Musicians Do Better than Nonmusicians in Both Auditory and Visual Timing Tasks

THOMAS H. RAMMSAYER
University of Bern, Bern, Switzerland

FRANZISKA BUTTKUS, & ECKART ALTENMÜLLER
University of Music, Drama and Media, Hannover, Germany

The present study was designed to investigate differences in auditory and visual temporal information processing between musicians and nonmusicians. For this purpose, timing performance on a set of six different psychophysical temporal tasks for both the auditory and visual sensory modalities was compared in 40 formally trained musicians and 40 controls without musical experience. Across modalities, superior temporal acuity for musicians compared to nonmusicians could be shown for all temporal tasks except for temporal generalization. When comparing the two sensory modalities, temporal acuity was superior to auditory stimuli as compared to visual stimuli, with the exception of the temporal generalization task in the 1-s range. The overall pattern of our findings is consistent with the notion that musicians’ long-lasting intensive music training, starting in childhood, improves general timing ability irrespective of sensory modality.

Received July 5, 2011, accepted December 21, 2011.

Key words: interval timing, rhythm perception, auditory, visual, sensory modality

The definition of musical ability has always been a highly disputed issue (for a review see Sloboda, 2005). There is no clear agreement on a definition yet (Bentley, 1966; Colwell, 1970; Lundin, 1967; Radocy & Boyle, 1979; Shutery-Dyson, 1999), although good timing ability is considered explicitly or implicitly as an important attribute of musical ability. Musical aptitude tests often include temporal discrimination or rhythm perception tasks, and are based on various aspects of auditory performance and tonal and rhythmic concepts (e.g., Drake, 1954; Kwalwasser, 1953; Seashore, 1919). For example, the Seashore Measures of Musical Talents (Seashore, Lewis, & Saetveit, 1956) were intended to assess musical aptitude by means of various tests referred to as Sense of Pitch, Sense of Intensity, Timbre, Tonal Memory, Sense of Time, and Sense of Rhythm. The construct validity of such musical aptitude tests has been doubted by several researchers (e.g., Anastasi, 1961; Henson & Wyke, 1982; Sloboda, 2005; Winner & Martino, 1993). One major objection refers to the fact that musical aptitude tests often assess isolated sensorimotor abilities rather than considering the complexity of musical abilities, such as emotional appreciation, which is required in a real musical context (Haroutonian, 2000; Henson & Wyke, 1982; Rainbow, 1965; Sloboda, 1985, 2005).

A positive functional relationship between musical ability and performance on temporal information processing is supported by various findings of superior timing accuracy in musicians compared to nonmusicians. Many of these findings were obtained with tasks based on the reproduction or production of temporal patterns (Aschersleben 2002; Drake, Penel, & Bigand, 2000; Repp, 2006). In these studies, temporal information processing was investigated in combination with motor skills, showing a good ecological validity for music performance skills. It is possible, therefore, that the superior performance on timing and time perception tasks observed with musicians largely depends on the strong central nervous coupling of auditory and sensory-motor representations.

In music performance, musicians frequently attain extremely precise timing control. Wagner (1971) assessed the rhythmic precision of playing a C-major scale in professional pianists. At a required speed of about six key-strokes per s, a standard deviation of 6 to 10 ms was found when calculating the temporal deviations of 30 subsequent keystrokes. An even higher degree of regularity of cyclic trill movements in a pianist was reported by Moore (1992) using MIDI-technology. Another study showed that cumulative practice time of
professional percussionists was correlated with the ability to produce regular sequences of synchrony (Trappe, Katzenberger, & Altenmüller, 1998). There is also a less pronounced negative asynchrony in professional musicians compared to nonmusicians when asked to tap in beat with a metronome (Aschersleben, 2002).

Investigation of temporal information processing should also include tasks of a purely perceptual nature. There are only few studies based on perception tasks without any motor component. Two studies by Jones and Yee (1997) and Yee, Holleran, and Jones (1994) found that musicians showed superior performance in detecting small time changes embedded in regular auditory sequences, although these results may apply only to specific aspects of temporal judgments. In another study, when musicians were asked to judge whether the last of six regular sequential auditory or tactile stimuli occurred 50 or 25 ms later or earlier, musicians did not perform better than nonmusicians (Lim, Bradshaw, Nicholls, & Altenmüller, 2003). A study by Rammsayer and Altenmüller (2006) also investigated temporal information processing of a purely perceptual nature in musicians and nonmusicians. Seven different auditory psychophysical timing tasks without a motor component were included. Superior temporal acuity for musicians compared to nonmusicians was shown for auditory fusion, rhythm perception, and three temporal discrimination tasks. The two groups did not differ, however, in terms of their performance on two tasks of temporal generalization. Unlike the other timing tasks, temporal generalization involves a longer period of time during which temporal information is stored in memory before being processed (McCormack, Brown, Maylor, Richardson, & Darby, 2002). Therefore, musicians’ superior performance appeared to be limited to aspects of timing that do not require long-term storage of temporal information.

Given musicians’ superior auditory temporal accuracy compared to nonmusicians, the major goal of the present study was to determine whether musical experience also exerts a beneficial effect on temporal information processing in the visual domain. The nature of temporal information processing represents a highly disputed issue concerning the underlying mechanisms of timing. Most researchers assume a general timing ability that is active in temporal information processing independent of the actual modality of the task and also independent of the task itself (cf., Grondin, 2003; Guttmann, Gilroy, & Blake, 2005; Merchant, Zarco, & Prado, 2008; Rammsayer & Ulrich, 2001). Nevertheless, several findings suggest a multidimensional model of temporal information processing, assuming several underlying mechanisms that may be active depending on the sensory modality and specific timing task used (cf., Block, 1990; Lapid, Ulrich, & Rammsayer, 2009; Penney, 2003). In the present study, therefore, also the dimensional structure of timing performance of musicians and nonmusicians was investigated with principal components analysis.

Our approach is based on a comparison of timing performance between formally trained musicians and participants without any specific musical experience, who were matched with regard to age, gender, and level of education. For this purpose, we employed a set of six basic timing tasks for psychophysical assessment of performance on different aspects of temporal information processing. As indicators of individual temporal acuity, performance measures on temporal fusion, rhythm perception, and interval timing in the range of seconds and milliseconds were obtained. In the following, a brief description of the six timing tasks will be given.

Temporal fusion refers to the size of the temporal interval between two sensory events that is required for them to be perceived as two separate events rather than fused as one event. Thus, temporal fusion thresholds represent a psychophysical indicator of temporal resolving power for central sensory information processing (Robin & Royer, 1987; van Wassenhove, 2009).

A focus of research on rhythm perception is on discrimination of serial temporal patterns (ten Hoopen et al., 1995). Commonly, in a rhythm perception task, a participant is presented with a simple pattern of brief auditory or visual stimuli. The participant’s task is to detect a deviation from regular, periodic interstimulus intervals.

For assessment of performance on interval timing, two temporal discrimination and two temporal generalization tasks were used. In a typical temporal discrimination task, a participant is presented with two intervals and his/her task is to decide which of the two intervals was longer. While timing of brief intervals in the range of milliseconds appears to be dependent on sensory processes beyond cognitive control, temporal processing of longer intervals is likely to be cognitively mediated (Lewis & Miall, 2003; Michon, 1985; Rammsayer, 1999; Rammsayer & Lima, 1991; Rammsayer & Ulrich, 2011). This latter mode of processing implies a cognitive representation of temporal information that draws on central executive resources and is subject to a limited-capacity attentional system (Rammsayer & Ulrich, 2011). Based on these considerations, two temporal discrimination tasks were employed: one task with a 50-ms standard duration and one task with a 1,000-ms standard duration.
In addition to the temporal discrimination tasks, two temporal generalization tasks were used with standard durations of 75 and 1,000 ms, respectively. Unlike temporal discrimination, temporal generalization requires a longer period of time during which temporal information is stored in memory. This is because with the latter task, participants are presented with a standard duration during a preexposure phase and are required to judge whether the durations presented during a subsequent test phase were the same as the standard duration that they have encountered earlier.

The short and long standard durations of the interval timing tasks were selected because the hypothetical shift from one timing mechanism to the other may be found at an interval duration somewhere between 100 and 500 ms (Abel, 1972; Buonomano & Karmarkar, 2002; Buonomano, Bramen, & Khodadadifar, 2009; Michon, 1985; Spencer, Karmarkar, & Ivry, 2009). Furthermore, when participants are asked to compare time intervals, many of them adopt a counting strategy. Since explicit counting becomes a useful timing strategy for intervals longer than approximately 1,200 ms (Grondin, Meilleur-Wells, & Lachance, 1999; Grondin, Ouellet, & Roussel, 2004), the “long” standard duration was chosen not to exceed this critical value.

Method

Participants
Two groups of participants, musicians and nonmusicians, participated in the study. The musician group included 20 male and 20 female musicians ranging in age from 18 to 30 years (mean age ± standard deviation: 22.7 ± 2.6 years). All participants of the musician group were graduate students at the University of Music, Drama and Media, Hannover, Germany, enrolled in the music master program with major “musical performance.” All musicians had played their instruments for 16.5 ± 0.5 years on average. The nonmusician group included 20 male and 20 female nonmusicians ranging in age from 18 to 28 years (mean age: 22.8 ± 2.6 years). All nonmusicians were graduate students at the University of Göttingen and reported that they had never been playing any musical instrument nor were they especially interested in music. Thus, none of the nonmusicians were occupied with music to a greater extent than occasionally listening to music. The level of education was matched between the two groups as both musicians and nonmusicians possessed the German Abitur, a high school degree required to enroll at German universities.

Psychometric Assessment of General Intelligence
As a psychometric index of general intelligence, the Zahlen-Verbindungs-Test (ZVT; Oswald & Roth, 1987) was used. The ZVT measures general intelligence by the assessment of information processing speed (Oswald & Roth, 1987). The ZVT is a trail-making test in which participants draw lines to connect, in order, circled numbers or letters that are positioned more or less randomly on a sheet of paper. Vernon (1993) introduced additional versions of the ZVT varying in task complexity. He found more complex versions to be more highly correlated with conventional measures of general intelligence. Therefore, we administered a ZVT version with high task complexity (Version 4) selected from Vernon’s (1993) survey. This version represents one of the most valid ones as it shows a substantial correlation of r = .71 with full-scale IQ obtained by Jackson’s (1983) Multidimensional Aptitude Battery. More recently, Rammsayer (2005) reported correlation coefficients of similar magnitude between ZVT Version 4 and a g factor of psychometric intelligence extracted from 15 subtests assessing different aspects of intelligence. With Version 4, participants are required to connect alternate numbers and letters backward (26-Z-25-Y-24-X… etc.). In the present study, ZVT Version 4 was administered in a speeded format that counted the number of items completed within 45 s. Test-retest reliability reported by Vernon (1993) for test administration under time limitations was r = .80.

Psychophysical Timing Tasks
Auditory and visual duration discrimination in the sub-second range. The presentation of the intervals to be judged and the recording of participants’ responses were controlled by a computer. The standard and the comparison stimuli were filled auditory or visual intervals. Auditory stimuli were white-noise bursts from a computer-controlled sound generator (Phylab Model 1), presented binaurally through headphones (Vivanco SR85) at an average intensity of 63 dB(A) SPL. Visual stimuli were generated by a red LED (diameter 0.48°, viewing distance 60 cm, luminance 48 cd/m²) positioned at eye level of the participant. The intensity of the LED was clearly above threshold, but not dazzling.

The duration discrimination task consisted of the presentation of one block of auditory and one block of visual intervals. The order of blocks was counterbalanced across participants. Each block consisted of 64 trials, and each trial consisted of one standard interval and one comparison interval. The duration of the comparison interval varied according to an adaptive rule (Kaernbach,
1991) to estimate x.25 and x.75 of the individual psychometric function; that is, the two comparison intervals at which the response “longer” was given with a probability of .25 and .75, respectively.

For both the auditory and the visual task, the standard interval was 50 ms and initial durations of the comparison interval were 15 ms below and above the standard interval for x.25 and x.75, respectively. To estimate x.25, the duration of the comparison interval was increased for Trials 1-6 by 3 ms if the participant had judged the standard interval to be longer and decreased by 9 ms after a “short” judgment. For Trials 7-32, the duration of the comparison interval was increased by 2 ms and decreased by 6 ms, respectively. The opposite step sizes were employed for x.75. In each experimental block, one series of 32 trials converging to x.75 and one series of 32 trials converging to x.25 were presented. Within each series, the order of presentation for the standard interval and the comparison interval was randomized and balanced, with each interval being presented first in 50% of the trials. Trials from both series were randomly interleaved within a block.

Each participant was seated at a table with a keyboard and a computer monitor. To initiate a trial, the participant pressed the space bar; auditory presentation began 900 ms later. The two intervals were presented with an interstimulus interval of 900 ms. The participant’s task was to decide which of the two intervals was longer and to indicate his decision by pressing one of two designated keys on a computer keyboard (two-alternative forced-choice technique). One key was labeled “First interval longer” and the other was labeled “Second interval longer.” The instructions to the participants emphasized accuracy; there was no requirement to respond quickly. After each response, visual feedback (“+”, i.e., correct; “−”, i.e., false) was displayed on the computer screen. The next trial started when the participant pressed the space bar again.

As a measure of performance, mean differences between standard and comparison intervals were computed for the last 20 trials of each series. Thus, estimates of the 25% and 75% difference thresholds in relation to the 50 ms standard intervals were obtained for the auditory and the visual task, respectively. In a second step, half the interquartile range, (75%−threshold value - 25%−threshold value)/2, representing the difference limen, DL (Luce & Galanter, 1963), was determined for both duration discrimination tasks. With this psychophysical measure, better performance on duration discrimination is indicated by smaller values.

**Auditory and visual duration discrimination in the second range.** Apparatus, stimuli, and the psychophysical procedure were the same as in the previous task except that the temporal intervals to be discriminated were longer. For duration discrimination of intervals in the second range, the standard interval was 1,000 ms and the initial values of the comparison interval were 500 ms and 1,500 ms for x.25 and x.75, respectively. To estimate x.25, the duration of the comparison interval was increased by 100 ms if the participant had judged the standard interval to be longer and decreased by 300 ms after a “short” response. For Trials 7-32, the duration of the comparison interval was increased by 25 ms and decreased by 75 ms, respectively. Again, the opposite step sizes were employed for x.75. As a psychophysical indicator of performance on auditory and visual duration discrimination, DLs were determined.

**Auditory and visual temporal generalization in the sub-second and second range.** In addition to the duration discrimination tasks, two auditory and visual temporal generalization tasks were used with base durations of 75 and 1,000 ms, respectively. Participants were presented with a standard duration during a pre-exposure phase and were required to judge whether the durations presented during the test phase were the same as the standard duration that they had encountered earlier.

Apparatus and stimuli were the same as in the previous experimental tasks. For auditory and visual temporal generalization of intervals in the sub-second range, the non-standard stimulus durations were 42, 53, 64, 86, 97, and 108 ms and the standard duration was 75 ms. For auditory and visual temporal generalization of intervals in the second range, the standard stimulus duration was 1,000 ms and the non-standard durations were 700, 800, 900, 1,100, 1,200, and 1,300 ms.

Performance on temporal generalization in the sub-second and second range was assessed separately for auditory and visual intervals. For both time ranges, order of the auditory and visual temporal generalization tasks was randomized and balanced across participants. With all generalization tasks, participants were required to identify the standard stimulus among the six non-standard stimuli. In the first part of the experiment, participants were instructed to memorize the standard stimulus duration. For this purpose, the standard interval was presented five times accompanied by the display “This is the standard duration.” Then participants were asked to start the test. Each generalization task consisted of eight blocks. Within each block, the standard duration was presented twice, while each of the six non-standard intervals was presented once. All duration stimuli were presented in randomized order.
On each test trial, one duration stimulus was presented. Participants were instructed to decide whether or not the presented stimulus was of the same duration as the standard stimulus stored in memory. Immediately after presentation of a stimulus, the display “Was this the standard duration?” appeared on the screen, requesting the participant to respond by pressing one of two designated response keys. Each response was followed by visual feedback. As a quantitative measure of performance on temporal generalization an individual index of response dispersion (cf., Wearden, Wearden, & Rabbitt, 1997) was determined. For this purpose, the proportion of total “yes” responses to the standard duration and the two non-standard durations immediately adjacent (e.g., 900, 1,000, and 1,100 ms) was determined. This measure would approach 1.0 if all “yes” responses were clustered closely around the standard duration.

**Auditory and visual rhythm perception.** Apparatus and stimuli were the same as in the previous experimental tasks. For the auditory rhythm perception task, the stimuli consisted of 3-ms clicks presented binaurally through headphones, while, for the visual task, light flashes with a duration of 3 ms were used. Participants were presented with rhythmic patterns, each consisting of a sequence of six clicks (auditory task) or six flashes (visual task) marking five beat-to-beat intervals. Four of these intervals were of a constant duration of 150 ms, while one interval was variable (150 ms + x). The magnitude of x changed from trial to trial depending on the participant’s previous response according to the weighted up-down procedure (Kaernbach, 1991) that converged on a probability of hits of .75. Correct responding resulted in a decrease of x and incorrect responses made the task easier by increasing the value of x. For each task, a total of 64 experimental trials were grouped in two independent series of 32 trials each. In Series 1, the third beat-to-beat interval was the deviant interval, while in Series 2 the fourth beat-to-beat interval was the deviant interval. Trials from both series were randomly interleaved.

The participant’s task was to decide whether the presented rhythmic pattern was perceived as “regular” (i.e., all beat-to-beat intervals appeared to be of the same duration) or “irregular” (i.e., one beat-to-beat interval was perceived as deviant). Participants indicated their decision by pressing one of two designated response keys. No feedback was given, as there were no perfectly isochronous (“regular”) patterns presented. As a psychophysical indicator of performance on auditory and visual rhythm perception, the 75% threshold for detection of irregularity was determined. Individual threshold estimates represented the mean threshold value across Series 1 and 2.

**Auditory and visual temporal fusion.** Apparatus and stimuli were the same as for the previous experimental tasks. The stimuli consisted of 25-ms noise bursts and 25-ms light flashes for the auditory and visual fusion tasks, respectively. Fusion threshold estimation consisted of 12 trials, and each trial consisted of two noise bursts (auditory fusion task) or two light flashes (visual fusion task) separated by a variable ISI ranging from 1 to 40 ms. After each trial, the participant’s task was to indicate by pressing one of two designated response keys whether he perceived the two successive noise bursts/flash lights as one tone/light or two separate events. The ISI was changed using an adaptive rule based on the Best PEST procedure (Pentland, 1980) to estimate the 75% auditory and visual fusion threshold.

**PROCEDURE**

The intelligence test and the experimental tasks were administered in a testing session of approximately 90 min. All experiments were carried out in a sound-attenuated room with constant ambient light. The testing session was initiated by the psychometric intelligence test followed by the duration discrimination tasks in the sub-second and second range, the temporal generalization tasks in the sub-second and second range, the rhythm perception tasks, and the two fusion tasks. For each type of timing task, order of auditory and visual versions of the task was counterbalanced across participants. Experimental trials of all tasks were preceded by practice trials to ensure that the participants understood the instructions and to familiarize them with the stimuli.

**Results**

Several studies suggest a positive association between individual level of general intelligence and performance on temporal information processing (e.g., Helmbold, Troche, & Rammsayer, 2007; Rammsayer & Brandler, 2007). Therefore, in a first step, musicians’ and nonmusicians’ ZVT scores as a psychometric indicator of general intelligence were compared by means of a t-test. There was no statistically significant difference in levels of general intelligence between the musician and nonmusician groups, t(78) = 0.41, p = .68; mean (± SEM) ZVT scores were 19.78 ± 1.02 and 19.22 ± 0.95 for musicians and nonmusicians, respectively. Thus, it appears highly unlikely that potential differences in timing performance between both groups will be due to differences in general intelligence.

In a second step, two-way analyses of variance were performed on all psychophysical timing tasks with
Group (two levels: musicians and nonmusicians) as an independent factor and Sensory Modality (two levels: auditory and visual) as a repeated measurement factor. Simple main effect means (± SEM) of musicians and nonmusicians for the auditory and visual versions of each temporal task are given in Table 1.

Two-way analyses of variance revealed significantly better timing performance for musicians compared to nonmusicians for duration discrimination in the sub-second, $F(1, 78) = 6.37, p < .05, \eta^2 = .08$, and in the second range, $F(1, 78) = 17.51, p < .001, \eta^2 = .18$, rhythm perception, $F(1, 78) = 17.74, p < .001, \eta^2 = .19$, and temporal fusion, $F(1, 78) = 7.54, p < .01, \eta^2 = .09$. No performance differences between both groups could be observed for temporal generalization in the sub-second, $F(1, 78) = 0.91, ns, \eta^2 = .01$, and second range, $F(1, 78) = 2.68, ns, \eta^2 = .03$.

A statistically significant main effect of Sensory Modality yielded superior auditory compared to visual timing acuity for duration discrimination in the sub-second, $F(1, 78) = 41.28, p < .001, \eta^2 = .48$, and in the second range, $F(1, 78) = 43.97, p < .001, \eta^2 = .36$, temporal generalization in the sub-second, $F(1, 78) = 10.44, p < .01, \eta^2 = .12$, rhythm perception, $F(1, 78) = 9.58, p < .01, \eta^2 = .19$, and temporal fusion, $F(1, 78) = 473.19, p < .001, \eta^2 = .86$. No modality-specific effect was found for temporal generalization in the second range, $F(1, 78) = 0.62, ns, \eta^2 = .01$.

A statistically significant interaction between Group and Sensory Modality could only be shown for duration discrimination of intervals in the range of milliseconds, $F(1, 78) = 4.98, p < .05, \eta^2 = .06$. Post-hoc Scheffé tests yielded reliably better temporal discrimination performance for musicians than for nonmusicians with visual intervals ($p < .001$) while no such difference was found for auditorily presented intervals. In addition, for both groups, timing performance was better with auditory than with visual stimuli ($p < .001$ each). There were no statistically significant interactions for all other temporal tasks applied in the present study.

In order to elucidate the dimensional structure of timing performance assessed in the present study, individual performance scores on the six different types of psychophysical timing tasks were subjected to a principal component analysis. The index of response dispersion obtained with the temporal generalization tasks was positively related to performance, i.e., better performance was indicated by higher values of response dispersion, while the other psychophysical measures based on threshold estimates were negatively associated with temporal performance, i.e., better performance was reflected by lower threshold values and lower DLs, respectively. Therefore, to enhance clarity of data presentation, the sign (+ or −) of the factor loadings presented in Table 2 has been adjusted such that positive values indicate a positive covariation of performance and respective factor loading.

Principal component analysis resulted in three factors with eigenvalues greater than unity; eigenvalues were 3.75, 1.40, and 1.28 for the first, second, and third component, respectively. The additional scree test (Cattell, 1966; Cattell & Vogelmann, 1977), however, unambiguously supported a one-factor solution. This first unrotated component accounted for 31.3% of total timing variability (see Table 2).

All temporal tasks, except the two temporal generalization tasks, consistently showed substantial factor loadings ranging from .55 to .77 irrespective of sensory modality. For the temporal generalization tasks, however, a rather ambiguous pattern of factor loadings emerged. For

---

**Table 1. Mean Performance Scores (± SEM) for Auditory and Visual Versions of Each Temporal Task for Musicians and Nonmusicians.**

<table>
<thead>
<tr>
<th>Temporal task</th>
<th>Indicator of performance</th>
<th>Auditory</th>
<th>Visual</th>
<th>Auditory</th>
<th>Visual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$M$</td>
<td>SEM</td>
<td>$M$</td>
<td>SEM</td>
</tr>
<tr>
<td>DD1</td>
<td>DL [ms]</td>
<td>6.35</td>
<td>0.30</td>
<td>23.73</td>
<td>1.34</td>
</tr>
<tr>
<td>DD2</td>
<td>DL [ms]</td>
<td>107.79</td>
<td>10.01</td>
<td>150.21</td>
<td>8.45</td>
</tr>
<tr>
<td>TG1</td>
<td>Response dispersion</td>
<td>.48</td>
<td>.02</td>
<td>.45</td>
<td>.02</td>
</tr>
<tr>
<td>TG2</td>
<td>Response dispersion</td>
<td>.50</td>
<td>.02</td>
<td>.47</td>
<td>.02</td>
</tr>
<tr>
<td>RP</td>
<td>75% threshold [ms]</td>
<td>42.32</td>
<td>2.00</td>
<td>44.78</td>
<td>2.77</td>
</tr>
<tr>
<td>Fusion</td>
<td>75% threshold [ms]</td>
<td>5.90</td>
<td>0.47</td>
<td>26.08</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.34</td>
<td>0.47</td>
<td>30.00</td>
<td>1.56</td>
</tr>
<tr>
<td></td>
<td></td>
<td>140.51</td>
<td>9.30</td>
<td>211.97</td>
<td>12.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.48</td>
<td>.02</td>
<td>.48</td>
<td>.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.45</td>
<td>.02</td>
<td>.45</td>
<td>.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.02</td>
<td>.02</td>
<td>.02</td>
<td>.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>54.50</td>
<td>2.97</td>
<td>63.66</td>
<td>3.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9.68</td>
<td>1.59</td>
<td>29.08</td>
<td>0.99</td>
</tr>
</tbody>
</table>

*Note: DD1: duration discrimination of intervals in the sub-second range, standard = 50 ms; DD2: duration discrimination of intervals in the second range, standard = 1,000 ms; TG1: temporal generalization, standard = 75 ms; TG2: temporal generalization, standard = 1,000 ms; RP: rhythm perception; Fusion: temporal fusion.*
temporal generalization in the sub-second range, factor loadings were .20 and .41 for auditory and visual intervals, respectively, while for temporal generalization in the second range, factor loadings were .49 and .27 for auditory and visual intervals, respectively. Since all timing tasks other than auditory temporal generalization in the sub-second range and visual temporal generalization in the second range markedly contributed to this first component, it appears reasonable to construe this factor as task- and modality-independent general timing ability.

In a final step, resulting factor scores of musicians and nonmusicians were compared by means of t tests. Mean factor score was reliably higher in musicians than in nonmusicians, t(78) = 5.24, p < .001, d = 1.19, indicating that the musician group scores notably higher in general timing ability than the nonmusician group.

Discussion

The first goal of the present study was to replicate and expand the finding that temporal information processing in the auditory modality is more accurate in musicians than in nonmusicians (Rammsayer & Altenmüller, 2006). The second aim of the study was to investigate whether the superior temporal information processing of musicians compared to nonmusicians also holds for the visual modality. For this purpose, timing performance on a set of six different psycho-physical temporal tasks for both the auditory and visual sensory modality was compared in 40 formally trained musicians and 40 controls without musical experience.

With respect to group differences, superior temporal acuity for musicians compared to nonmusicians was found for all temporal tasks except for the two temporal generalization tasks. When comparing the two sensory modalities, temporal acuity was superior with auditory stimuli as compared to visual stimuli with the exception of the temporal generalization task in the second range. With this latter task, timing performance was almost the same for both sensory modalities.

The musicians’ superior performance compared to the nonmusicians confirms the results reported by Rammsayer and Altenmüller (2006). The fact that performance on various aspects of temporal perception such as rhythm perception, temporal fusion, and duration discrimination was consistently superior in the musician group may lead to the assumption that one general internal timing mechanism underlies all of these different aspects of timing performance. To further elucidate this assumption, a principal component analysis was performed to identify the dimensional structure of timing performance. All timing tasks, except for auditory temporal generalization in the sub-second range and visual temporal generalization in the second range, exhibited substantial loadings on the first unrotated principal component. This finding also supports the notion of a task- and modality-independent general internal timing mechanism.

The notion of a hypothetical general timing mechanism that operates independent of sensory modality is in line with both some experimental findings and a major theoretical account of temporal information processing. Because perceptual timing tasks require processing of events or changes in information over time, several authors (e.g., Burle & Bonnet, 1997, 1999; Rammsayer & Brandler, 2007; Surwillo, 1968) have put forward the idea that a general internal timing mechanism in the brain is responsible for various aspects of temporal information processing such as rhythm perception or interval timing. More specifically, performance on interval timing is often explained by the general assumption of a hypothetical internal clock based on neural counting (e.g., Creelman, 1962; Gibbon, 1977; Killeen & Weiss, 1987; Rammsayer & Ulrich, 2001; Treisman, Faulkner, Naish, & Brogan, 1990). The main features of such an internal clock mechanism are a pacemaker and an accumulator. The pacemaker emits pulses and the number of pulses relating to a physical time interval is recorded by the accumulator. Thus, the number of pulses counted during a given time interval is the internal temporal representation of the interval. The higher the clock rate, the finer the temporal resolution of the internal clock will be, which is equivalent to higher

<table>
<thead>
<tr>
<th>Temporal task</th>
<th>Factor 1</th>
<th>Eigenvalue</th>
<th>Explained variance [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>DD1_a</td>
<td>.65</td>
<td>.55</td>
<td>31.29</td>
</tr>
<tr>
<td>DD1_v</td>
<td>.59</td>
<td>.57</td>
<td></td>
</tr>
<tr>
<td>DD2_a</td>
<td>.55</td>
<td>.59</td>
<td></td>
</tr>
<tr>
<td>DD2_v</td>
<td>.69</td>
<td>.55</td>
<td></td>
</tr>
<tr>
<td>TG1_a</td>
<td>.41</td>
<td>.57</td>
<td></td>
</tr>
<tr>
<td>TG1_v</td>
<td>.41</td>
<td>.55</td>
<td></td>
</tr>
<tr>
<td>TG2_a</td>
<td>.49</td>
<td>.55</td>
<td></td>
</tr>
<tr>
<td>TG2_v</td>
<td>.27</td>
<td>.59</td>
<td></td>
</tr>
<tr>
<td>RP_a</td>
<td>.68</td>
<td>.55</td>
<td></td>
</tr>
<tr>
<td>RP_v</td>
<td>.77</td>
<td>.55</td>
<td></td>
</tr>
<tr>
<td>Fusion_a</td>
<td>.57</td>
<td>.59</td>
<td></td>
</tr>
<tr>
<td>Fusion_v</td>
<td>.58</td>
<td>.55</td>
<td></td>
</tr>
<tr>
<td>Eigenvalue</td>
<td>3.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Explained variance [%]</td>
<td>31.29</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: a: auditory; v: visual.
temporal sensitivity, as indicated by better performance on duration discrimination and rhythm perception (Pashler, 2001; Rammsayer & Brandler, 2007).

Within the framework of a general, modality-independent timing mechanism, better timing performance with auditory than with visual stimuli can be envisioned as an increase in neural firing rate in the case of auditory temporal stimuli (cf. Grondin, 2001; Wearden, Edwards, Fakhri, & Percival, 1998). This higher pacemaker rate yields finer temporal resolution and, thus, better timing accuracy for auditory compared to visual stimuli. In a recent fMRI study, Shih, Kuo, Yeh, Tzen, and Hsieh (2009) identified the supplementary motor area and the basal ganglia as a common neural substrate involved in temporal processing of both auditory and visual intervals in the subsecond range. Both brain structures and their precise interactions in the millisecond range have been shown to be extremely sensitive to timing demands in musicians and nonmusicians (cf., Chen, Penhune, & Zatorre, 2008; Haslinger et al., 2005; Krause, Schnitzler, & Pollok, 2010).

In a recent psychophysical study, Stauffer, Haldemann, Troche, and Rammsayer (2011) confirmed higher temporal sensitivity for rhythm perception and duration discrimination in the range of milliseconds for the auditory compared to the visual sensory modality. Furthermore, their data also provide empirical evidence for a hierarchical model with modality-specific visual and auditory temporal processing at a first level and a superordinate, modality-independent processing system at a second level of the hierarchy.

In the present study, musicians also performed superior to nonmusicians in the set of temporal processing tasks presented in the visual sensory modality. If a timing mechanism, such as the hypothetical internal clock, underlies temporal information processing independent of sensory modality, the influence of extensive music training may indirectly shape also the timing ability in the visual sensory modality.

An alternative possible explanation of the musicians’ superior performance in the visual sensory modality is that the processing of visual temporal clues more directly benefited from the years of music training. Musicians rely very strongly on visual clues when playing with a conductor in an orchestra as well as when playing in small ensembles. Synchronizing to and anticipating the movements of the conductor and of other members of the ensemble are crucial to establish optimal timing while making music together (cf., Pecenka & Keller 2011; Repp, 2006). This latter explanation does not necessitate a general, modality-independent timing mechanism but would also be consistent with the notion of two distinct, modality-specific timing mechanisms.

Supposedly converging evidence for this latter notion can be derived from the significant interaction between Group and Sensory Modality for the duration discrimination task in the sub-second range. Post-hoc tests revealed a superior temporal discrimination performance for musicians compared to nonmusicians for visually presented but not for auditorily presented intervals. For both groups, timing performance was better with auditory than with visual stimuli. Another study comparing timing performance in blind and sighted participants also found performance differences in most of the applied auditory temporal tasks except for discrimination of brief tones in the sub-second range (Rammsayer, 1994). This lack of a performance difference has been attributed to the fact that temporal discrimination of brief auditory intervals can be considered a process highly overlearned in everyday life. The ability to discriminate the duration of sounds is very critical for speech perception (Ackermann, Graber, Hertrich, & Daum, 1999; Drullman, 1995; Liberman, Delattre, Gerstman, & Cooper, 1956; Scott, 1982; Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995; Tallal et al., 1996). Therefore, temporal discrimination of extremely brief tones in the range of milliseconds may represent an overlearned perceptual function that does not benefit from additional music training. Such an explanation may account for the observed lack of a performance difference between blind and sighted individuals (Rammsayer, 1994), as well as between musicians and nonmusicians in the present study.

Eventually, from a theoretical point of view, crossmodal perceptual learning from the auditory to the visual sensory modality might also account for musicians’ superior timing performance. It should be noted, however, that in the present study, reliably superior timing accuracy was revealed for musicians compared to nonmusicians for visual, but not for auditory duration discrimination in the range of milliseconds. Thus, apparently no crossmodal transfer from the auditory to the visual sensory modality occurred in nonmusicians. This conclusion is consistent with the outcome of a study employing a perceptual learning paradigm to examine potential crossmodal transfer in a duration discrimination task. In this study, employing nine testing sessions extending over two weeks, Lapid et al. (2009) showed transfer within the auditory modality but failed to confirm a transfer from the auditory to the visual modality. One should keep in mind, though, that unlike nonmusicians, all musicians tested in the present study
had long lasting, intensive music training starting in childhood. This extensive practice may represent a crucial prerequisite for a crossmodal transfer from the auditory to the visual domain to become effective. From this perspective, Lapid et al.’s (2009) failure to show crossmodal transfer in timing performance could be attributed to insufficient practice rather than challenging the idea of crossmodal perceptual learning in temporal information processing.

In the present study, musicians performed superior to nonmusicians on all temporal tasks except for the temporal generalization task. This result replicates former findings for the auditory domain (Ramsayer & Altenmüller, 2006) and may be indicative of, at least partly, qualitatively different timing mechanisms involved in temporal generalization, on the one hand, and duration discrimination, rhythm perception, and temporal fusion, on the other. With the latter class of timing tasks, temporal information has to be stored in memory for a much shorter period of time than with the temporal generalization paradigm. Based on this consideration, our findings suggest that extensive music training may exert a positive effect on timing performance by reducing variability or noise associated with the timing process itself. However, this advantage of musicians compared to nonmusicians appears to wear off with increasing memory retention time. Thus, temporal judgments which cannot be derived directly from perceptual processing seem to be less sensitive to music training.

While it is plausible that the differences between musicians and nonmusicians are the result of music training, a cautionary note, however, is that innate differences in timing ability between these groups cannot be ruled out completely. Another possible moderating variable that could account for the difference between the two groups is that the musicians were more motivated when performing the timing tasks (cf., McAuley, Henry, & Tuft, 2011). Also, given that various cognitive measures predict timing ability (e.g., Helmbold et al., 2007; Ramsayer & Brandler, 2007; Troche & Ramsayer, 2009), an additional alternative explanation could be that there is a selection bias in a sense that individuals with higher cognitive ability go on to study music. In the present study, however, all musicians and nonmusicians were graduate students and the two groups did not differ in mean ZVT score, a global measure of general intelligence.

Taken together, the present study confirmed previous findings of superior auditory timing performance in musicians compared to nonmusicians. Furthermore, we were able to expand these findings by providing first direct experimental evidence that musicians’ superior temporal information processing also holds for the visual modality. The overall pattern of our results is consistent with the notion that musicians’ long-lasting extensive music training, starting in childhood, can enhance general timing ability irrespective of sensory modality.

Author Note

We thank Kati Klopfeisch for her assistance in data collection.

Correspondence concerning this article should be addressed to Thomas Ramsayer, Department of Psychology, University of Bern, Muesmattstrasse 45, CH-3000 Bern 9, Switzerland. E-mail: thomas.ramsayer@psy.unibe.ch

References


