Acta Psychologica 147 (2014) 25–33

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

Musical training generalises across modalities and reveals efficient and adaptive mechanisms for reproducing temporal intervals

David Aagten-Murphy [a,*], Giulia Cappagli [a], David Burr [a,b]

[a] Department of Neuroscience, University of Florence, Florence 50125, Italy
[b] CNR Institute of Neuroscience, Pisa 56100, Italy

ARTICLE INFO

Article history:
Received 22 December 2012
Received in revised form 8 October 2013
Accepted 14 October 2013
Available online 31 October 2013

PsycINFO classification:
2320 Sensory Perception
2340 Cognitive Processes

Keywords:
Music
Multisensory
Temporal reproduction

ABSTRACT

Expert musicians are able to time their actions accurately and consistently during a musical performance. We investigated how musical expertise influences the ability to reproduce auditory intervals and how this generalises across different techniques and sensory modalities. We first compared various reproduction strategies and interval length, to examine the effects in general and to optimise experimental conditions for testing the effect of music, and found that the effects were robust and consistent across different paradigms. Focussing on a 'ready-set-go' paradigm subjects reproduced time intervals drawn from distributions varying in total length (176, 352 or 704 ms) or in the number of discrete intervals within the total length (3, 5, 11 or 21 discrete intervals). Overall, Musicians performed more veridical than Non-Musicians, and all subjects reproduced auditory-defined intervals more accurately than visually-defined intervals. However, Non-Musicians, particularly with visual stimuli, consistently exhibited a substantial and systematic regression towards the mean interval. When subjects judged intervals from distributions of longer total length they tended to regress more towards the mean, while the ability to discriminate between discrete intervals within the distribution had little influence on subject error. These results are consistent with a Bayesian model that minimizes reproduction errors by incorporating a central tendency *prior* weighted by the subject's own temporal precision relative to the current distribution of intervals. Finally a strong correlation was observed between all durations of formal musical training and total reproduction errors in both modalities (accounting for 30% of the variance). Taken together these results demonstrate that formal musical training improves temporal reproduction, and that this improvement transfers from audition to vision. They further demonstrate the flexibility of sensorimotor mechanisms in adapting to different task conditions to minimise temporal estimation errors.

1. Introduction

During a concert performance, expert musicians are required to time their actions accurately and consistently. Yet to reach the level where they can do this with great precision involves hundreds of hours of practice and musical training, ostensibly resulting in modifications to the underlying neural processes involved. Indeed, there is increasing evidence that musical training may have long-term influences on seemingly unrelated, non-musical cognitive abilities (Schellenberg, 2001, 2009). Although humans are constantly perceiving and utilising the temporal features of their environment in everyday life, there is a growing literature suggesting that differences can exist between the perceived and actual duration of an event (Eagleman, 2008).

One of the first to identify a bias in temporal perception was Karl von Vierordt (1868; as cited in Wearden & Lejeune, 2008), who reported that humans perceive short and long durations respectively longer and shorter than they really are, with the exception of durations around 700 ms, termed the "indifference point", where representations were veridical (Wearden & Lejeune, 2008). Based on this study, Hollingworth (1910) demonstrated that the location of the indifference point depends on the range of durations employed. For that purpose he stated that Vierordt's Law was not an illusion, but a consequence of the formation of some sort of "central tendency of judgement": according to him, when we estimate the duration of an interval embedded in a series of intervals of different lengths, we tend to estimate and reproduce it as the mean value of the length range of the intervals presented, which corresponds to the indifference point. This means that the estimate of duration is not absolute but relative, because it depends on the context in which stimuli are embedded. For longer durations (1–10 s) it has been demonstrated that the location of the indifference point varies with the tested range of durations presented in the auditory modality but not in the visual modality (Noulhiane, Pouthas, & Samson, 2009), hinting at the prospect of modality specificity. However the authors acknowledge that this may be due to longer intervals before response being more detrimental to visual than auditory memory. Using a reproduction paradigm Jazayeri and Shadlen (2010) hypothesized that this behaviour might be a Bayesian

* Corresponding author at: CNR Institute of Neuroscience, Pisa 56100, Italy. Tel.: +39 050 3153175.
E-mail address: David.AagtenMurphy@gmail.com (D. Aagten-Murphy).

0001-6918/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.actpsy.2013.10.007

strategy to optimise temporal processing abilities: extracting the mean value of a range of presented intervals and relying on it to estimate similar intervals is a useful strategy when perception is uncertain.

Although many of these studies focused primarily on the method of reproduction to demonstrate the central tendency effect for durations, other methods such as recognition (e.g., Hollingworth, 1910) or matching and comparison (e.g., Stevens & Greenbaum, 1966) have been used. Indeed a detailed review of the timing literature by Wearden and Lejeune (2008) concluded that both within the same methodology and when comparing across methodologies there is mixed evidence for a proportional representation of time. While some studies provide clear evidence of scalar variance with temporal judgements, the results of others show non-scalar effects of central tendency (although the presence of central tendency may not exclude scalar variance, see Wearden and Lejeune (2008)). It has been argued that none of those methodologies can claim consistent superiority (Allan, 1979) and that the appropriateness of a method may depend on the range of duration (short or long intervals) and on the paradigm (prospective or retrospective) under investigation (Grondin, 2010) in terms of the accuracy, precision and presence of central tendency in the results. Furthermore, the presentation of the interval can have a large impact on how the interval is timed. Especially in the auditory modality, a clear distinction is made between empty (marked by two brief tones) and filled intervals (containing a continuous tone) with psychophysical studies indicating better accuracy for the timing of empty intervals (Grondin, 2010; Rammsayer & Altenmueller, 2006). As such both the methodology used for response as well as the way the stimulus is defined appears to impact performance on temporal tasks, although the mechanisms through which these changes cause alterations in behaviour are not well understood.

Recently Cicchini, Arrighi, Cecchetti, Giusti, and Burr (2012) extended the Bayesian model proposed by Jazayeri and Shadlen (2010) by examining both audio and visual temporal reproduction performances in skilled drummers, skilled string musicians and non-musicians. By using the variance within these subgroups in their sensory precision they sought to examine how central tendency in responses can optimise performance. They found that skilled percussionists outperform non-musicians in both auditory and visual temporal reproduction tasks, but importantly they demonstrated that subjects were indeed using an adaptive strategy to minimise their total error. This means that when the task was difficult for subjects they tended to bias their estimates towards the average stimulus they had seen previously, resulting in not only a slight decrease in reproduction accuracy but also a (proportionally greater) increase in reproduction precision. This overall resulted in all subjects performing with less total error than would be anticipated from their sensory precision (as measured by an independent temporal bisection task). These results suggest a cognitive mechanism through which musical experience induces a refinement of the neural codes involved resulting in a more precise sensory estimate and a freeing from the necessity to utilise a prior, resulting in more veridical reproduction of time.

Musicians' superior performance is usually limited to aspects of timing that are considered to be automatic and immediately derived from perceptual processing related to temporal information, suggesting that extensive musical training may act to reduce variability or noise associated with time processing (Rammsayer & Altenmueller, 2006). Indeed, compared with people without musical expertise, musicians are often found to have better temporal discrimination (Kraus & Chandrasekaran, 2010; Strait & Kraus, 2011). However, some authors have also observed improvements in overall IQ and other general cognitive abilities (Hyde et al., 2009; Kraus & Chandrasekaran, 2010). Although the exact generalizability of the improvements from musical training and their specificity for particular modalities is contentious, neurodevelopmental studies with musical training in children have consistently suggested that improvements can occur in the visual domain (Bilhartz, Bruhn, & Olson, 1999; Brochard, Dufour, & Despres, 2004; Costa-Giomi, 1999; Rauscher et al., 1997). Furthermore, a recent review (Kraus & Chandrasekaran, 2010) found that improvements to musical ability improve a wide range of different magnitude estimations, both within the auditory domain and across the senses, further suggesting that these processes share some common circuitry.

There are inherent difficulties in inferring causality from correlational studies: it could be that people with better timing skills gravitate towards a music career. However, there is some evidence suggesting that early and long-term musical experience can lead to real permanent morphological and functional cerebral neuro-plastic changes, also associated with improved timing skills (Herholz & Zatorre, 2012; Hyde et al., 2009; Skoe & Kraus, 2012; for a review see Habib and Besson (2009)). To further investigate whether musical training is causal in both neurological and performance-based changes researchers have conducted longitudinal studies randomly assigning children to musical training groups and assessing them periodically over time (Hyde et al., 2009, Moreno et al., 2009). For example, 15 months of intensive musical training was sufficient to induce detectable structural changes in primary auditory and motor areas which were associated with the improvements in auditory and motor skills respectively (Moreno et al., 2009). Overall this suggests that musical training can cause long-lasting alterations in brain structures and functioning through experience and that these changes are specifically related to improvements in functioning.

Additionally there have been consistent findings across a range of studies using neurophysiological and behavioural tasks which indicate that children trained in music perform better in verbal, mathematical and visuospatial tasks (Bilhartz et al., 1999; Brochard et al., 2004; Costa-Giomi, 1999; Rauscher et al., 1997). However, these non-auditory abilities may just reflect short-term advantages during development, as a form of practice enhancement, not necessarily reflecting long-lasting improvements or refinements in abilities that persist into adulthood, even if some authors are reporting that the same brain structural changes persist in adulthood (Hyde et al., 2009). For example, Gaser and Schlaug (2003) found differences in grey matter volume in motor, auditory and visual–spatial areas of the brain when comparing professional keyboardists with amateur and non-musicians. Furthermore, in opposition to the argument that the neurological differences may have simply been due to an innate predisposition, they found strong correlations between the degree of change, level of musical ability and amount of practice, suggesting a more experienced-based role of neural plasticity in these structural differences. The process of musical training has been proposed to be an ideal model for studying neuroplasticity, given the availability of a wide spectrum of subjects with different levels of expertise and the clearly observable changes that occur within the brain (Gaser & Schlaug, 2003; Munte, Altenmuller, & Jancke, 2002).

1.1. The present study

In the present study we aim to investigate multisensory temporal processing in a range of individuals with varying levels of musical expertise to understand the influence that their musical studies have on their temporal reproduction. In particular we will examine the stereotyped bias subjects typically display when reproducing temporal intervals from different distributions wherein, as the task increases in difficulty, subjects will tend to respond with systematic shifts towards the distribution mean, thereby reducing their overall error.

In experiment one we will examine this regression towards the mean with a variety of different temporal stimuli and at different ranges to observe the robustness of the effect and generality between methodologies. Afterwards, in experiment two, we will investigate the reproduction of both auditory and visual events with musicians of different skill levels to examine the effects both in the predominantly trained auditory modality and if there is any transfer to the visual modality. While it may be hypothesized that predominantly auditory musical training will have the greatest effect on auditory performance, studies have often found cross-sensory effects yielding benefits also in the visual modality. Furthermore we will examine two distribution manipulations in order to test the appropriateness of a Bayesian framework in describing the data. Firstly we will change the length of the distribution, while maintaining the mean, in a condition predicted to yield more regression at smaller distributions. Secondly we will change the number of unique values within the distribution to see whether the degree of ability to discriminate or separate the temporal intervals plays a role in the degree of regression observed — a manipulation which, according to recent models (Cicchini et al., 2012; Jazayeri & Shadlen, 2010), should yield no effect. Finally we will examine the proportion of variance within regression index scores that a subject's musical level can explain.

2. Experiment One: Response variation on time reproduction tasks

In this experiment we aim to examine three different techniques for presenting temporal intervals and measuring subject's response. This is to gauge the robustness of the central tendency effect across methodologies and to determine if a particular methodology is better suited for the examination of the impact of different musical levels in experiment two. Three visual temporal reproduction stimulus and response paradigms were examined (sustained stimulus presentation, two interval stimulus presentation and ready-set-go presentation; see method for details) with three different interval distributions (short, medium and long interval distributions).

2.1. Method

2.1.1. Subjects

Seven experimental subjects (2 males, mean age 25 ± 1.7 years) including two professional musicians (with more than 10 years of musical training at the Conservatory of Music in Genoa) took part in the experiment. Subjects had normal hearing and normal or corrected-to-normal visual acuity, and gave their informed consent to participate in the study, which was conducted in accordance with the guidelines of the University of Florence.

2.1.2. Stimuli

Visual stimuli were generated and presented using Matlab version 7.9 and functions provided by the Psychophysics toolbox 3.0 (Brainard, 1997; Pelli, 1997). The visual stimuli were white discs with raised cosine edges and a diameter of 7.7° of visual angle (120 cd/m^2) displayed on a black background 3 cm above fixation and viewed at 57 cm from the screen.

The stimuli were either presented for the entire test stimulus duration, as in the sustained stimulus presentation condition, or for 83 ms during either ready-set-go or two interval presentation conditions. To further eliminate environmental distractions, subjects wore both disposable earplugs (Howard Leight Laser Lite, SNR 35 dB, 32 dB NRR) and earmuffs (Howard Leight, Leightning L1, SNR 30 dB, 25 dB NRR) throughout the experiment.

2.1.3. Procedure

The experiments were performed in a dimly lit room on an Acer TravelMate 6292 13″ (32-bit operating system, 1280 × 800 screen resolution, 60 Hz refresh rate) laptop operating with an external USB keyboard for response. Throughout the experiments timing data was recorded and analysed to ensure that the experiment precisely kept to the required intervals. At the beginning of each trial, subjects were required to fixate on the white square at the centre of the screen and indicate with the space bar when they were ready to proceed, after which the fixation change to a red dot and, after a random delay (0–1.5 s), the first stimulus occurred.

Three different methods of temporal interval presentation and response were examined; sustained stimulus presentation, two interval stimulus presentation and ready-set-go presentation. In sustained stimulus presentation conditions a single visual event was displayed on the screen from the initiation to the end of the test interval. Subjects then, at their leisure, reproduced this single event by holding the space key on the keyboard for the equivalent period of time. For the two event stimulus presentation conditions a brief flash appeared at the beginning and end of the test interval and subjects responded, again at their leisure, by striking the keyboard twice to represent both stimuli and, as such, the interval between them. Intervals presented using the ready-set-go condition had subjects observe a temporal interval demarcated by two brief events, similar to two event presentations, however subjects were immediately required to indicate by a single key-press when a third event should occur in order to replicate interval between the first and second events. Throughout all experimental conditions subjects received no feedback and after response no additional stimuli occurred. Instead after a 0.5–1.5 s delay the fixation reverted back to a white square prompting subjects to initiate the next trial when they were ready.

The three distributions examined were selected to be similar to those examined by both Jazayeri and Shadlen (2010) and Cicchini et al. (2012) in their investigations of temporal interval reproduction. As such, in this initial experiment, we examined three partially overlapping distributions, each spanning a range of 352 ms and divided into 11 equidistant sample intervals, based around a different mean (671, 847 or 1023). This resulted in three test distributions that entailed a short distribution (494–847 ms), an intermediate distribution (671–1023 ms) and a long distribution (847–1200 ms). We examined each of these distributions in the three different presentation and response conditions of sustained stimulus, two event and ready-set-go presentations, allowing us to examine the generality of this behaviour on reproduction tasks examined in different ways.

The experiment was broken up into multiple sessions with each session lasting approximately 10 min and comprising 11 intervals of a specific distribution (746 ms, 816 ms, 845 ms…etc.) being examined 10 times each for each experimental condition (short, medium or long). This was repeated 3 times for a total of 330 trials for each of the three distributions in each experimental condition. Thus, with each condition consisting of three distributions (and hence 990 trials), subjects performed just under 3000 trials in total across the 3 conditions examined. The order of testing was randomised across both distribution type and methodology; however, subjects were prevented from repeating the same interval and paradigm on the same day. Overall subjects underwent 3–5 days of testing.

2.2. Analysis

Subject's responses were analysed with respect to the presented value by taking the average of the reproductions for each sample interval. In order to prevent outliers or lapses in attention any responses more than 4 standard deviations from their mean were excluded before further analysis. For each distribution and each condition the best linear fit, inclusive of a constant bias parameter, was calculated and the difference in slope between this and the equality line is defined as the regression index (RI). This metric provides a measure of the degree of linear regression where 0 represents veridical performance and 1 represents complete regression to the mean.

The data was additionally analysed with a 2-factor repeated measures ANOVA to examine both within-subject effects for the three distributions (short, medium or long) or three experimental conditions (sustained condition, two event condition or ready-set-go) and between-subject effects related to their musical expertise (Non-Musician or Expert Musician).

2.3. Results and discussion

Fig. 1 shows the average reproduction times for the 5 Non-Musicians (Fig. 1A) and the 2 Expert Musicians (Fig. 1B), for the three different

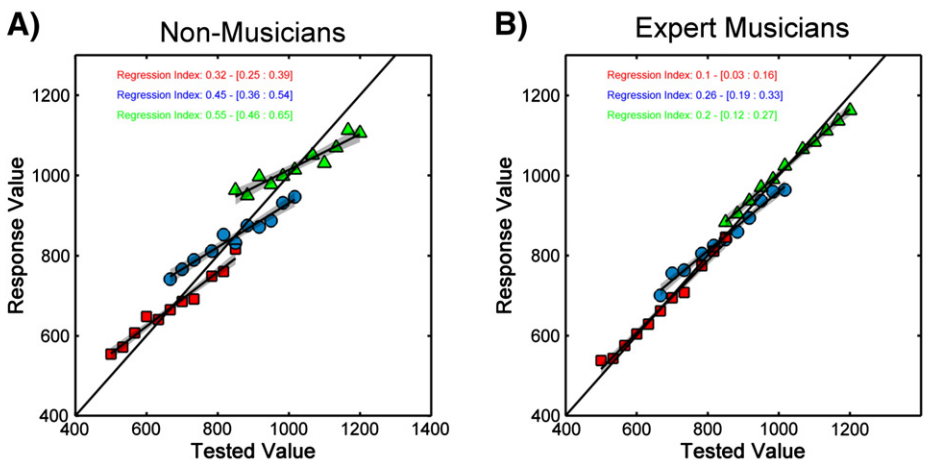

Fig. 1. The presented and reproduced intervals for Non-Musicians (A) and Musicians (B) within three different distributions for the ready-set-go paradigm. As there was no significant variances between the paradigms this provides an accurate account of performance for all paradigms. As can be seen, Non-Musicians exhibited significantly more regression towards the mean (as shown by the tendency away from the veridical) than Musicians across all ranges. Furthermore there is a tendency for the degree of regression to increase as the distribution it is drawn from increases. Errors bars represented the standard error between subjects.

distributions tested, with error bars representing the standard error between subjects. Although the focus of experiment one was on the difference in methodologies, even with this small sample group results were remarkably consistent.

As has been previously reported, Non-Musician subjects showed substantial regression towards the mean, as evidenced by the relatively large regression indexes, and by the fact that as the mean of the distribution increased so did the magnitude of their regression. In contrast, the Expert Musicians showed negligible regression to the mean with low regression indexes and data that closely follows the veridical. Importantly, the 850 ms temporal interval occurred in all 3 distributions and as such can be directly used to examine the influence of the distribution context on temporal reproduction. While there is less than 50 ms separating the Expert Musicians' average response across the 3 distributions, the Non-Musicians have more than 100 ms of difference between their performance on this identical temporal interval, depending solely on the context of the distribution from which it was drawn (short, medium or long). This result is again in keeping with previous studies (Cicchini et al., 2012; Jazayeri & Shadlen, 2010), which also found that the distributions with a higher mean show more regression, in keeping with the scalar property of temporal estimation whereby the noise of the task increases as the mean increases resulting in a stronger reliance on regression to the mean to keep reproduction noise minimal.

The mean regression index for the two groups for each experimental condition and for all three distributions is illustrated in Fig. 2. As can be seen, there is very little difference between the three conditions at any of the three distribution means for all subjects. Indeed a repeated measures ANOVA revealed no significant main effect of experiment ($F(2,10) = 0.397$, $p > 0.05$) indicating that the subject's regression towards the mean was independent of the method used. Furthermore, this tendency is preserved even when the ANOVA excludes the two Expert Musicians and considers only the 5 Non-Musician subjects ($F(2,8) = 0.881$, $p \geq 0.05$). The main effect for distribution approached, but did not reach, significance ($F(2,10) = 3.599$, $p = 0.06$) indicating that overall the regression index did not differ significantly between the small, medium and large distributions. However, as Fig. 2 shows, all subjects showed less regression on the shortest distribution, while their performances on the medium and long distributions were almost identical.

That all three methodologies of stimulus presentation and response yielded near-equivalent degrees of regression from our subjects highlights the robustness of this regression effect in temporal reproduction and suggests that, at least for the methodologies tested, there are no

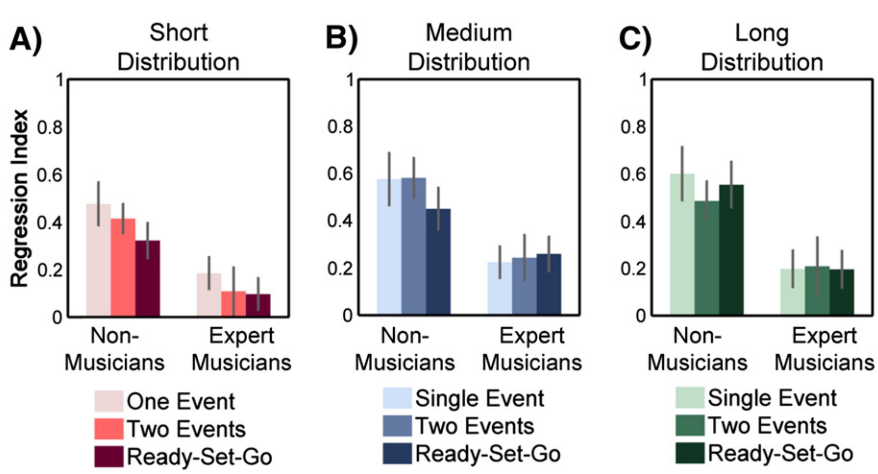

Fig. 2. A comparison of the three different paradigms for the short (A), medium (B) and long (C) distributions. Different shaded bars represent different paradigms while the two groups represent Non-Musicians and Musicians respectively. As can be seen, there is little variation between the paradigms but a substantial and consistent difference between the skilled and Non-Musicians at all distribution lengths. Errors bars represented the standard error between subjects.

substantial differences in the degree to which subjects demonstrate central tendency. For the second experiment, we focus on the ready-set-go paradigm, which has the advantage of being the fastest paradigm to test and the only 'time pressured' task in which subjects must respond immediately (as the other methods allow the subject to initiate the reproduction at leisure). This not only contributes to reducing the overall time it takes to test subjects, but also helps minimise any memory-based effects by preventing subjects from pausing for variable periods before responding. Furthermore, the task involves a single motor response instead of two (as the other methods), making it overall a simpler and more controlled task for analysis.

In the second experiment we aim to both investigate the role of musical experience on temporal reproduction of both auditory and visual intervals and further investigate the properties of the test distribution by examining this effect within the bounds of the model proposed by Cicchini et al. (2012).

3. Experiment Two: The role of musical expertise on audio and visual temporal reproduction

This is the main experiment, aimed at explicitly investigating the role musical expertise has on both visual and auditory temporal reproductions and, more specifically, on the tendency for subject to utilise knowledge of test distributions to reduce errors. As the experimental conditions investigated in experiment 1 proved equivalent, for the second study we focus on the ready-set-go condition as this paradigm was both faster and marginally more consistent between subjects than the other conditions. Furthermore, we will examine the model proposed by Cicchini et al. (2012) looking at how the conditions of the experiment influence the degree of regression index. To this end we will focus on the medium temporal distribution from experiment 1 centred on 847 ms.

We will examine both manipulations in the length of the distribution, where a shorter distribution would be expected to yield more regression, and the role the number of discrete temporal intervals has within a distribution, which according to the Bayesian framework (see Eq. 12; Cicchini et al., 2012) would have no effect on the degree of regression.

3.1. Method

3.1.1. Subjects

We tested 33 subjects (14 males, mean age 23.7 ± 4.4 years) of which 11 were Non-Musicians (2 males, mean age 26.9 ± 4.6) and 22 were trained Musicians. Of these 22 musicians, 11 were determined to be Intermediate Musicians (3–8 years of musical training; 5 males, mean age 20.8 ± 3.6 years) while 11 were Expert Musicians (9–12 years of formal musical training; 6 males, mean age 23.7 ± 3.6). Subjects considered Non-Musicians had either had no musical experience (8 subjects) or had less than 3 years of casual musical training without any qualifications and had not returned to an instrument in more than 7 years (3 subjects). All musicians were students of the Conservatory of Music in Genoa where they were receiving formal training in music (with their years of training used as a proxy for musical ability). The different categories of instruments the Musicians specialised in are shown in Table 1, with Musicians belonging to four groups playing wind, string or percussion instruments or the piano. Furthermore all subjects had normal hearing and normal or corrected-to-normal visual acuity. Six subjects (4 Non-Musicians and 2 Musicians) previously participated in experiment 1.

Table 1
The number of Intermediate and Expert Musicians and the class of instrument that they play. Subjects were categorised as Intermediate Musicians if they had more than 3, but less than eight, years of formal music training, while those with 9 or greater were classified as Expert Musicians.

	Intermediate Musicians	Expert Musicians
Strings	5	2
Wind	2	3
Percussion	2	3
Piano	2	3
Total	11	11

3.1.2. Stimuli

The experimental stimuli were identical to those of experiment one with the addition of acoustic stimuli in the auditory condition. The acoustic stimuli were pure tones of 400 Hz presented for 83 ms (5 frames) with transitions smoothed by a 10 ms raised cosine filter and digitalised at 45 kHz and were listened to over headphones (Sony MDR-E818LP) with an average intensity of 75 dB measured at the sound source.

3.1.3. Procedure

Experiment two followed the same experimental procedure as experiment one, however here we focused on just the intermediate distribution, which spanned 352 ms with a mean of 847 ms and was divided into 11 equidistant samples (between 671 and 1023 ms). In experiment two, instead of manipulating the mean as in experiment one, we manipulated the range (176 ms, 352 ms, 704 ms all with 11 samples) or the number of samples within this range (5, 11, 21 in a range of 352 ms) with 352 ms and 11 samples being the common standard to both manipulations. Finally, we also examined all 5 conditions in a visual condition, when the intervals were defined by visual flashes (as the ready-set-go conditions of experiment one) and in an auditory condition, when the intervals were defined by auditory tones. As such we had 12 conditions in total, 6 in vision and 6 in audition, examining 3 different ranges and 3 different sampling densities within the distributions. Importantly the central condition was identical to the intermediate distribution utilised in experiment 1 and in previous temporal reproduction studies (Cicchini et al., 2012; Jazayeri & Shadlen, 2010).

The presentation of the 12 conditions was randomly intermingled, such that each subject performed each range, sampling density and modality in a different order to minimise any learning effects in the data, and repeated 3 times for a total of 330 trials per condition. As in experiment one the experiment was conducted in a dimly lit room and subjects wore earmuffs (Howard Leight, Leightning L1, SNR 30 dB, 25 dB NRR) throughout the experiment, either over the top of their headphones during auditory conditions or over the top of disposable earplugs (Howard Leight Laser Lite, SNR 35 dB, 32 dB NRR) during visual conditions.

3.2. Analysis

The analyses for experiment 2 followed that of experiment one. The data for the two conditions, manipulating distribution length or manipulating the number of samples in a distribution, were measured in two separate 2-factor repeated measures ANOVAs.

Each ANOVA examined within-subject effects for two stimulus modalities (auditory or visual) and the three levels of the experimental condition as well as between-subject effects related to their musical expertise (Non-Musician, Intermediate Musician or Expert Musician) or their musical instrument. Finally, the relationship between musical expertise and both visual and auditory temporal reproductions was analysed through a linear regression investigating the proportion of variance that musical expertise can explain within subjects reproduction data (calculating the variance with the adjusted-R2).

3.3. Results

Fig. 3 shows the average reproduction times for the Non-Musicians, Intermediate Musicians and Expert Musicians on the auditory (Fig. 3A) and visual (Fig. 3B) temporal reproduction tasks, for the common intermediate distribution (as shown in experiment 1). As the figure shows, there were substantial differences between the different musical groups, with Expert Musicians showing less regression in both modalities than Non-Musicians. Further, it can be seen that all subjects exhibited more tendency to regress towards the mean in the visual modality than in the auditory condition.

Fig. 4A shows the first experimental condition which involved manipulating the size of the interval distribution while maintaining a fixed mean. The repeated measures ANOVA showed a highly significant main effect for modality ($F(1,24) = 63.001$, $p < 0.01$). All interactions were non-significant with the exception of a distribution size and musical expertise interaction ($F(2,48) = 3.041$, $p < 0.05$). Bonferroni-corrected pair-wise tests revealed that as the distribution size decreases the regression index increases significantly for both the large to medium ($F(1,24) = 6.199$, $p < 0.05$) and medium to small ($F(1,24) = 7.238$, $p < 0.05$) comparisons. However, this interaction between distribution size and musical expertise was driven by variation in the amount of regression different Musicians have at the smallest distribution size. The analysis of between-subject effects revealed a significant ($F(1,24) = 6.321$, $p < 0.05$) effect of musical expertise group on regression but no effect of musical instrument type ($F(2,23) = 0.083$, $p > 0.05$) and Bonferroni-corrected tests reveal that this is driven by significant differences between the Expert Musicians and both the Intermediate and Non-Musicians ($p < 0.05$) but not between Intermediate and Non-Musicians ($p > 0.05$). Thus it appears that the length of the test distribution has a substantial influence on the degree of regression towards the mean subjects exhibited.

Fig. 4B shows the mean regression index for Musicians and Non-Musicians varying the number of samples defining each distribution, from 5 to 7 to 21. From this figure and the repeated-measures ANOVA we can see that there is no effect of the number of samples ($F(2,48) = 0.6$, $p > 0.05$), with subjects showing identical degrees of regression towards the mean regardless of condition, but again a significant effect of stimulus modality ($F(2,24) = 32.761$, $p < 0.001$), with no significant interactions. Within-subjects there is a strong effect for the musical expertise group ($F(1,24) = 6.852$, $p < 0.05$) and again no effect for musical instrument ($F(3,24) = 0.017$, $p > 0.05$). Once again pair-wise comparisons showed that Expert Musicians were substantially different from Intermediate and Non-Musicians ($p < 0.5$), while Non-Musicians and Intermediate Musicians were not significantly different ($p > 0.05$). Thus overall there was no effect of the density of stimuli within a distribution on the degree of regression towards the mean subjects exhibited. Furthermore, across both experiments while the level of musical expertise was a strong predictor of the degree of regression, type of musical instrument was not.

To investigate this point further each individual's average central tendency (across conditions differing in sampling density) was regressed with the continuous measure of their musical level in years of formal musical training. The results shown in Fig. 6 for auditory (blue) and for visual (red) modalities demonstrate a substantial and significant contribution with years of musical expertise predicting from 37% ($F(1,20) = 11.836$, $p < 0.013$) to 22% ($F(1,20) = 5.763$, $p < 0.05$) of the degree to which they exhibit regression on a temporal interval task for audition and vision respectively. Furthermore, it appears that undergoing primarily auditory-based musical training yields gradual improvements in auditory reproduction abilities but, interesting, further training seems to be needed before concurrent improvements in vision occur.

Overall the results show that individuals with musical training, particularly in the acoustic domain, show substantially less regression towards the mean and more veridical performance on a temporal reproduction task. While the length of the distribution is important, the number of stimuli within this length is not, two results which are in line with a Bayesian model of central tendency. Furthermore these results are suggestive of plastic changes occurring as musicians acquire expertise, which is at least partly responsible for the reduced reliance on regression strategies to minimise their error, as their own increased precision allows them to respond more veridical and rely less on central tendency.

4. Discussion

In this study we examined central tendency of time reproduction with multiple paradigms and manipulations of the testing distribution. To ensure that the reproduction effects are relatable between different studies we first attempted to determine if stimulus and response patterns yield any substantial differences in performance. We found that, regardless of methodology, subjects showed the same degree of systematic regression towards the mean of the distribution. Given that the recent papers by Jazayeri and Shadlen (2010) and Cicchini et al. (2012) used different methodologies, it was important to determine if the rates of central tendency are stable across paradigms in order to focus on the differences due to music ability in experiment two. For a small sample, we tested each methodology multiple times with 3 different distributions to obtain a robust estimate of central tendency.

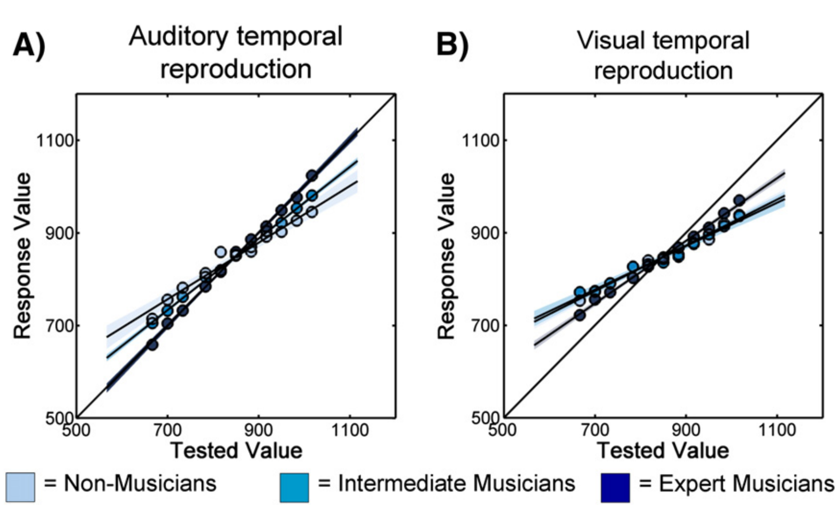

Fig. 3. The presented and reproduced intervals for Non-Musicians (light blue), Intermediate Musicians (blue) and Expert Musicians (dark blue) for audition (A) and vision (B). As can be seen, with increased expertise subjects perform with less bias (and more veridical) in both modalities, however overall all subjects perform the task better within the auditory modality than the visual modality. Errors bars represented the standard error between subjects.

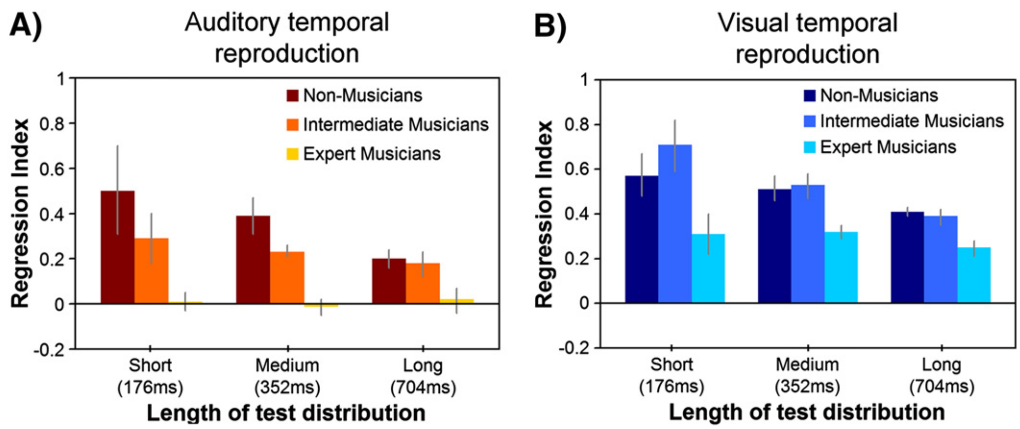

Fig. 4. Regression index for Non-Musician, Intermediate Musicians and Expert Musicians for audition (A) and vision (B) modalities and for short, medium and long test distribution lengths. While musical expertise always results in more veridical performance the results demonstrate that short intervals induce more regression towards the mean in all subjects than longer intervals. Errors bars represented the standard error between subjects.

We found a remarkable consistency between subjects across the three conditions and no-significant group differences. This was despite the ready-set-go paradigm requiring only a single motor command (as opposed to the other requiring two to initiate and end the reproduced interval) and requiring an immediate response (while the other two were self-initiated and possibly involved short-term memory processes). Furthermore, only the sustained response methodology utilised filled stimulus intervals which, unlike the 'empty' stimulus intervals of the other two methodologies, are usually found to be less accurate than the timing of empty intervals (Grondin, 2010, Rammsayer & Altenmueller, 2006). However the results demonstrated similar levels of central tendency with, if anything, a slight but non-significant decrease in regression index for the filled stimulus interval condition. This finding demonstrates the robustness of this effect as it generalised across various similar tasks and, in light of this general similarity, the simpler (and more rapid) ready-set-go paradigm was chosen for use in the second part.

Two different manipulations of the test distribution, which theoretically should result in differences in the prior, were studied in experiment two to further examine the stereotyped patterns of error. The first was changing the length of the distribution so the 11 different intervals stretched across a wider range of times. This results in the same mean but a larger standard deviation as the range increases. In terms of the Bayesian framework, with all else equal, this results in a broader, less strong prior being formed and as such a greater reliance on the sensory estimate yielding less central tendency. This becomes intuitive when considering that as the distribution length is decreased it becomes more similar to the mean of the distribution and hence making a response at the mean becomes an increasingly accurate prediction. As demonstrated by Cicchini et al. (2012; Eq. 12) the optimal prior width can be demonstrated as:

$$\hat{\sigma}_P^2 = \begin{cases} 2\sigma_S^2 - \sigma_L^2 & \text{for} \quad \sigma_L \leq \sqrt{2}\sigma_S \\ 0 & \text{otherwise} \end{cases}$$

where both the prior and likelihood function are Gaussians with mean and standard deviations (μ_P,σ_P^2) and (μ_L,σ_L^2) respectively. This predicts that the influence of the prior, and hence the degree of central tendency, should be stronger for both smaller ranges and when subjects' sensory representations are imprecise (large σ_L^2), while invariant to changes in distribution density. As was anticipated, increasing the interval length increased the amount of regression for all levels of musical experience. However, altering the range of the distribution while maintaining 11 discrete intervals within each distribution means that the ability to discriminate between any neighbouring duration within the distribution also increases. Although intuitively it may be thought that this allows responders to create a less overlapped and hence more clear representation of the distribution, the models proposed by Jazayeri and Shadlen (2010) and Cicchini et al. (2012) suggest that only the mean and standard deviation of the distribution are important — both of which remain fixed with this manipulation. Hence three different numbers of samples (matching the spacing found within the previous

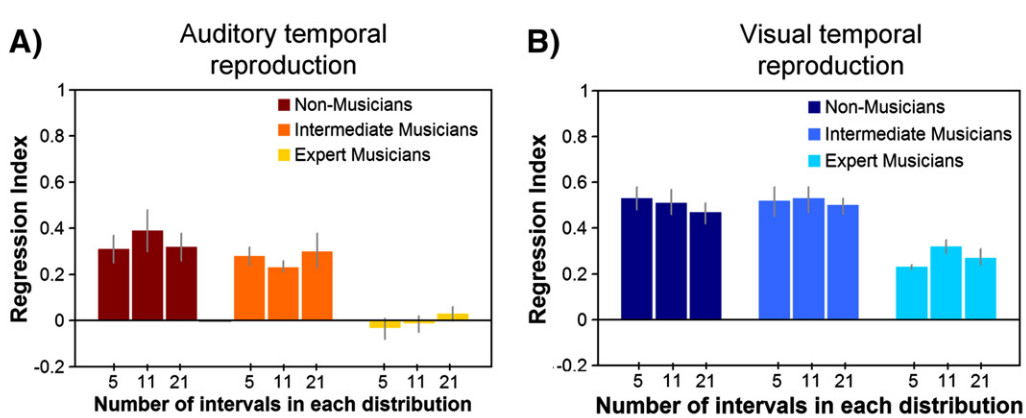

Fig. 5. Regression index for Non-Musician, Intermediate Musicians and Expert Musicians for audition (A) and vision (B) modalities and short, medium and long test distribution lengths. While musical expertise always results in more veridical performance the results demonstrate that short intervals induce more regression towards the mean in all subjects than longer intervals. Errors bars represented the standard error between subjects.

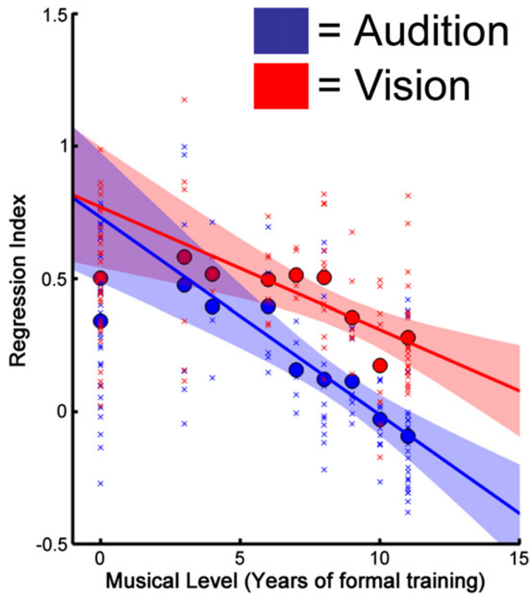

Fig. 6. The amount of regression towards the mean as a function of the number of years of musical training for audition (blue) and vision (red). While there is a significant reduction in the regression index with increasing expertise for both modalities, audition is decreasing at a faster rate than vision, suggesting that while musical expertise reduces systematic bias for both modalities, it does so preferentially in the audition. Errors bars represented the standard error between subjects.

distribution length manipulations) were also examined. The data supports the model as the number of samples was found to have no influence on the degree of regression towards the mean. Furthermore we can confirm that the previous effect of the size of the range is due solely to the magnitude of the range and not dependent on the distribution density. The consistency of the responses, with the three density distributions thus being essentially three independent examinations of the effect within the same test distribution, further illustrates the robustness of the central tendency effect and subjects' consistency in their strategies when approaching these tasks. Finally, these results support previous studies with Non-Musicians showing clear evidence for a systematic regression towards the mean of their test distribution — particularly when dealing with their own noisy estimates of long temporal intervals (Cicchini et al., 2012; Jazayeri & Shadlen, 2010). An intuitive way of understanding what is happening is that when faced with a difficult task an estimate of the mean will be less noisy than any individual estimate, so incorporating this estimate can, under the right circumstances, reduce overall noise.

Within the Bayesian framework of the reproduction task, while the previous manipulations dealt with the specifics of the prior, selecting subjects with different levels of musical expertise allows for a population with predictably changing sensory abilities (Gaser & Schlaug, 2003; Munte et al., 2002). Indeed musicians are frequently found to show enhanced performance particularly in auditory or temporal tasks (Kraus & Chandrasekaran, 2010; Strait & Kraus, 2011). Recently Cicchini et al. (2012) found that skilled percussionists outperform non-musicians in both auditory and visual temporal reproduction tasks, but importantly they demonstrated that subjects were indeed using an adaptive strategy to minimise their total error. This means that when the task was difficult for subjects they tended to bias their estimates towards the average stimulus they had seen previously, resulting in not only a slight decrease in reproduction accuracy but also a (proportionally greater) increase in reproduction precision. This overall resulted in all subjects performing with less total error than would be anticipated from their sensory precision. Although previously Cicchini et al. (2012) reported that the degree of bias differed in musicians according to the instrument they played, here we instead demonstrated a large role of musical expertise (as measured by years of formal training). Indeed across the different levels of musical knowledge there was little difference attributable to the musical instrument studied, although a larger number of subjects varying in both main instrument and in expertise would be needed to formally understand the relationship between the two. Furthermore, other factors such as whether the musicians are training and performing as soloists or as concert players may also prove influential in the degree of regression behaviour they exhibit, as this also dictates the importance of precise temporal performance and as such the importance to their further acquisition of musical expertise. Regardless, simply knowing the musical level of a musician alone was sufficient to explain 37% of their variance in the performance on the auditory temporal task.

Interestingly there is also a significant benefit in temporal performance with visual stimuli dependent on musical knowledge, with expertise explaining 22% of the variance in the amount of regression to the mean on the visual temporal reproduction task. Although substantially less, this does suggest that predominantly auditory musical training is able to improve visual temporal abilities and lends support to the idea of estimation areas being partially multisensory or at least the neural underpinnings relying on shared circuits (Noulhiane et al., 2009). This is in agreement with the commonly reported finding that for temporal estimates audition performs better than vision system (Aschersleben & Bertelson, 2003; Burr, Banks, & Morrone, 2009; Fendrich & Corballis, 2001; Vroomen & de Gelder, 2004). Although a caveat of any correlation studies is to the inability to directly infer causality, that musical training is causal has been supported from other studies demonstrating long-lasting effects of musical training (Habib & Besson, 2009; Herholz & Zatorre, 2012, Hyde et al., 2009; Skoe & Kraus, 2012) and with longitudinal studies where musical training was randomly-assigned and performance measured several months later (Hyde et al., 2009; Moreno et al., 2009). Indeed Moreno et al. (2009) found, in a study with 6 months of randomly-assigned art or music training in 8 year old children, that enhanced reading and pitch discrimination abilities in speech were presented only after musical training. This suggests that musical training may indeed have a unique, causal role in the improvement of other abilities. However, future studies investigating musical training would benefit from the inclusion a control group of individuals highly trained in another domain (such as, for example, a spatial discrimination task) in order to more explicitly demonstrate the unique contribution of musical training.

Interestingly, while Intermediate Musicians were still showing less regression than Non-Musicians for auditory intervals, they performed similarly to Non-Musicians for visual intervals. Examining Fig. 5, which shows the degree of regression against years of experience, it is clear that Musicians show improvement in audition much sooner than they do in vision. Although it would be expected that predominantly auditory temporal training would benefit auditory reproductions more, it is interesting to observe this delay before benefits extend across modalities. A recent study by Noulhiane and colleagues (2009) demonstrated that for longer durations (1–10 s) there is some suggestion of cognitive processes differentially involved in the reproduction of visual and auditory durations, with the indifference point varying with the tested range of durations presented in the auditory modality but not in the visual modality. These studies are important in the context of the neurodevelopmental studies of the role of musical training, which have hinted at improvements in the visual domain from musical training in children (Bilhartz et al., 1999; Brochard et al., 2004, Costa-Giomi, 1999; Rauscher et al., 1997). Furthermore, a recent review (Kraus & Chandrasekaran, 2010) outlines the evidence for permanent effects on the brain and neural system from a variety of studies including longitudinal and imaging studies, as well as evidence for an enhancement favouring identifying patterns and detecting regularities. Together with findings demonstrating a wide range of different magnitude estimations showing improvement with musical ability, both within the auditory domain and across the senses, this lends support to the idea that these processes share some common circuitry.

That all three presentations and reproduction conditions yielded equivalent degrees of regression from our subjects highlights the robustness of this regression effect in temporal reproduction. However, here we have found exceptionally good performance even in Musicians of different backgrounds. This suggests that Cicchini et al.'s (2012) conclusion that instrument plays a strong role in the degree of regression towards the mean may require further investigation, as their sample of musicians specialising in different instruments may also have varied in musical training.

5. Conclusion

Thus we have demonstrated the remarkable robustness of the central tendency effect that persists despite variations in stimulus presentation and reproduction requirements. Furthermore, we have shown a strong association between musical expertise and both auditory and visual reproduction abilities, which appear to be not based on the type of instrument (Cicchini et al., 2012) but instead on the level of musical expertise. That these plastic changes can generalise across modalities lends further support to the argument that there is an amodal component to at least pattern-analysing aspects of temporal perception. Finally further support was found, within an experimental design that included variations to prior strength (through specific distribution manipulations) and subjects' sensory ability (by selecting subjects based on musical expertise), that temporal reproduction behaviour and central tendency can be accurately modelled within a Bayesian framework. Overall these results demonstrate the plasticity occurring as musicians acquire expertise, with their improved sensory estimate allowing for a reduced reliance on regression strategies, and the flexibility of sensorimotor mechanisms in adapting to different task conditions in order to minimise temporal estimation errors.

Acknowledgment

This research was supported by the Italian Ministry of Universities and Research (PRIN2009) and by ERC project "STANIB" (FP7 ERC).

References

Allan, L. (1979). The perception of time. *Perception & Psychophysics*, *26*, 340–354.

Aschersleben, G., & Bertelson, P. (2003). Temporal ventriloquism: crossmodal interaction on the time dimension — 2. Evidence from sensorimotor synchronization. *International Journal of Psychophysiology*, *50*, 157–163.

Bilhartz, T. D., Bruhn, R. A., & Olson, J. E. (1999). The effect of early music training on child cognitive development. *Journal of Applied Developmental Psychology*, *20*, 615–636.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Brochard, R., Dufour, A., & Despres, O. (2004). Effect of musical expertise on visuospatial abilities: Evidence from reaction times and mental imagery. *Brain and Cognition*, *54*, 103–109.

Burr, D., Banks, M. S., & Morrone, M. C. (2009). Auditory dominance over vision in the perception of interval duration. *Experimental Brain Research*, *198*, 49–57.

Cicchini, G. M., Arrighi, R., Cecchetti, L., Giusti, M., & Burr, D. C. (2012). Optimal encoding of interval timing in expert percussionists. *Journal of Neuroscience*, *32*, 1056–1060.

Costa-Giomi, E. (1999). The effects of three years of piano instruction on children's cognitive development. *Journal of Research in Music Education*, *47*, 198–212.

Eagleman, D.M. (2008). Human time perception and its illusions. *Current Opinion in Neurobiology*, *18*, 131–136.

Fendrich, R., & Corballis, P.M. (2001). The temporal cross-capture of audition and vision. *Perception & Psychophysics*, *63*, 719–725.

Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *Journal of Neuroscience*, *23*, 9240–9245.

Grondin, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception, & Psychophysics*, *72*, 561–582.

Habib, M., & Besson, M. (2009). What do music training and musical experience teach us about brain plasticity? *Music Perception*, *26*, 279–285.

Herholz, S.C., & Zatorre, R. J. (2012). Musical training as a framework for brain plasticity: Behavior, function, and structure. *Neuron*, *76*, 486–502.

Hollingworth, H. L. (1910). The central tendency of judgment. *The Journal of Philosophy, Psychology and Scientific Methods*, *7*, 461–469.

Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A.C., et al. (2009). Musical training shapes structural brain development. *Journal of Neuroscience*, *29*, 3019–3025.

Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature Neuroscience*, *13* (1020-U152).

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews. Neuroscience*, *11*, 599–605.

Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex*, *19*, 712–723.

Munte, T. F., Altenmuller, E., & Jancke, L. (2002). The musician's brain as a model of neuroplasticity. *Nature Reviews. Neuroscience*, *3*, 473–478.

Noulhiane, M., Pouthas, V., & Samson, S. (2009). Is time reproduction sensitive to sensory modalities? *European Journal of Cognitive Psychology*, *21*, 18–34.

Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

Rammsayer, T., & Altenmueller, E. (2006). Temporal information processing in musicians and nonmusicians. *Music Perception*, *24*, 37–47.

Rauscher, F. H., Shaw, G. L., Levine, L. J., Wright, E. L., Dennis, W. R., & Newcomb, R. L. (1997). Music training causes long-term enhancement of preschool children's spatial–temporal reasoning. *Neurological Research*, *19*, 2–8.

Schellenberg, E. G. (2001). Music and nonmusical abilities. In R. J. Zatorre, & I. Peretz (Eds.), *Biological foundations of music*.

Schellenberg, E. G. (2009). Music training and nonmusical abilities: Commentary on Stoesz, Jakobson, Kilgour, and Lewycky (2007) and Jakobson, Lewycky, Kilgour, and Stoesz (2008). *Music Perception*, *27*, 139–143.

Skoe, E., & Kraus, N. (2012). A little goes a long way: How the adult brain is shaped by musical training in childhood. *Journal of Neuroscience*, *32*, 11507–11510.

Stevens, S. S., & Greenbaum, H. B. (1966). Regression effect in psychophysical judgment. *Perception & Psychophysics*, *1*, 439–446.

Strait, D., & Kraus, N. (2011). Playing music for a smarter ear: Cognitive, perceptual and neurobiological evidence. *Music Perception*, *29*, 133–146.

Vroomen, J., & de Gelder, B. (2004). Temporal ventriloquism: Sound modulates the flash-lag effect. *Journal of Experimental Psychology. Human Perception and Performance*, *30*, 513–518.

Wearden, J. H., & Lejeune, H. (2008). Scalar properties in human timing: Conformity and violations. *Quarterly Journal of Experimental Psychology*, *61*, 569–587.