception.

Temporal perception

The investigation of the human capacity to discriminate temporal events is one of the oldest topics in experimental psychology. Indeed, this topic has witnessed a variety of methodologies in multiple sensory modalities (Bald, Berrien, Price, & Sprague, 1942; Boring, 1929; Exner, 1875).

¹ It should be noted, however, that any such findings must be interpreted cautiously as studies that recruit already “trained” participants such as VGPs and musicians cannot infer that the training itself is the cause for such enhancements. Only an intervention study that administers such training in a controlled setting can make such strong claims in regards to trainings, and specifically to a segregated attentional system as discussed here.

The rich history and varied empirical approaches to understanding temporal perception exemplifies the fundamental implications of temporal processing on how humans perceive and interact with their environment. Indeed, it influences audiovisual integration, speech recognition, and how humans generally integrate multiple stimuli into coherent percepts of events (Navarra et al., 2005). For instance, audiovisual integration has been found to depend on attention and to have specific temporal constraints (Alsius, Navarra, Campbell, & Soto-Faraco, 2005). This is highlighted by the seemingly simple act of speech recognition, which has been shown to rely on detecting temporal cues, thereby necessitating accurate temporal discrimination (Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995; Tallal, Miller, & Fitch, 1993).

The temporal order judgment (TOJ) task has been widely used as a tool to measure temporal processing. From this task, two measures of perceptual processing can 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 stimuli 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). It is interesting to note that, in general, humans are quite proficient at temporal discrimination. For instance, in studies examining within and cross-modal (visual, auditory, and tactile) temporal order judgments (TOJ), Hirsh and Sherrick Jr. (1961) demonstrated that participants could discriminate temporal order between stimuli (JND) when presented as quickly as 20ms apart. Interestingly however, in crossmodal tasks (i.e., audio-visual presentations) the visual stimuli had to lead auditory stimuli by approximately 40-80ms for participants to perceive them as being presented simultaneously (PSS; Hirsh & Sherrick Jr, 1961; Zampini, Shore, & Spence, 2003). Other research also suggests

that the baseline resolution of temporal acuity is better in the auditory modality than in the visual or tactile modalities (Chen & Yeh, 2009; Gebhard & Mowbray, 1959; O'Connor & Hermelin, 1972; Welch, DuttonHurt, & Warren, 1986).

Furthermore, performance on TOJ tasks can also be sensitive to procedural variations, changing according to response requirements (e.g., “which side came first?”, “which modality came first?”, etc.), as well as whether task stimuli are presented from different or same locations. In an audiovisual TOJ study by Zampini et al. (2003) for instance, it was found that participants could discriminate up to a threshold of 22ms when the task required them to respond which ‘modality’ occurred first (i.e., visual, auditory, or tactile), as compared to 62ms when responding which ‘side’ occurred first (i.e., left or right). This would indicate that determining the spatial resolution of temporal order may be more difficult than simply determining the order of modality of presentation. Additionally, Zampini et al. noted that across different modalities, an average increase in temporal resolution of 10-30ms could be observed when stimuli was presented from different spatial locations (i.e., first on the left and second on the right), as opposed to when they were repeatedly presented from the same location (i.e., first on the left and second also on the left), suggesting that spatial information can provide additional help in discriminating temporal order.

Given the efficacy of which humans can discriminate temporally, it is surprising to note that significant gains or losses in processing can occur as a result of injury, age, or training. This would imply that temporal perception is perhaps dependent on attentional mechanisms, and not a purely sensory-based process. This has been demonstrated by research with brain-injured patients. For example, deficits in temporal discrimination have been linked to stroke patients suffering unilateral visual neglect (Husain, Shapiro, Martin, & Kennard, 1997). Here patients

were required to detect specific target letters embedded within a rapid serial visual presentation (RSVP) stream of letters. Brain-injured patients exhibited diminished performance when compared with controls at short temporal lags between targets, suggesting that their ability to detect stimuli appearing in close temporal proximity is greatly impaired. More recent work by Sinnott et al. (2007) also found deficits in both visual and auditory TOJ tasks for patients with right-hemisphere lesions, with JND scores of up to 250ms. Lastly, age has also been observed as a factor that affects temporal order judgment, with general decline in perceptual acuity occurring with increased age (Dinnerstein & Zlotogura, 1968). Further demonstrating the malleability of temporal processing, numerous studies have shown that *enhancements* can occur from repeated video game playing (Donohue, et al., 2010; Green & Bavelier, 2003; Green, Li, et al., 2010; I. Spence & Feng, 2010). Interestingly however, recent research has also extended findings of better performance in an auditory TOJ task to musical-conductors (Hodges, Hairston, & Burdette, 2005).

In a study using both fMRI and behavioral measures, Hodges et al. (2005) examined the effects of musical expertise by comparing ten conductors to age and education matched controls. Overall, conductors outperformed the control group on several tasks. For instance, they found that conductors had better pitch discrimination skills as well as shorter auditory temporal thresholds in a TOJ task. Specifically, conductors required less time between two sounds to correctly discriminate which one had occurred first, and also performed better on a crossmodal TOJ task involving visual targets and concurrent auditory clicks. It should also be noted that previous findings indicated that receiving concurrent information via additional sensory modalities (i.e., a multisensory condition) can actually enhance performance on a temporal order

task (Morein-Zamir, Soto-Faraco, & Kingstone, 2003). These novel finding suggests that conductors may have more efficient and enhanced levels of multisensory processing.

If musical training does indeed lead to enhancements in information processing, then temporal discrimination may be better than control performance both within and across sensory modalities, which would be manifested in lower temporal thresholds (JNDs) in the temporal order judgment task. While this would be expected in the auditory modality, as it corresponds to the medium used in musical training, it is unclear as to what will happen in the visual modality (but note potential enhancements for visual processing in musicians, see Lim & Sinnett, 2011) or even in crossmodal situations. Other patterns of differences (e.g., prior entry effects, seen in PSS scores) within and between musicians and controls will likewise have implications on sensory dominance as well as how attentional orienting occurs across sensory modalities (see discussion section of Experiment 2).

The investigation described below compared trained musicians with controls (who did not have any musical training) on a series of TOJ tasks that were presented in either the visual or auditory modalities, or across modalities. Additionally, within or across modality cues that non-predictively cued target order were incorporated. Altogether this created seven different conditions. While all of the participants completed each of these tasks (in a randomized and counterbalanced order), for ease of presentation the conditions will be separated and described in two experiments. Experiment 1 will discuss visual, auditory, and crossmodal TOJ tasks that did not include spatial clues. Of main interest here is the possible modulations of detection thresholds (JND) in all conditions. The PSS however, should theoretically be close to zero (i.e., show no bias) in the unimodal conditions, while in the crossmodal condition the PSS should indicate that visual stimuli needs to be presented before auditory stimuli for simultaneity to be

perceived (see Hirsh & Sherrick Jr, 1961; Zampini et al., 2003). It is possible that this effect would be reduced in musicians if they do indeed process multisensory information more efficiently. The cued conditions will be discussed in Experiment 2. These conditions involve adding a spatial cue that could be in either the same, or a different, modality as the target task. For instance, the visual TOJ task could have a visual (peripheral flash) or auditory cue (lateral beep) preceding the presentation of the stimuli, which would characterize them as unimodal and crossmodal conditions, respectively. While the JND continues to be of interest here, it is also important to measure the effect of the cue (via PSS scores) and whether there are any differential effects in attentional capture across groups.

Experiment 1: Measuring temporal perception

Participants

A total of 20 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 placed throughout campus. To qualify as a musician, participants were required to have at least three years of formal training in music, and to report a regular practice schedule of at least six hours/week over the past six months (see Appendix 1 for instrument types and musical experience reported). These figures were chosen based on similar cutoffs for experts used in studies with video game players and musicians (Green & Bavelier, 2003; Helmbold, et al., 2005).

Control participants ($n = 20$; age = 22 ± 5 , 16 females)² were recruited from undergraduate courses at the University of Hawaii at Manoa. All control participants had little or

² Given the difference in male to female ratio for controls (4:16) versus musicians (15:5), we pooled together all male and female participants across both musicians and controls for t-test comparisons. No differences were found in accuracy scores between male and female for any of the conditions in both Experiment 1 and 2 (all $p > .1$).

no formal training in music and normal or corrected to normal hearing and vision, as assessed by an initial survey. Control participants were offered course credit for their participation, and musicians were given \$10 in order to facilitate recruitment. Ethical approval was obtained from the Committee on Human Subjects at the University of Hawaii at Manoa.

Stimuli and apparatus

The basic temporal order judgment 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 following the staircase setup of Stelmach and Herdman (1991), also explained in a psychometric paper by Levitt (1971). The SOA for each successive trial either decreases or increases in a stepwise manner dependent on whether the participant answers the previous trial correctly. These stepwise increments occurred at intervals of 16.7ms (monitor refresh rate). The experiment started off with a relatively easy SOA of 167ms, and as it progressed, each trial's SOA decreased until the stimuli become very close and 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 then terminates once a cutoff number of turning points is reached (12 in this study; see West et al., 2008 for a similar approach, albeit one with a less conservative cutoff).

Stimuli were presented on a 21inch Core2Duo 2.4 GHz iMac computer using DMDX software (Forster & Forster, 2003), with visual stimuli occurring on screen and audio stimuli presented via external speakers placed directly besides the monitor. Participants were seated at

an eye to monitor distance of approximately 60cm. From this distance, all presented auditory stimuli occurred at approximately 75 decibels, as measured by a sound meter.

Visual, auditory, and crossmodal stimuli were used to compare musicians and controls' responses in various modalities. Prior to each trial a fixation-cross flanked by two square placeholders on the left and right was presented (see Figure 1). The fixation cross was 0.5° visual angle wide. The placeholder squares were 1.4° wide and situated 4° from the central fixation.

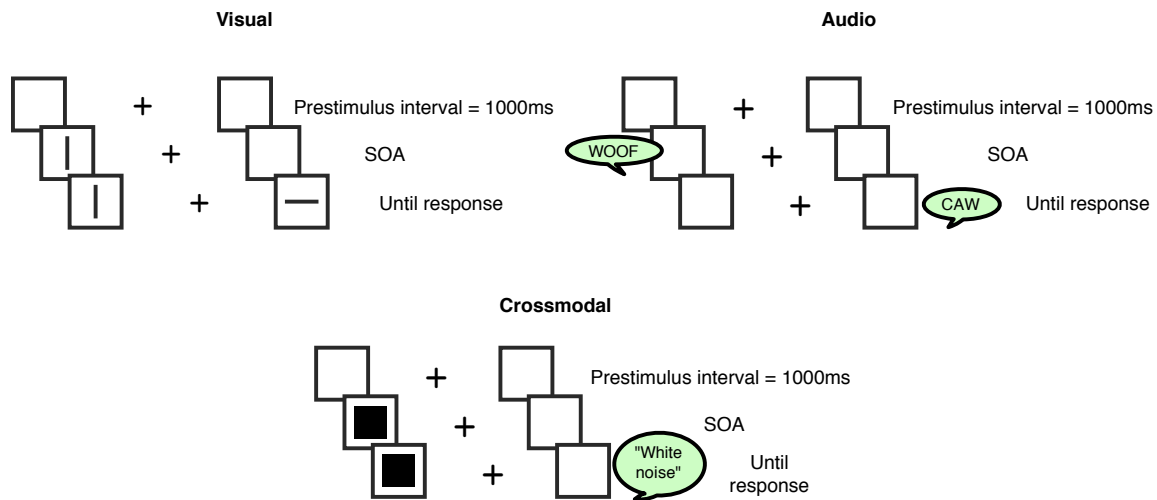


Figure 1. Stimuli for the three TOJ tasks in Experiment 1. On each trial participants were presented with two placeholder boxes on either side of fixation. After 1000ms one randomly chosen stimulus appeared in one of the placeholders. The onset times of the second stimulus were determined by the step-function which monitored each participant's accuracy rate.

Stimuli for the visual task were horizontal and vertical lines subtending 0.9° and occurring centrally within the placeholder squares. For the auditory stimuli, processed samples of a dog and crow sound, both lasting for 350ms were used. In the audio-visual (crossmodal) condition, the visual stimulus consisted of a black square of width 0.9° (appearing within the placeholder), whereas the auditory stimulus was 50ms of white noise. Responses were made via key presses to one of two buttons on a keyboard.

Procedure

The procedure was identical across all three modality conditions. Participants made unspeeded responses signaling which stimulus they believed had appeared first using one of two keyboard buttons corresponding to each stimulus (e.g., horizontal or vertical line, etc.). In all three modality conditions, participants were first presented with onscreen instructions followed by a short sequence of practice trials, with accuracy feedback directly appearing after each trial. The experimenter monitored completion of the practice trials and ensured that participants understood the task requirements (repeating the practice session if needed). Presentation side (i.e., left or right) and stimuli order (e.g., horizontal or vertical line first) were all randomized, as was the order of experimental conditions (e.g., audio, visual, crossmodal) for each participant. Each condition took approximately 5-8 minutes to complete.

Results

The results from the TOJ task can be analyzed to determine the minimum amount of time that must separate two events such that they are still accurately perceived as occurring successively (and not simultaneously), that is, the just noticeable difference (JND). This is essentially a measure of the resolution or threshold of temporal discrimination. A second measure can be extracted from the results, referred to as the point of subjective simultaneity (PSS). The PSS is the point in time in which one stimulus can be presented before the other such that they are still perceived as being simultaneous. Although this measure is usually expected to fall at 0ms (or close to it) in unimodal conditions (unless there is a bias in response), it is more informative in the crossmodal condition as it indicates whether auditory or visual events must be presented first

for them to be perceived as simultaneous. Recall that Zampini et al. (2003) demonstrated that normally a visual target must precede an auditory target for simultaneity to be perceived.

The calculation of both the JND and PSS was based on approaches used by previous research (for examples of other studies using similar methodologies and analyses, see C. Spence, Baddeley, Zampini, James, & Shore, 2003; Stelmach & Herdman, 1991). To begin with, the data from musicians and controls were pooled into separate groups. The average ratio of responses "horizontal line first" (or "crow first"/"auditory first") was then plotted as a function of the time in which the horizontal line preceded the vertical line (or the crow sound preceded the dog sound, etc.). For TOJ tasks, response rates typically follow a sigmoidal curve, from which data can be fit using the following logistic function:

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

where the response rate is mapped as a function of the SOA (x), with two estimated parameters of central tendency (a), and slope (b ; see C. Spence, et al., 2003).

Data was fit to this equation by minimizing the weighted sum of squares to obtain parameter estimates for a and b . The PSS, or SOA at which the participants considered the two stimuli to be simultaneous, corresponds to parameter a . The JND, or smallest interval between two stimuli giving a correct judgment probability 75% of the time, is directly related to parameter b (analogous to the slope of the central portion of logistic function). Here the relationship is that a steep slope will result in a smaller JND, and a shallow slope in a larger JND. The JND can be obtained using the slope according to the following formula (see C. Spence, et al., 2003):

$$\text{JND}_{75} = 1.099 \cdot b$$

Confidence intervals (95%) for each group statistic were calculated using a parametric bootstrap method with 999 replications (Efron & Tibshirani, 1993). For between group comparisons a permutation bootstrap method was used, where data from both groups were combined and resampled to construct a distribution from which the likelihoods (p-value) of obtaining the observed differences between each groups' scores from a mixed population pool were estimated (for an example of another TOJ study that employed a similar bootstrap approach, see Azañón & Soto-Faraco, 2007).

Given the unique constraints of our dataset, we decided to use a bootstrap resampling approach for the statistical analyses due to particular benefits over more traditional means. That is, due to the varied number of observations and different response patterns resulting from the adaptive staircase paradigm, each individual's data points could vary significantly, and fitting the logistic function individually did not always converge or yield meaningful estimates. Thus, combining data from all participants in each group allowed for a better distribution of scores across all SOAs for the logistic fit from which we were able to extrapolate the overall JND and PSS values for musicians and controls using the abovementioned functions. Furthermore, using bootstrap resampling enabled a direct comparison of these parameters as well as confidence interval estimations (for a study using a similar bootstrap approach with pooled data, see Jeon, Hamid, Maurer, & Lewis, 2010). Unfortunately however, the combining of scores across participants did not allow for a traditional ANOVA analysis for main effects or interactions due to the lack of estimates for individual scores. Nevertheless, we strove to compare group scores both within and across experiments to provide the most comprehensive scope of analyses.

Auditory condition: In the auditory condition, the 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; $p = 0.07$; see Figure 2).

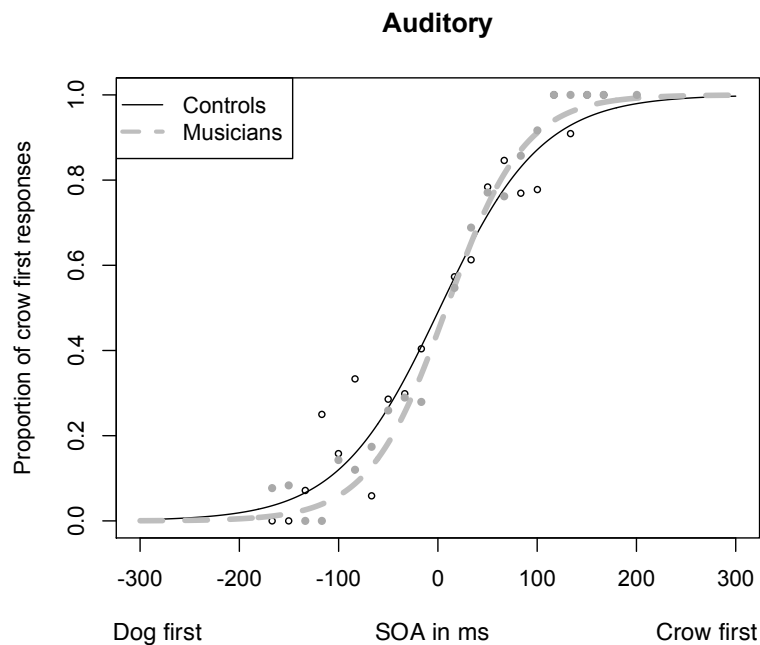


Figure 2. Pooled data for the auditory condition in Experiment 1.

Visual condition: In the 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 the negative PSS values indicating a possible bias in responses towards horizontal lines for both groups. Nevertheless, the confidence intervals for both groups straddle 0ms, and thus may not actually differ from 0ms. The average musicians' JND score was also significantly lower than controls by 18ms (29ms, CI = 23 to 35ms; and 47ms, CI = 37 to 56ms; $p = 0.006$; see Figure 3).

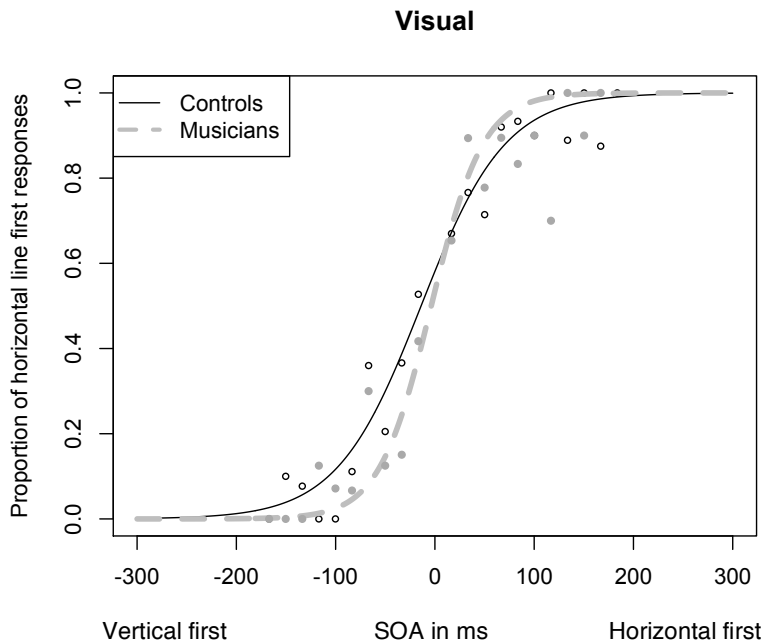


Figure 3. Pooled data for the visual condition in Experiment 1. Note the steeper curve for the musicians indicating better temporal resolution (i.e., lower JND).

Crossmodal condition: In the crossmodal condition, musicians' average PSS score was not significantly different than controls (-43ms, CI = -60 to -25ms; and -63ms, CI = -93 to -30ms; $p = 0.261$; respectively). It is also worth noting here that the negative PSS results indicate a bias in response towards auditory stimuli. That is, the visual stimuli needed to be presented prior to the auditory stimuli in both groups for simultaneity to be perceived (see also Zampini et al., 2003). In regards to the average JND score, musicians' were significantly lower than controls by 59ms (104ms, CI = 80 to 127ms; and 163ms, CI = 112 to 207ms; $p = 0.021$; see Figure 4).

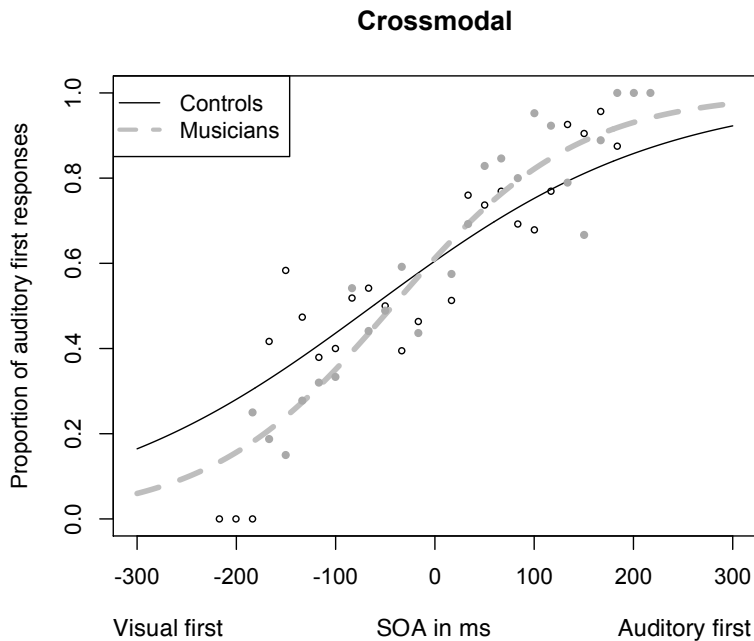


Figure 4. Pooled data for the crossmodal condition in Experiment 1.

Discussion

There were no differences between musicians and controls for auditory PSS scores, although this is to be expected given that PSS should be relatively close to zero, unless there is a bias of some sort to one of the targets, or if a spatial cue is incorporated (i.e., Experiment 2). Furthermore, and contrary to our initial hypothesis, the difference in temporal thresholds between musicians and controls in the auditory condition was only marginally lower ($p = 0.07$). This is surprising given that musical training is arguably largely auditory in nature, and therefore a more robust difference in auditory JND scores between musicians and control participants was expected. On the other hand, in the visual condition musicians did have both significantly lower PSS as well as JND scores, compared to controls. Although visual PSS scores for both musicians and controls were biased towards the horizontal line, the significant difference between groups demonstrates that control participants were slightly more biased towards responding to the horizontal bar (i.e.,

the vertical bar needed to be presented 14ms prior to the horizontal bar for simultaneity to be perceived). It is difficult to speculate as to why participants favored the horizontal bar overall, or why control participants favored the horizontal bar more than the musicians. However, it should be noted that PSS scores were low in general (i.e., close to zero, with confidence intervals also straddling 0ms). More importantly, the significant JND enhancements seen in the visual modality may suggest a modulation of attention.

Despite the differences observed in the visual modality, musicians only showed a marginal difference when compared to non-musicians in the auditory modality. It is possible for instance, that the “realistic” auditory stimuli used in our experiment may be more difficult than simpler tones, and therefore any effect might be masked. Furthermore, it might be possible that as the sounds were non-tonal (musical tones have been used in past research, see Hodges, et al., 2005), musicians may not have had a distinct advantage. This could also have lead to a ceiling effect. Indeed, given that humans discriminate temporal events better in the auditory modality when compared to the visual modality, it is possible that both musician and control participants performed similarly due to a ceiling effect (Chen & Yeh, 2009; Gebhard & Mowbray, 1959; O'Connor & Hermelin, 1972; Welch, et al., 1986). Note however, that much of this is based on speculation, and therefore future research is needed.

The largest differences were seen in the crossmodal condition. As can be seen in Figure 5, it is worth noting that PSS scores were substantially larger in the crossmodal condition when compared with those in the auditory or visual conditions, with a shift towards visual first SOAs. That is, for both groups, visual stimuli had to precede auditory stimuli (by 43 and 63ms) for them to be perceived as occurring simultaneously (although there were no differences between musicians and controls for PSS). These findings coincide with previous crossmodal TOJ

findings, where visual events generally preceded auditory events by 40 to 80ms (Hirsh & Sherrick Jr, 1961; Zampini, et al., 2003) for simultaneity to be perceived. Additionally, musicians had significantly lower JND scores, indicating that the temporal resolution of multisensory processing was better for the musicians who took part in this study. Speculatively, this could potentially be due to the requirements of reading musical notation while concurrently listening to auditory information. And lastly, also supported by Zampini et al. (2003), JND scores increased nearly three-fold when compared to unimodal conditions, demonstrating that the task was more difficult.

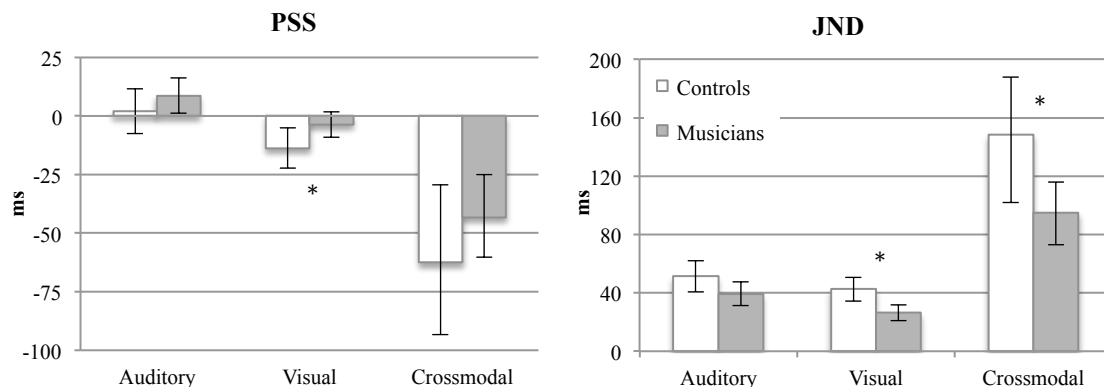


Figure 5. PSS and JND scores for Experiment 1. Positive PSS values indicate a bias towards the dog sound in the auditory condition, vertical line in the visual condition, and visual stimuli in the crossmodal condition. Negative values indicate biases towards opposite stimuli (crow/horizontal line/auditory stimuli). Asterisks indicate significant between group differences, and error bars indicate 95% confidence intervals.

In addition to the basic TOJ setup used in Experiment 1 to measure temporal perception, spatial cues can also be incorporated into these tasks, thereby allowing a measure of how attention is oriented and captured for the two groups (in addition to still being able to measure JND). This is ideal as the additional spatial cues not only provide an opportunity for a better understanding of how information processing is modulated, but are perhaps more analogous to real world situations requiring attention to be directed to a task while at the same time being presented with irrelevant within and across modality stimuli. The presentation of exogenous cues

prior to stimuli onset in a TOJ task creates a ‘prior entry’ effect, whereby 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 (i.e., in our task the cue is only correct half of the time; see also Shore, Spence, & Klein, 2001; C. Spence, Shore, & Klein, 2001; Zampini, Shore, & Spence, 2005).

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 capture attention). By presenting an exogenous cue in the TOJ task prior to the onset of the first stimuli, participants’ attention should theoretically be involuntarily 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. The PSS score in this case would indicate how much in advance the uncued side must be presented before the cued for subjective simultaneity to be perceived. Analyses generally reveal the PSS as deviating from the central tendency by shifting towards the cued side (Shore, et al., 2001). Thus, in Experiment 2 we replicated the unimodal conditions of Experiment 1 with additional within sensory and cross-modal cues to determine whether spatial attention, as measured by the cued TOJ task, will differ between musicians and non-musicians. Based on findings from Experiment 1, it is hypothesized that musicians should have a smaller orienting affect, which would manifest in lower PSS and JND scores than controls across all conditions. That is, smaller PSS and JND scores 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 (although for preliminary findings in the visual modality using exogenous and endogenous cues, see Lim & Sinnett, 2011).

Experiment 2: Measuring temporal and spatial perception

Participants

The same participants in Experiment 1 also took part in Experiment 2. Recall that all conditions from both Experiment 1 and Experiment 2 were interleaved and randomized into a battery of seven tasks for each participant in order to counterbalance for any possible training effects. The two experiments are separated here only to facilitate coherence in presentation and analyses.

Stimuli and apparatus

The stimuli and conditions were identical to those in Experiment 1, except for the use 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 (see Figure 7). In the auditory condition, the cue was a laterally presented 500Hz sinusoidal wave lasting for 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 will also be randomly determined by the step staircase algorithm, and thus will have an equal chance of validly or invalidly cuing the target stimuli.

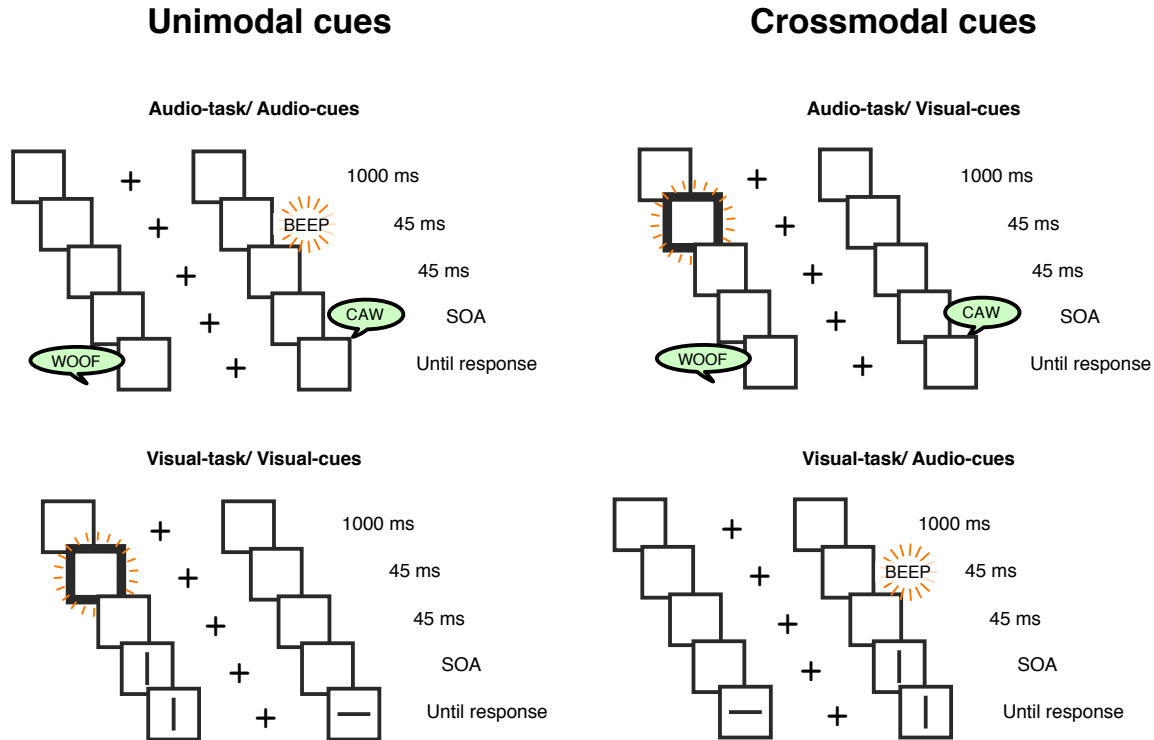


Figure 6. Stimuli for the four TOJ tasks in Experiment 2. On each trial participants were presented with two placeholder boxes on either side of fixation. After 1000ms one randomly chosen placeholder was cued for 45ms followed by the onset of either stimulus in one of the placeholders after a cue-target interval of 45ms. The onset time of the second stimulus was determined by the step-function.

Procedure

The procedure was identical to that used in Experiment 1, except that participants were notified of the cues and that they were non-predictive in nature (i.e., they were instructed to make their judgments based on temporal order alone). The order of experimental conditions (e.g., auditory, visual, crossmodal) was randomized for each participant.

Results

The JND and PSS scores were calculated using similar methods as in Experiment 1. First, data was pooled over 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., see Figure 7 and 8). Separate PSS and JND estimates (a and b parameters) were then calculated for each of the two cued curves. The overall PSS value for each condition was computed as half the distance between each of the PSS values for the two curves,

$$PSS_{overall} = \frac{PSS_2 - PSS_1}{2}$$

and the average of the two JND values for each curve was then used as the overall JND score.

This approach essentially calculates the PSS for each type of stimulus cued, and averages the effect. In order to gauge the influence of the cue, the two fitted curves are 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 (see Shore, et al., 2001), although as the slope, JND_1 and JND_2 are expected to be similar for each stimulus type.

$$JND_{overall} = \frac{JND_1 + JND_2}{2}$$

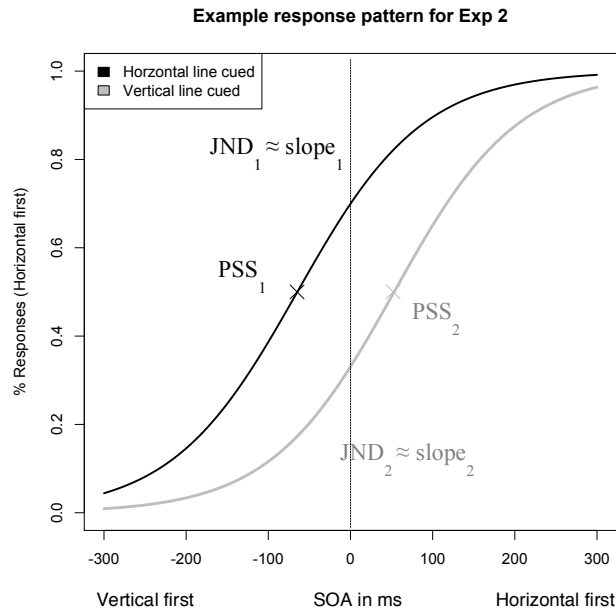


Figure 7. An example of a typical response curve from Experiment 2 (here using the visual task), where data from each condition is grouped according to the stimulus cued, with corresponding parameters used to determine overall PSS and JND. Note the shift from 0 of each curve, indicating the effect that the cue had on the task.

Similar to Experiment 1, confidence intervals (95%) and p-values were computed using appropriate bootstrap and permutation resampling methods for each statistic.

Unimodal cues

Auditory condition: To assess whether the auditory cues had any effect on temporal judgment, crow-cued trials were compared to dog-cued trials. This comparison revealed significant differences in PSS scores for both musicians and controls (both $p < 0.01$), suggesting that the auditory cues effectively captured attention.³ The magnitude of this effect, however, was not significantly different between musicians and controls (23ms, CI = 12 to 37ms; and 29ms, CI = 17 to 42ms; $p = 0.259$; respectively). Similarly, there were no differences in JND scores

³ Note, this analysis does not look at whether the crow or dog was preferred (i.e., analogous to horizontal bars being preferred in the visual condition of Experiment 1), but merely confirms whether there was indeed a cueing effect.

between the two groups (92ms, CI = 77 to 110ms; and 109ms, CI = 93 to 125ms; $p = 0.106$; see Figure 9).

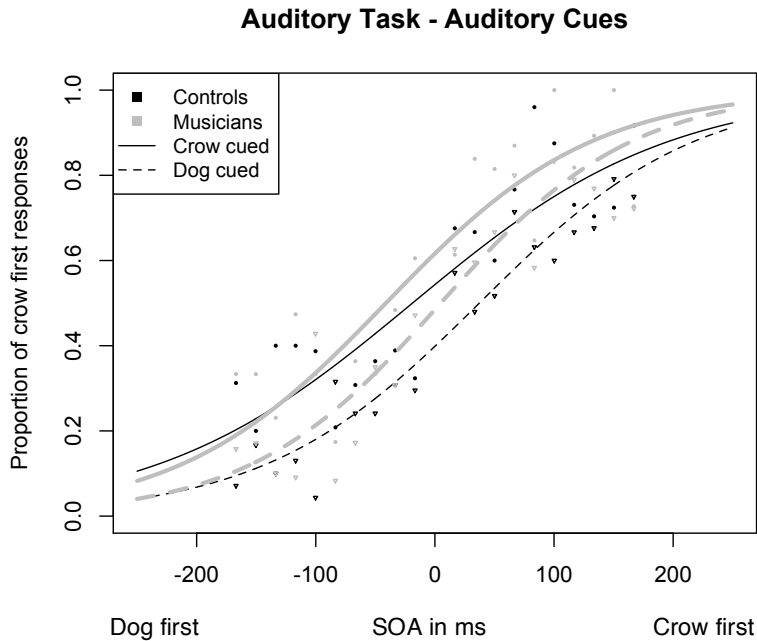


Figure 8. Pooled data for the auditory-task/auditory-cues condition in Experiment 2.

Visual condition: To assess whether the visual cues had any effect on temporal judgment, horizontal-cued trials were compared to vertical-cued trials. This comparison revealed significant differences in PSS scores for both musicians and controls (both $p = 0.001$), suggesting that the visual cues effectively captured attention. Much like Experiment 1, the magnitude of this effect, however, was significantly lower for musicians than controls by 29ms (30ms, CI = 10 to 46ms; and 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; and 84ms, CI = 75 to 96ms; $p = 0.29$; see Figure 10).

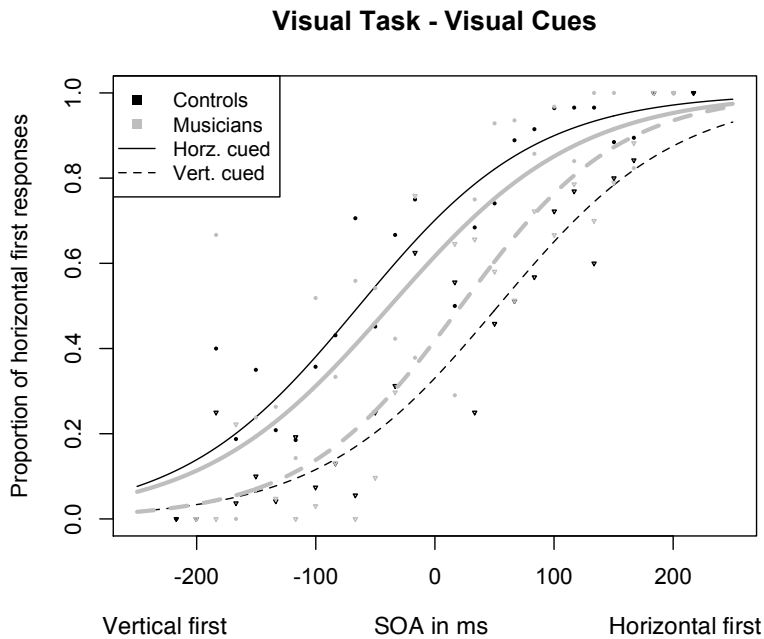


Figure 9. Pooled data for the visual-task/visual-cues condition in Experiment 2.

Crossmodal cues

Auditory TOJ with visual cues: To assess whether the crossmodal cues had any effect on temporal judgment, crow-cued trials were compared to dog-cued trials. This comparison revealed significant differences in PSS scores for both musicians and controls (both $p < 0.05$), suggesting that the visual cues effectively captured attention. However, the magnitude of the PSS scores did not significantly differ between musicians and controls (8ms, CI = 1 to 14ms; and 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; and 63ms, CI = 53 to 73ms; $p = 0.014$).

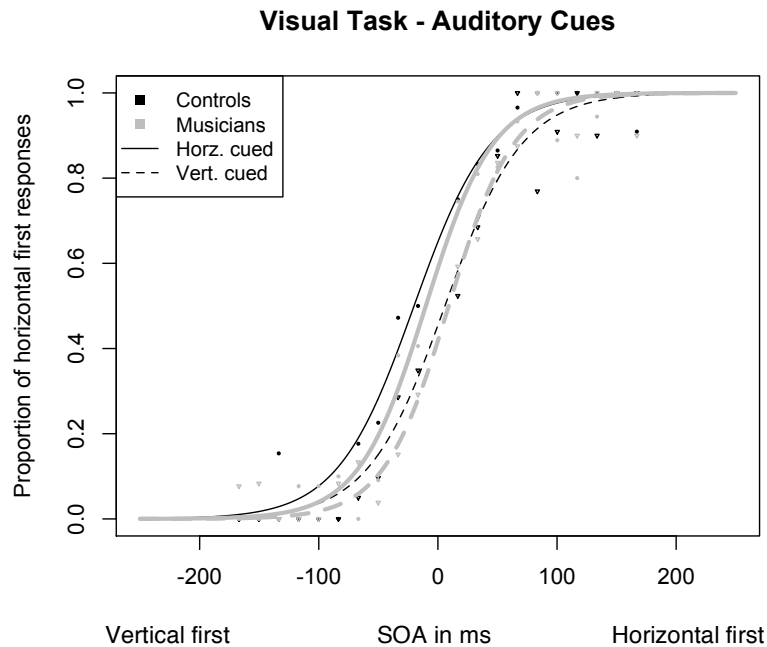


Figure 11. Pooled data for the visual-task/auditory-cues condition in Experiment 2.

Within vs. across modality cue comparisons: To assess any differential effects of within modality or cross modality cues, the respective TOJ tasks were compared. Comparing the auditory TOJ task that had auditory cues to the auditory TOJ task that had visual cues revealed significantly lower PSS and JND scores for the crossmodal condition for both musicians and controls (both groups PSS, $p < .05$ and JND, $p < .001$). Comparing the visual TOJ task that had visual cues to the visual TOJ task that had auditory cues also revealed significantly lower PSS and JND scores for the crossmodal condition for both musicians and controls (both groups PSS, $p < .05$ and JND, $p < .001$).

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.

A comparison of cued JND scores to no-cue scores revealed that they 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). This was reflected by differences between cue conditions in both the auditory (crow or dog-cued) and visual (horizontal or vertical bar-cued) unimodal conditions of Experiment 2 when compared with their respective counterparts in Experiment 1 (auditory and visual no cues condition), for both musicians and controls (all $p < .01$, using the same bootstrap comparison method as conducted throughout this study). However, when the cues were presented in a separate modality (i.e., the crossmodal conditions of Experiment 2), JND scores were indistinguishable from the unimodal no cue conditions (Experiment 1) for both musicians and non-musicians. This was reflected by a lack of differences between crossmodal TOJ tasks and their analogous no cue conditions in Experiment 1 (all $p > .05$). Collectively, this may suggest that a difficult unimodal task can be made easier as a crossmodal task (Sinnott, et al., 2006; Toro, Sinnott, & Soto-Faraco, 2005; Wickens, 1984).

For PSS scores, comparing the no-cue conditions of Experiment 1 to the cued conditions of Experiment 2 yielded less consistent results across conditions. In the unimodal auditory task, the cues had an effect in shifting PSS scores when compared to the non-cued condition for the control participants ($p = 0.041$ and $p = 0.001$ for each target type, crow and dog), whereas for musicians only the crow-cued condition differed from the no-cue condition ($p = 0.001$ and $p = 0.643$). This would suggest that for musicians, auditory cues did not have an effect on the detection of the dog stimulus. In the crossmodal auditory task (with visual cues), cued PSS scores did not differ from the no-cue condition for controls (all $p > 0.05$), but for musicians the crow-cued condition differed from the no-cue condition ($p = 0.04$), while the dog-cued condition

did not ($p = 0.66$). This result is curious in and of itself, as it suggests that visual cues only had an effect on musicians when they appeared before the crowd sound.

Varying trends were seen for visual PSS scores. In the unimodal visual task, horizontal-cued and vertical-cued PSS scores differed significantly from the no-cue conditions for both musicians and controls (all $p < 0.01$). Interestingly, in the crossmodal visual task similar differences occurred for musicians and controls, where horizontal-cued PSS scores did not differ from the no-cue condition (both $p > .1$), while vertical-cued PSS scores did (both $p = 0.004$). Curiously, this finding suggests that auditory cues had a stronger effect on vertical lines than horizontal lines for both groups. It should be emphasized however, that despite the inconsistent PSS findings across cueing and no cueing conditions, the cues nevertheless were effective in capturing attention as based on the within experiment PSS analysis for the cued conditions.

Discussion

Robust findings from cross-experiment analyses broadly suggest that unimodal cues have a detrimental affect on JND scores, whereas crossmodal cues do not. These results were similar for both musicians and controls. Excluding cross experiment analyses, the only observed significant difference 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 cross-modal processing was easier for musicians, but only when judging temporal order for auditory targets that were cued visually (see Figure 12). Note however, that this result does not fully align with the lack of JND differences for both the no-cue auditory condition

(although it was approaching significance) and in the auditory-task/auditory-cue condition of Experiment 2.

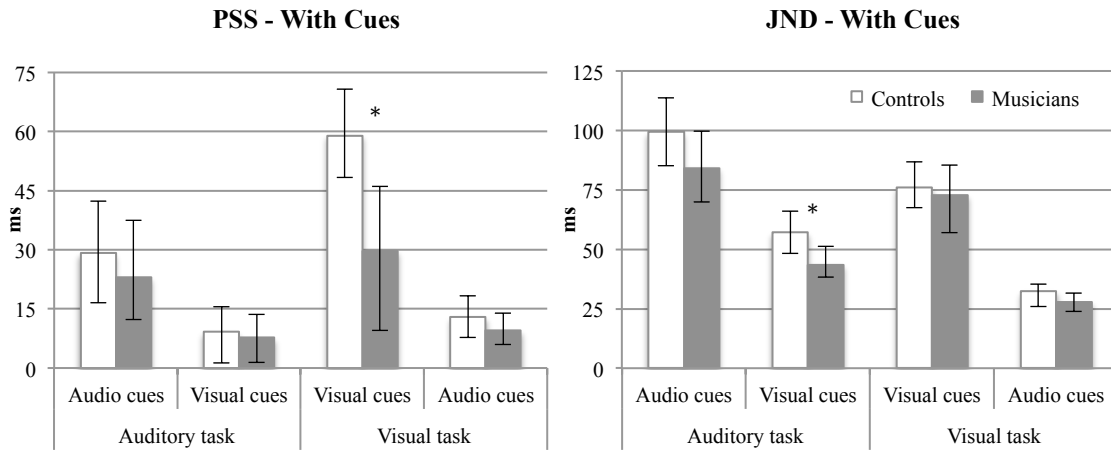


Figure 12. PSS and JND scores for Experiment 2. Asterisks indicate significant between group differences, and error bars indicate 95% confidence intervals.

General Discussion

There are a number of important findings that merit discussion. To begin with, performance between musicians and controls were mixed across both experiments in the auditory condition (musicians did have significantly lower JND scores in the auditory-task/visual-cues condition 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, see Figure 13). 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. The difference in findings 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 (440Hz and 660Hz) were used in Hodges et al.'s studies. Given that our auditory stimuli did not differ as much in frequency (both at

approximately 500Hz), 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.

Experiment 1	PSS	JND
Auditory	<i>ns</i>	Marginally lower for musicians
Visual	Lower for musicians	Lower for musicians
Crossmodal	<i>ns</i>	Lower for musicians
Experiment 2		
Auditory with unimodal cues	<i>ns</i>	<i>ns</i>
Auditory with crossmodal cues	<i>ns</i>	Lower for musicians
Visual with unimodal cues	Lower for musicians	<i>ns</i>
Visual with crossmodal cues	<i>ns</i>	<i>ns</i>

Figure 13. Combined table of findings for Experiment 1 and 2 (*ns* = not significant).

It is worth noting however, that across all task types, JND scores for musicians were lower than those for controls. Although these differences were statistically significant in only three out of the seven conditions (Exp 1: visual and crossmodal; Exp 2: Audio-task/visual cues), it may be the case that with a larger sample size, more results would reach significance. Nevertheless, we do see a trend towards lower thresholds of temporal discrimination in musicians.

Tentative speculation for a supramodal account of attentional resources may also be 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. 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 other modalities however, could be attributed to 1) better attentional resources, and 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 cause however, these results can only be seen as tentative support for a supramodal account, pending further investigation with specific training conditions (see also Footnote 1, page 6). 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 a segregation of attentional systems (Sinnott, et al., 2006; Sinnott, et al., 2007; C Spence & Driver, 1996). Nevertheless, the current set of data makes it difficult to arrive at a decisive claim on either side of the debate. Indeed, it is likely that many of the findings supporting one theoretical account or the other may indeed be constrained by the varying methodologies used.

Another related novel finding that was observed across both musicians and controls is 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. These results suggest 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.

While between group differences in the auditory task in Experiment 1 were not observed (although note that it was approaching significance), 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, which other studies with expert populations have also shown (for an example with

VGPs, see Donohue, et al., 2010). Importantly, the visual enhancements observed in JND in Experiment 1 lend support to the idea that attentional allocation, and thereof the improvement through training, may not be constrained within particular sensory modalities, but instead distributed to multiple modalities as the current task demands. Nevertheless, an important criticism of studies that used “trained” populations such as musicians and VGPs, is the extent to which observed differences in experimental settings can actually be attributed to prior training (Boot, Blakely, & Simons, 2011). In a recent article examining studies that looked at the cognitive effects of video game training, Boot and colleagues in fact extend this criticism to studies using pre-post training procedures. These authors claim that participants are often not blind to the purpose of the study, and that this awareness and potential motivational factor may very well influence performance. Unfortunately, our recruitment strategy for finding 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.

Learning an instrument is a complex task that takes years to master, as it requires a variety of complex motor, auditory and multimodal skills (Ericsson, Krampe, & Tesch-Römer, 1993). Any study aiming to administer musical training to participants in order to measure the cognitive effects then must either choose between realistic instrumental training that requires a considerable duration of time for progress to be made (months to years), or a shorter regimen of training that focuses perhaps on more specific aspects of music learning. It should be noted that most studies to date have examined the effects of training using longitudinal approaches, in order

to realistically replicate the process that trained musicians go through (e.g., for studies with one year training durations, see Fujioka, Ross, Kakigi, Pantev, & Trainor, 2006; Schlaug, Norton, Overy, & Winner, 2005). It would also be informative, however, to train participants on subsets of musical tasks that may require a shorter time period to master, as this would allow us to focus on the more specific effects of the various subcomponents of musical training, and the cognitive mechanisms they potentially invoke. Although many well known studies have examined the effects of listening to excerpts of music on cognitive performance (e.g., Rauscher, et al., 1993; Steele, et al., 1999; Thompson, et al., 2001), we are not aware of any studies that trained participants on actual subsets of musical skills (e.g., pitch discrimination, rhythmic training, notation reading, etc.). Although the full extent of instrumental training, and the multimodal processes involved therein could only be acquired through realistic training, gaining a better understanding of the subcomponents may also help to piece together a picture of what makes musical training unique from a cognitive standpoint and any relevant side effects thereof.

References

- Alsius, A., Navarra, J., Campbell, R., & Soto-Faraco, S. (2005). Audiovisual integration of speech falters under high attention demands. *Current Biology*, *15*(9), 839-843.
- Azañón, E., & Soto-Faraco, S. (2007). Alleviating the crossed-hands deficit by seeing uncrossed rubber hands. *Experimental brain research*, *182*(4), 537-548.
- Bald, L., Berrien, F., Price, J., & Sprague, R. (1942). Errors in perceiving the temporal order of auditory and visual stimuli. *Journal of Applied Psychology*, *26*(3), 382-388.
- Boot, W. R., Blakely, D. P., & Simons, D. J. (2011). Do Action Video Games Improve Perception and Cognition? [Perspective]. *Frontiers in Psychology*, *2*.
- Boring, E. (1929). *A history of experimental psychology*. New York: Appleton-Century.
- Borsook, D., Becerra, L., Fishman, S., Edwards, A., Jennings, C., Stojanovic, M., et al. (1998). Acute plasticity in the human somatosensory cortex following amputation. *Neuroreport*, *9*(6), 1013.
- Chen, K., & Yeh, S. (2009). Asymmetric cross-modal effects in time perception. *Acta psychologica*, *130*(3), 225-234.
- Cohen, J. E., Green, C. S., & Bavelier, D. (2008). Training visual attention with video games: Not all games are created equal. *Computer games and team and individual learning*, 205-227.
- Črnčec, R., Wilson, S. J., & Prior, M. (2006). The cognitive and academic benefits of music to children: Facts and fiction. *Educational Psychology*, *26*(4), 579-594.
- Dinnerstein, A., & Zlotogura, P. (1968). Intermodal perception of temporal order and motor skills: effects of age. *Perceptual and Motor Skills*, *26*, 987-1000.
- Donohue, S., Woldorff, M., & Mitroff, S. (2010). Video game players show more precise multisensory temporal processing abilities. *Attention, Perception, & Psychophysics*, *72*(4), 1120.
- Dye, M. W., Green, C. S., & Bavelier, D. (2009). The development of attention skills in action video game players. *Neuropsychologia*, *47*(8-9), 1780-1789.
- Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap* (Vol. 57): Chapman & Hall/CRC.

- Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363.
- Exner, S. (1875). Experimentelle Untersuchung der einfachsten psychischen Prozesse [Experimental research on simple physical processes]. *Pflugers Archive*, *11*, 403-432.
- Farah, M., Wong, A., Monheit, M., & Morrow, L. (1989). Parietal lobe mechanisms of spatial attention: modality-specific or supramodal? *Neuropsychologia*, *27*(4), 461-470.
- Forster, K., & Forster, J. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*(1), 116.
- Fujioka, T., Ross, B., Kakigi, R., Pantev, C., & Trainor, L. J. (2006). One year of musical training affects development of auditory cortical-evoked fields in young children. *Brain*, *129*(10), 2593.
- Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *Journal of Neuroscience*, *23*(27), 9240.
- Gebhard, J., & Mowbray, G. (1959). On discriminating the rate of visual flicker and auditory flutter. *The American journal of psychology*, *72*(4), 521-529.
- Granek, J., Gorbet, D., & Sergio, L. (2010). Extensive video-game experience alters cortical networks for complex visuomotor transformations. *Cortex*, *46*(9).
- Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature*, *423*(6939), 534-537.
- Green, C. S., & Bavelier, D. (2007). Action-video-game experience alters the spatial resolution of vision. *Psychological Science*, *18*(1), 88-94.
- 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.
- Husain, M., Shapiro, K., Martin, J., & Kennard, C. (1997). Abnormal temporal dynamics of visual attention in spatial neglect patients. *Nature*, 385(6612), 154.
- Jeon, S. T., Hamid, J., Maurer, D., & Lewis, T. L. (2010). Developmental changes during childhood in single-letter acuity and its crowding by surrounding contours. *Journal of Experimental Child Psychology*, 107, 423-437.
- 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.
- Kolb, B., Gibb, R., & Robinson, T. (2003). Brain plasticity and behavior. *Current Directions in Psychological Science*, 12(1), 1.
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, 49(2), 467-477.
- Lim, A., & Sinnott, S. (2011). *Exploring Visual Attention in Musicians: Temporal, Spatial and Capacity Considerations*. Paper presented at the 33rd Annual Conference of the Cognitive Science Society, Boston, MA.
- Morein-Zamir, S., Soto-Faraco, S., & Kingstone, A. (2003). Auditory capture of vision: examining temporal ventriloquism. *Cognitive Brain Research*, 17(1), 154-163.
- Münte, T., Altenmüller, E., & Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. *Nature Reviews Neuroscience*, 3(6), 473-478.
- Navarra, J., Vatakis, A., Zampini, M., Soto-Faraco, S., Humphreys, W., & Spence, C. (2005). Exposure to asynchronous audiovisual speech extends the temporal window for audiovisual integration. *Cognitive Brain Research*, 25(2), 499-507.
- O'Connor, N., & Hermelin, B. (1972). Seeing and hearing in space and time. *Perception & psychophysics*, 11, 46-48.
- 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.

- Posner, M., & Rothbart, M. (2007). Research on attention networks as a model for the integration of psychological science. *Psychology*, 58(1), 1.
- Rauscher, F. (2003). Can Music Instruction Affect Children's Cognitive Development? *ERIC Clearinghouse on Early Education and Parenting* Retrieved 10 October, 2011, from <http://www.ericdigests.org/2004-3/cognitive.html>
- Rauscher, F., Shaw, G., & Ky, K. (1993). Music and spatial task performance. *Nature*, 365(6447), 611.
- Rettenbach, R., Diller, G., & Sireteanu, R. (1999). Do deaf people see better? Texture segmentation and visual search compensate in adult but not in juvenile subjects. *Journal of cognitive neuroscience*, 11(5), 560-583.
- 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.
- Sachs, C. (1943). *Rise of music in the ancient world east and west*: W.W. Norton.
- Schellenberg, E. (2001). Music and nonmusical abilities. *Annals of the New York Academy of Sciences*, 930(1), 355-371.
- Schlaug, G., Norton, A., Overy, K., & Winner, E. (2005). Effects of music training on the child's brain and cognitive development. *Annals of the New York Academy of Sciences*, 1060(1), 219-230.
- Shannon, R., Zeng, F., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, 270(5234), 303.
- Shimojo, S., & Shams, L. (2001). Sensory modalities are not separate modalities: plasticity and interactions. *Current Opinion in Neurobiology*, 11(4), 505-509.
- 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 inattention 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.
- Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *Journal of Experimental Psychology: General*, 130(4), 799-832.
- Spence, I., & Feng, J. (2010). Video games and spatial cognition. *Review of General Psychology*, 14(2), 92.
- Steele, K., Dalla Bella, S., Peretz, I., Dunlop, T., Dawe, L., Humphrey, G., et al. (1999). Prelude or requiem for the Mozart effect? *Nature*, 400(6747), 827.
- 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.
- Tallal, P., Miller, S., & Fitch, R. (1993). Neurobiological basis of speech: a case for the preeminence of temporal processing. *Annals of the New York Academy of Sciences*, 682(1), 27-47.
- Thompson, W., Schellenberg, E., & Husain, G. (2001). Arousal, mood, and the Mozart effect. *Psychological Science*, 12(3), 248.
- Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition*, 97(2), B25-B34.
- Welch, R., DuttonHurt, L., & Warren, D. (1986). Contributions of audition and vision to temporal rate perception. *Perception & psychophysics*, 39(4), 294-300.
- 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.
- Zampini, M., Shore, D., & Spence, C. (2005). Audiovisual prior entry. *Neuroscience letters*, 381(3), 217-222.

Appendix 1

Table 1.

Musician participants questionnaire results

Education	<i>n</i>	Percentage
High school graduate	2	10
Currently in college	13	65
College graduate	3	15
Currently in graduate school	2	10
Principle instrument		
Guitar	10	50
Bass guitar	3	15
Piano	3	15
Saxophone	1	5
French horn	1	5
Voice	1	5
Years of study on instrument		
3 – 10	9	45
11 – 20	8	40
21 – 30	1	5
Over 30	2	10
Hours/week of practice		
6 – 10	13	65
11 – 20	4	20
21 – 30	1	5
Over 30	2	10
Video games (hours/week)⁴		
1 – 4	7	35
None	13	65

⁴ It is worth noting that all “video game” experience reported for both musicians and controls were on non-action type video games. Thus, this small amount of experience was not a concern, as the only type of video games that have been shown to lead to attentional enhancements are action video games (Green & Bavelier, 2003). Also, the ratio of gaming to non-gaming experience were similar for participants of both groups. Furthermore, the cut-off in video gaming experiments generally only considers a participant to be in the gaming group if they play action video games for more than 4 hours a week (e.g., see Green & Bavelier, 2007).

Table 2.

Control participants questionnaire results

Education	<i>n</i>	Percentage
Currently in college	20	100
<hr/>		
Activities (> 10 hours/week)		
None	8	40
Sports & exercise	8	40
Other (sewing, reading, dancing and drama)	4	20
<hr/>		
Video games (hours/week)		
1 – 4	8	40
None	12	60