


SCIENTIFIC REPORTS



OPEN

Unimodal and cross-modal prediction is enhanced in musicians

Eliana Vassena^{1,*}, Katty Kochman^{2,*}, Julie Latomme¹ & Tom Verguts¹

Received: 14 December 2015

Accepted: 06 April 2016

Published: 04 May 2016

Musical training involves exposure to complex auditory and visual stimuli, memorization of elaborate sequences, and extensive motor rehearsal. It has been hypothesized that such multifaceted training may be associated with differences in basic cognitive functions, such as prediction, potentially translating to a facilitation in expert musicians. Moreover, such differences might generalize to non-auditory stimuli. This study was designed to test both hypotheses. We implemented a cross-modal attentional cueing task with auditory and visual stimuli, where a target was preceded by compatible or incompatible cues in mainly compatible (80% compatible, predictable) or random blocks (50% compatible, unpredictable). This allowed for the testing of prediction skills in musicians and controls. Musicians showed increased sensitivity to the statistical structure of the block, expressed as advantage for compatible trials (disadvantage for incompatible trials), but only in the mainly compatible (predictable) blocks. Controls did not show this pattern. The effect held within modalities (auditory, visual), across modalities, and when controlling for short-term memory capacity. These results reveal a striking enhancement in cross-modal prediction in musicians in a very basic cognitive task.

Music is a universal attribute to all human cultures and pervasive in daily life^{1–3}. Advanced musical practice involves skills in processing various kinds of stimuli. First, musicians are highly trained in the memorization of auditory stimuli, from simple tones to complex rhythms and harmonic structures⁴. Second, musicians read symbolic visual stimuli associated with those tones. Third, musicians produce tones by performing automatized, refined actions involving haptic feedback in coordination with their environment^{5,6}. These components are combined in a rapidly evolving processing stream, and yet organized in a meaningful sequence, which produces the pleasant stimulus the listener perceives as music^{7,8}.

The complexity of musical training suggests that consistent exposure and expertise may be associated with measurable effects in several cognitive functions⁹, as well as on brain plasticity¹⁰. Musicians show enhanced auditory-perception skills, such as pitch discrimination, temporal order judgment¹¹ and discrimination of psychoacoustic features¹². Moreover, structural and functional changes in brain regions dedicated to auditory processing have been consistently reported¹³. These auditory-perceptual advantages suggest that musical expertise is associated with improved temporal discrimination and attentional capacity¹⁴.

The generalization of these benefits to other cognitive functions remains debated¹⁵. A popular study reported that exposure to a 10-minute fragment by Mozart improved spatial reasoning¹⁶. This result had great media resonance and was named “Mozart effect” by the press, conveying the idea that classical music could in fact improve cognitive skills. Although not consistently replicated¹⁷, this result stimulated further research testing whether musical training provides benefits beyond auditory perception, yielding controversial results.

On the one hand, several studies have reported advantages for musicians in diverse cognitive domains. Musicians have shown better reproduction of time intervals for both auditory and visual intervals¹⁸, as well as better reproduction of multimodal sequences¹². Musicians have also performed better in judging whether auditory and visual information was presented synchronously or asynchronously in musical video clips¹⁹. Musicians also outperform controls in attention and visuo-spatial tasks such as detecting single elements in complex objects, detecting letters among digits²⁰, and line bisection²¹. Finally, children exposed to 9-months of musical training have shown increased reading abilities as well as pitch discrimination in speech²². These findings suggest a cross-modal transfer of benefits for musicians beyond the musical domain and beyond the auditory modality.

On the other hand, several studies have found selective benefits to musical and auditory processing, with no generalization to other modalities. For example, advantages in attentional performance were reported only for auditory, but not for visual attention tasks^{12,23}. Also, musicians proved selectively better at reproducing auditory

¹Department of Experimental Psychology, Ghent University, Belgium. ²Institute for Psychoacoustics and Electronic Music, Ghent University, Belgium. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to E.V. (email: eliana.vassena@ugent.be)

sequences but not audio-visual sequences²⁴. Additionally, no advantage was reported in learning sequence structure after passive listening. Lastly, musicians showed improved detection of audio-visual asynchrony but only with music and not with speech²⁵. However, an important caveat is that most of these studies present correlational evidence, showing better performance in musicians compared to controls. Although informative, one cannot infer a causal influence of musical training on cognitive skills (be it selective to the auditory and musical domain or more general). Causal evidence remains sparse in the literature and should receive further attention in future research.

This overview suggests auditory processing benefits, possibly deriving from the extensive expertise that musicians acquire in fast-scale temporal processing. However, the evidence for advantages beyond the auditory domain remains varied.

A potential benefit on the core cognitive process of prediction has been hypothesized, but not directly tested. Prominent cognitive theories such as predictive coding and reinforcement learning suggest that cognitive processing proceeds by prediction^{26–30}. In this framework, each stimulus or sequence leads to a prediction of the upcoming stimulus. Predicted and actual stimulus are compared, and this discrepancy is termed prediction error. Prediction errors drive both cognitive processing and learning³¹. Perception of musical rhythm and meter have been framed in the context of predictive theories³², as well as the relationship between perception and action in musical performance³³. Error (and prediction error) minimization is the core concept shared by these accounts. Furthermore, it has been proposed that the surprise associated with prediction error carries an affective component, and an ideal amount of surprise (not too much but not too little) drives affective reaction to music, as well as guiding expert musicians in pleasing their audience³⁴. From the empirical point of view, studies reported an advantage for musicians in detecting auditory prediction errors, both with simple sounds and complex harmonic structures^{35,36}. The neural signature of deviance detection also reflects this facilitation. Rhythmic deviance also elicits error-related neural activity³⁷. Moreover, expert musicians have shown neural correlates of error detection even before performing and incorrect action, supposedly arising from a continuous fast monitoring of predictions and outcomes³⁸.

Taken together, these findings support a pivotal role of prediction in music perception and performance. An further intriguing possibility is that musical expertise might be associated with improved prediction skills. The goal of the current study was to test this hypothesis with a standard cognitive task involving basic stimulus-outcome prediction skills, outside the musical domain. Given that prediction applies to any stimulus sequence irrespective of its modalities, we hypothesized that this facilitation may extend to non-auditory (e.g. visual) and even cross-modal sequences (i.e., when an auditory stimulus is predictive of a visual one and vice-versa). The advantage should manifest as increased sensitivity to prediction errors, as a consequence of increased encoding of the statistical structure of the environment in both unimodal and cross-modal conditions. To test this, we implemented a cross-modal cueing paradigm with auditory and visual stimuli. We implemented different levels of predictability and thus prediction errors by different frequencies of compatible and incompatible cue-target pairs. Moreover, we administered two control tasks to measure verbal and visuo-spatial short-term memory. The goal was to determine whether the hypothesized difference would be specific to prediction, or simply accountable to differences in short-term memory capacity.

Results

In the cross-modal cueing task, overall accuracy was $78\% \pm 0.4$. Accuracy rates were averaged for each subject and for each condition, and subjected to a rANOVA. No significant effect of group was found ($F_{(1,27)} = 0.524, p = 0.48$), showing that accuracy did not differ between musicians and controls. A significant main effect of compatibility was observed ($F_{(1,28)} = 4.57, p < 0.05, \eta^2 = 0.14$), with higher accuracy for compatible trials ($M = 0.96 \pm 0.02$) relative to incompatible trials ($M = 0.95 \pm 0.02$).

RTs were averaged for each subject and for each condition. Error trials (4.7%) were excluded from further analysis. To minimize the impact of outliers, trials with RTs higher or lower than 2.5 standard deviations of the individual mean were also excluded (2.9%). Trimming means by removing outlying observations is a common way of making the mean a more robust measure of central tendency³⁹, and a 2.5 standard deviations cut-off is a commonly used convention in the field⁴⁰. Subsequently, we tested the assumption of normality of the residuals with the Shapiro-Wilk test. All p-values were larger than 0.05, confirming that the residuals were normally distributed. For 2 out of the 16 conditions the p-values were still rather small (0.08 and 0.06). To test the robustness of our results, we log-transformed the data and run the main analysis again. All significant main effects and interactions reported in the main analysis were preserved when tested on the log-transformed data.

Crucially, the rANOVA revealed a significant interaction group \times compatibility frequency \times compatibility ($F_{(1,28)} = 6.24, p < 0.05, \eta^2 = 0.18$, see Fig. 1), with musicians showing a stronger influence of compatibility frequency (enhanced compatibility effect in the 80/20 condition) as compared to controls.

Pairwise comparisons revealed a significant difference for musicians between compatible and incompatible trials in the 80/20 condition ($t_{(14)} = -20.17, p = 0.001$), but not in the 50/50 condition ($t_{(14)} = -4.9, p = 0.21$). Thus, musicians were relatively disadvantaged for incompatible targets, but only in the 80/20 condition. Conversely, controls showed no difference between compatible and incompatible trials in the 80/20 condition ($t_{(14)} = -1.68, p = 0.12$) but did show a small difference between compatible and incompatible trials in the 50/50 conditions ($t_{(14)} = -2.61, p = 0.02$). However, only the difference for the musicians between compatible and incompatible trials in the 80/20 condition remained significant after Bonferroni correction for multiple comparisons. This interaction shows increased sensitivity in musicians to compatibility frequency, suggesting a better representation of the statistical structure of the block. This translated in an increased compatibility effect, when incompatible trials were less frequent. Crucially, the cue modality ($F_{(1,28)} = 0.01, p = 0.93$) or target modality ($F_{(1,28)} = 0.73, p = 0.4$) did not interact with the 3-way group \times compatibility frequency \times compatibility interaction, indicating that the effect holds across cue and target modalities.

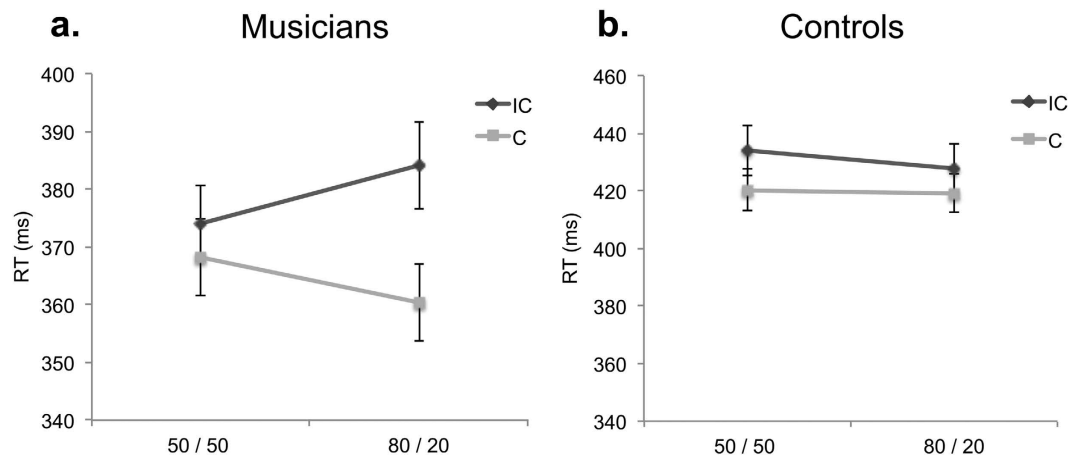


Figure 1. Group \times compatibility frequency \times compatibility interaction. (a) Average RTs for musicians in compatible (C) and incompatible (IC) trials, as a function of compatibility frequency in the block (50/50, 80/20). The error bars represent one standard error of the mean. (b) Average RTs for controls in compatible (C) and incompatible (IC) trials, as a function of compatibility frequency in the block (50/50, 80/20). The error bars represent one standard error of the mean.

Furthermore, a main effect of group was observed ($F_{(1,28)} = 10.49, p < 0.01, \eta^2 = 0.27$), with musicians showing overall faster RTs than controls. Pairwise comparisons across compatibility frequency and compatibility frequency conditions revealed that musicians reacted faster than controls in all conditions: to compatible targets and incompatible targets in random blocks (C $t_{(28)} = -3.39, p = 0.002, IC t_{(28)} = -3.46, p = 0.002$), and to compatible and incompatible targets in mainly compatible blocks (C $t_{(28)} = -3.6, p = 0.001, IC t_{(28)} = -2.37, p = 0.025$). This last comparison however, was not significant after applying a Bonferroni correction for multiple comparisons.

Additionally, there was a main effect of target modality ($F_{(1,28)} = 189.51, p < 0.001, \eta^2 = 0.87$), with faster RTs to visual targets, and a main effect of compatibility ($F_{(1,28)} = 16.61, p < 0.001, \eta^2 = 0.37$), with faster RTs in compatible trial. A significant group \times target modality interaction was also observed ($F_{(1,28)} = 5.6, p < 0.05, \eta^2 = 0.17$): Musicians responded faster to visual targets compared to auditory targets ($t_{(14)} = -7.81, p < 0.001$, mean difference -87.6 ms); controls also responded faster to visual targets ($t_{(14)} = -11.8, p < 0.001$, mean difference -123.9 ms); however, the difference for controls was larger, thus driving the interaction. This interaction might reflect a facilitation for musicians in responding to auditory stimuli. Furthermore, there was a significant cue modality \times target modality interaction ($F_{(1,28)} = 32.93, p < 0.001, \eta^2 = 0.54$), with faster RTs to visual as compared to auditory targets ($t_{(29)} = -12.79, p < 0.001$), but no significant difference between visual and auditory cues ($t_{(29)} = 0.21, p < 0.84$). A target modality \times compatibility frequency interaction was also reported ($F_{(1,28)} = 5.29, p < 0.05, \eta^2 = 0.16$): RTs were faster for visual compared to auditory targets in both the 80/20 condition ($t_{(29)} = -12.38, p < 0.001$, mean difference -101.37 ms) and 50/50 condition ($t_{(29)} = -12.56, p < 0.001$, mean difference -110.18 ms), with a larger difference for the latter.

Subsequently, performance on the short-term memory tasks was analyzed. The overall verbal short-term memory capacity score was 74.03 ± 6.01 . No significant differences were observed between musicians ($M = 76.2 \pm 7.54$) and controls ($M = 71.87 \pm 9.6, t_{(28)} = 0.34, p = 0.73$). The overall visuo-spatial short-term memory capacity score was 74.4 ± 4.36 . No significant differences were reported between musicians ($M = 79.8 \pm 7.01$) and controls ($M = 74.4 \pm 4.36, t_{(28)} = 1.28, p = 0.22$).

Although there was no group effect in short-term memory, in order to further ensure that the RT effects in the cross-modal cueing task could not be explained by differences in short-term memory capacity, the main rANOVA on RTs was repeated, including both short-term memory scores as covariates. No significant interaction of any factor with short-term memory scores was observed. Moreover, the most relevant group \times compatibility frequency \times compatibility was preserved ($F_{(1,28)} = 7.07, p < 0.05, \eta^2 = 0.21$), showing that the core finding (Fig. 1) is not driven by differences in short-term memory capacity.

Discussion

This study investigated the basic cognitive skill of prediction in musicians and non-musicians. We hypothesized an advantage for musicians in encoding predictable event sequences. The results can be summarized as follows. First, musicians showed enhanced prediction relative to controls, expressed as increased sensitivity to statistical block structure (compatibility frequency) in a very basic cueing task. Second, modulation by musical expertise held across modalities, revealing a striking cross-modal generalization. This shows increased prediction skills in musicians as compared to controls irrespective of event modality. Third, enhanced prediction could not be explained by short-term memory differences.

Earlier work addressed the role of prediction in music and auditory processing. One conclusion was that regular sound sequences generate predictions and prediction errors at several hierarchical levels, which can determine the pleasurable of the sequence⁴¹. Showing better prediction skills in musicians in a very basic cueing task indicates that the prediction machinery used in musical processing is rooted in basic cognitive prediction mechanisms.

Additionally, we reported an overall advantage for musicians, who responded faster in all conditions and irrespective of modality (although no differences in accuracy at any of the tasks were found). Moreover, the group \times target modality interaction, suggested to some extent faster processing of auditory targets. On the one hand, this result is compatible with prominent accounts stating that musical training results in fine-tuning and increased efficiency and precision of the auditory system¹³, also generalizing to speech^{42,43}. On the other hand, we report an advantage in prediction for musicians irrespective of cue and target modality, thus advocating for a non-selective effect.

In conclusion, this study provides several avenues for future work. First and foremost, we did not investigate causality. Our data is correlational in nature and does not allow for the drawing of inferences on the effect of musical training on cognition. In fact, a plausible alternative interpretation could be that people with better prediction skills become interested in music as a consequence of their natural abilities, and are presumably more suited to pursue music professionally in life. In order to disentangle the origin of such differences between musicians and non-musicians future studies should administer musical training to naive subjects and measure prediction before and after training. In addition to testing for causal relationship, this would allow investigating the amount of training required to induce measurable benefits. This might be particularly relevant in the context of longitudinal studies addressing the benefits of musical education.

Second, we did not distinguish between types of musical expertise or years of training, variables that were relevant in earlier research⁴⁴. The extent of the advantage might depend on musical instrument, as well as vary as a function of years of training.

Third, our sample size was rather limited. Future studies should aim for a larger amount of participants in each group to increase power and generalizability of the results.

Finally, the role of prediction in cognitive processing has been widely studied in computational neuroscience (e.g., reinforcement learning⁴⁵, predictive coding^{26,31,46}). Here, prediction is central in perception, learning, memory, decision-making, and action selection. It is an exciting open question to what extent training prediction skills as implemented in musical practice, may improve such domain-general abilities.

Methods

Participants. Thirty subjects participated to the study (age range 17–33, $M = 19.5$, $SD = 3.2$), recruited among Ghent University students who earned credits for participation. The sample size was determined a priori, following earlier conventions in music research^{11,18,21}. All participants provided written informed consent. Fifteen were selected for their musical expertise (nine males) with the following requirements: a minimum of five years of playing a musical instrument; having followed formal musical training in a music school; practicing the instrument on a daily basis; and currently playing the instrument at the time of the study. Fifteen control participants were selected (five males), where the listed requirements were exclusion criteria. Control participants and musicians did not differ in age ($t_{(28)} = 0.11$, $p = 0.91$) group. All participants gave written informed consent before participation. The experiment was conducted under the General Ethical Protocol for scientific research at the Department of Psychology and Educational Sciences of Ghent University, approved by the department's ethical committee. The procedure was in accordance with the guidelines provided within such protocol.

Procedure. After completing the informed consent, participants performed the main task (cross-modal cueing task), followed by two control tasks measuring verbal short-term memory (verbal span task) and visuo-spatial short-term memory (Corsi block tapping task). All tasks were programmed in E-prime 2.0 (Psychology Software Tools, Pittsburgh, PA) and presented on a 15" computer screen. Headphones were used to present auditory stimuli.

Cross-modal cueing task. Each trial started with a fixation cross (see Fig. 2). A first stimulus was presented as a cue (650 ms). The cue was followed by a target (650 ms), to which the participants had to respond as quickly and as accurately as possible, with a maximum response time limit of 1650 ms. In all trials, pressing the response key terminated the trial. Cue and target could be compatible or incompatible. In auditory-auditory trials, cue and target were auditory stimuli. Each of them could be a low pitch tone (800 Hz) or high pitch tone (1600 Hz, 650 ms). At the target, participants responded by indicating if the tone was low (left key press) or high (right key press). In compatible trials, the target matched the pitch of the cue, while in incompatible trials it did not. In visual-visual trials, two squares were presented, one to the left and one to the right of the fixation cross. The cue consisted of an arrow pointing left or right appearing in the center of the screen. The target stimulus consisted of an X, appearing either in the left square or in the right square. Participants had to indicate if the target stimulus X was on the left (left key press) or on the right (right key press). In compatible trials, the X appeared in the same direction as the arrow pointed, while in incompatible trials it appeared on the opposite side. In auditory-visual trials, the auditory cue was followed by the visual target. Trials were considered compatible when a low tone was followed by an X on the left, and a high tone by an X on the right. Trials were considered incompatible when a low tone was followed by an X on the right and when a high tone was followed by an X on the left. In visual-auditory trials the visual cue was followed by the auditory target. Trials were considered compatible when an arrow pointing left was followed by a low tone and an arrow pointing right by a high tone. Trials were considered incompatible when a left-pointing arrow was followed by a high tone and a right-pointing arrow by a low tone.

Cue modality and target modality were manipulated across blocks, yielding AA, AV, VA and VV blocks. Each of these block types occurred with 80% compatible and 20% incompatible trials (80/20 condition, 40 trials per block, 8 of which incompatible), or with 50% compatible and 50% incompatible trials (50/50 condition, 28 trials per block). This represents the crucial manipulation, as the 80/20 condition is characterized by a statistical structure that prompts prediction. This resulted in 8 blocks, repeated 4 times each, for a total of 32 blocks and 1088 trials. The order of presentation of these blocks was randomized, as well as the order of presentation of trials within a block.

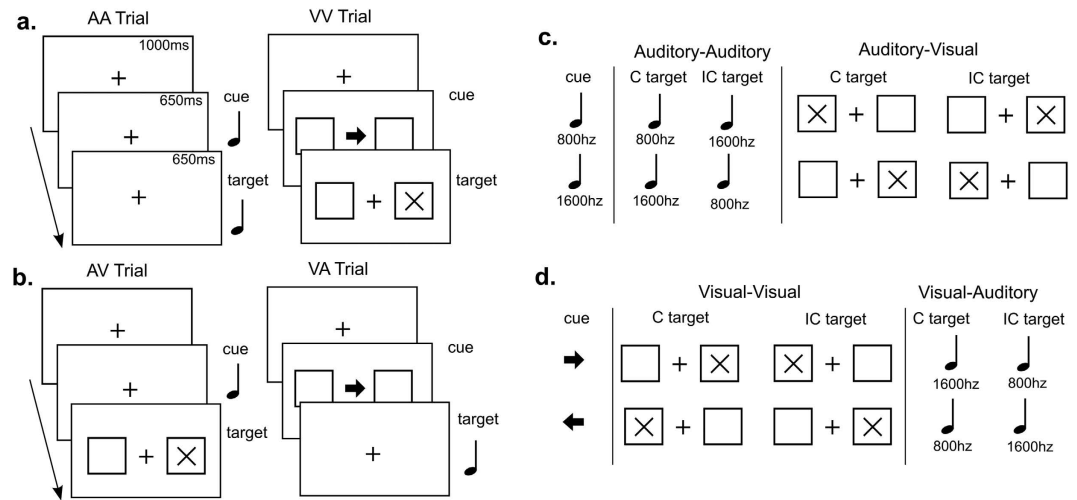


Figure 2. Task structure. (a) Task timing with an example of the two unimodal trial types: auditory-auditory (AA) with auditory cue and auditory target (tones); visual-visual (VV) with a visual cue (arrow) and visual target (X). (b) Example of the two cross-modal trial types: auditory-visual (AV) with auditory cue (tone) and visual target (X); visual-auditory (VA) with visual cue (arrow) and auditory target (tone). (c) Cue-target combinations and compatibility for AA and AV trials. From left to right: Auditory cues (low tone 800 Hz, high tone 1600 Hz); compatible (C) and incompatible (IC) auditory targets (AA trial); compatible (C) and incompatible (C) visual targets (AV trial). (d) Cue-target combinations and compatibility for VV and VA trials. From left to right: Visual cues (left or right pointing arrow); compatible (C) and incompatible (C) visual targets (VV trial); compatible (C) and incompatible (C) auditory targets (VA trial).

To summarize, the design implemented the following factors: cue modality (auditory or visual), target modality (auditory or visual), compatibility frequency in the block (80% or 50%), and compatibility in the trial (compatible or incompatible). The task lasted about 45 minutes.

Verbal short-term memory task. To control for verbal short-term memory capacity, a letter span task was administered⁴⁷. At the start of every trial, a blank screen was presented (1500 ms). A sequence of 4 consonants followed, each remaining on the screen for 1200 ms (inter-letter blank of 250 ms). After a 1500 ms retention period, participants were instructed to reproduce the sequence by pressing the corresponding keys on the keyboard in the correct order. The length of the string to be retained increased by one every time that the participant correctly reported three sequences of the current length. The maximum string length was 9. Failing to reproduce a sequence for three consecutive trials would terminate the task, constituting the span length for that participant. A first practice trial with feedback was presented in the beginning. The total task duration was 5 minutes on average.

Visuo-spatial short-term memory task. To control for visuo-spatial short-term memory capacity, a Corsi block tapping task was administered to all participants⁴⁸. At the start of every trial, a grid of 9 grey-colored squares (35 × 35 mm) was presented (1200 ms). Three of the squares would sequentially turn black (each for 1000 ms with 500 ms in between). After a 1000 ms blank screen, participants were asked to report the order of appearance of the squares, by clicking on the square in the order in which they were presented. Correctly reproducing the order in three trials would increase the number of squares by one. The maximum number of squares was 4. Failing to reproduce the order of presentation in three consecutive trials would terminate the task, providing the span length for that participant. A first practice trial with feedback was presented in the beginning of the task. The total task duration was on average 5 minutes.

Analysis. First, accuracy at the cross-modal cueing task was analyzed. A repeated measures analysis of variance (rANOVA) was performed on the accuracy data, with between-subjects factor group (musicians, controls), and within-subjects factors cue modality (auditory, visual), target modality (auditory, visual), compatibility frequency in the block (80/20, 50/50), and compatibility in the trial (compatible, incompatible). Second, the reaction times (RTs) of the cross-modal cueing task were analyzed. A rANOVA was conducted on this data, with between-subjects factor group (musicians, controls), and within-subjects factors cue modality (auditory, visual), target modality (auditory, visual), compatibility frequency in the block (80/20, 50/50), and compatibility in the trial (compatible, incompatible.) Third, the scores at the short-term memory tasks were computed. Following conventions from the literature, in both tasks accuracy was calculated as the longest sequence correctly reproduced, multiplied by the total number of correctly reproduced sequences. For example, in the letter span tasks, a participant who correctly reproduced a 6-letter span and overall reproduced 11 spans correctly received an accuracy score of 66. Two-sample t-tests were performed on the verbal short-term memory task scores and visuo-spatial short-term memory task to test for differences in short-term memory capacity between musicians and controls.

References

1. Cross, I. Music in the evolution of the mind. *Trends Neurosci.* **24**, 190 (2001).
2. Mithen, S. Singing in the brain. *New Sci.* **197**, 38–39 (2008).
3. Savage, P. E., Brown, S., Sakai, E. & Currie, T. E. Statistical universals reveal the structures and functions of human music. *Proc. Natl. Acad. Sci. USA* **112**, 8987–92 (2015).
4. Vuust, P., Gebauer, L. K. & Witek, M. A. G. Neural underpinnings of music: the polyrhythmic brain. *Adv. Exp. Med. Biol.* **829**, 339–356 (2014).
5. Maes, P.-J., Leman, M., Palmer, C. & Wanderley, M. M. Action-based effects on music perception. *Front. Psychol.* **4**, 1008 (2014).
6. Zatorre, R. J., Chen, J. L. & Penhune, V. B. When the brain plays music: auditory-motor interactions in music perception and production. *Nat. Rev. Neurosci.* **8**, 547–558 (2007).
7. Zatorre, R. Music, the food of neuroscience? *Nature* **434**, 312–315 (2005).
8. Zatorre, R. J. & Salimpoor, V. N. From perception to pleasure: music and its neural substrates. *Proc. Natl. Acad. Sci.* **110**, 10430–10437 (2013).
9. Zuk, J., Benjamin, C., Kenyon, A. & Gaab, N. Behavioral and neural correlates of executive functioning in musicians and non-musicians. *PLoS One* **9**, e99868 (2014).
10. Herholz, S. C. & Zatorre, R. J. Musical Training as a Framework for Brain Plasticity: Behavior, Function, and Structure. *Neuron* **76**, 486–502 (2012).
11. Hodges, D. A., Hairston, W. D. & Burdette, J. H. Aspects of multisensory perception: the integration of visual and auditory information in musical experiences. *Ann. N. Y. Acad. Sci.* **1060**, 175–185 (2005).
12. Carey, D. *et al.* Generality and specificity in the effects of musical expertise on perception and cognition. *Cognition* **137**, 81–105 (2015).
13. Kraus, N. & Chandrasekaran, B. Music training for the development of auditory skills. *Nat. Rev. Neurosci.* **11**, 599–605 (2010).
14. Lim, A. & Sinnett, S. Exploring Visual Attention in Musicians: Temporal, Spatial and Capacity Considerations. In Carlson, L., Hölscher, C., & Shipley, T. *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 580–585). Austin, TX: Cognitive Science Society (2011).
15. Moreno, S. & Bidelman, G. M. Examining neural plasticity and cognitive benefit through the unique lens of musical training. *Hear. Res.* **308**, 84–97 (2014).
16. Rauscher, F. H., Shaw, G. L. & Ky, K. N. Music and spatial task performance. *Nature* **365**(6447), 611 (1993).
17. Schellenberg, E. G. Cognitive Performance after listening to music: A review of the Mozart Effect. In MacDonald, R., Kreuz, G. & Mitchell, L. *Music, health, and wellbeing* 324–338. Oxford: Oxford University Press (2012).
18. Aagten-Murphy, D., Cappagli, G. & Burr, D. Musical training generalises across modalities and reveals efficient and adaptive mechanisms for reproducing temporal intervals. *Acta Psychol. (Amst.)* **147**, 25–33 (2014).
19. Bishop, L. & Goebel, W. Context-specific effects of musical expertise on audiovisual integration. *Front. Psychol.* **5**, 1123 (2014).
20. Helmbold, N., Rammsayer, T. & Altenmüller, E. Differences in primary mental abilities between musicians and nonmusicians. *J. Individ. Differ.* **26**, 74–85 (2005).
21. Patston, L. L., Hogg, S. L. & Tippett, L. J. Attention in musicians is more bilateral than in non-musicians. *Laterality* **12**, 262–272 (2007).
22. Moreno, S. *et al.* Musical training influences linguistic abilities in 8-year-old children: more evidence for brain plasticity. *Cereb. Cortex* **19**, 712–723 (2009).
23. Strait, D. L., Kraus, N., Parbery-Clark, A. & Ashley, R. Musical experience shapes top-down auditory mechanisms: evidence from masking and auditory attention performance. *Hear. Res.* **261**, 22–29 (2010).
24. Tierney, A. T., Bergeson-Dana, T. R. & Pisoni, D. B. Effects of early musical experience on auditory sequence memory. *Empir. Musicol. Rev. EMR* **3**, 178 (2008).
25. Lee, H. & Noppeney, U. Long-term music training tunes how the brain temporally binds signals from multiple senses. *Proc. Natl. Acad. Sci.* **108**, E1441–E1450 (2011).
26. Friston, K. & Kiebel, S. Cortical circuits for perceptual inference. *Neural Netw.* **22**, 1093–1104 (2009).
27. Lee, H. & Noppeney, U. Temporal prediction errors in visual and auditory cortices. *Curr. Biol.* **24**, R309–R310 (2014).
28. DenOuden, H. E. M. D., Friston, K. J., Daw, N. D., McIntosh, A. R. & Stephan, K. E. A dual role for prediction error in associative learning. *Cereb. Cortex* **19**, 1175–1185 (2009).
29. Vassena, E., Krebs, R. M., Silvetti, M., Fias, W. & Verguts, T. Dissociating contributions of ACC and vmPFC in reward prediction, outcome, and choice. *Neuropsychologia* **59**, 112–123 (2014).
30. Clark, A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* **36**, 181–204 (2013).
31. Summerfield, C. & Egner, T. Expectation (and attention) in visual cognition. *Trends Cogn. Sci.* **13**, 403–409 (2009).
32. Vuust, P. & Witek, M. A. G. Rhythmic complexity and predictive coding: a novel approach to modeling rhythm and meter perception in music. *Audit. Cogn. Neurosci.* **5**, 1111 (2014).
33. Maes, P.-J. Sensorimotor Grounding of Musical Embodiment and the Role of Prediction: A Review. *Front. Psychol.* **308** doi:10.3389/fpsyg.2016.00308 (2016).
34. Schaefer, R. S., Overy, K. & Nelson, P. Affect and non-uniform characteristics of predictive processing in musical behaviour. *Behav. Brain Sci.* **36**, 226–227 (2013).
35. Kuchenbuch, A., Paraskevopoulos, E., Herholz, S. C. & Pantev, C. Effects of musical training and event probabilities on encoding of complex tone patterns. *BMC Neurosci.* **14**, 51 (2013).
36. Oechslin, M. S., Van De Ville, D., Lazeyras, F., Hauert, C.-A. & James, C. E. Degree of Musical Expertise Modulates Higher Order Brain Functioning. *Cereb. Cortex* **23**, 2213–2224 (2013).
37. Vuust, P., Ostergaard, L., Pallesen, K. J., Bailey, C. & Roepstorff, A. Predictive coding of music – Brain responses to rhythmic incongruity. *Cortex* **45**, 80–92 (2009).
38. Mähdorf, C., Rieger, M., Prinz, W. & Koelsch, S. Nobody Is Perfect: ERP Effects Prior to Performance Errors in Musicians Indicate Fast Monitoring Processes. *PLoS ONE* **4**, e5032 (2009).
39. Wilcox, R. R. & Keselman, H. J. Modern robust data analysis methods: measures of central tendency. *Psychol. Methods* **8**, 254–274 (2003).
40. Fischer, R., Dreisbach, G. & Goschke, T. Context-sensitive adjustments of cognitive control: conflict-adaptation effects are modulated by processing demands of the ongoing task. *J. Exp. Psychol. Learn. Mem. Cogn.* **34**, 712–718 (2008).
41. Schmidhuber, J. In *Anticipatory Behavior in Adaptive Learning Systems* (eds. Pezzulo, G., Butz, M. V., Sigaud, O. & Baldassarre, G.) 48–76 (Springer Berlin Heidelberg, 2009).
42. Patel, A. D. The OPERA hypothesis: assumptions and clarifications. *Ann. N. Y. Acad. Sci.* **1252**, 124–128 (2012).
43. Patel, A. D. Why would Musical Training Benefit the Neural Encoding of Speech? The OPERA Hypothesis. *Front. Psychol.* **2**, 142 (2011).
44. Cameron, D. J. & Grahn, J. A. Neuroscientific investigations of musical rhythm. *Acoust. Aust.* **42**, 111 (2014).
45. O’Doherty, J. P., Dayan, P., Friston, K., Critchley, H. & Dolan, R. J. Temporal difference models and reward-related learning in the human brain. *Neuron* **38**, 329–337 (2003).
46. Wacongne, C., Changeux, J.-P. & Dehaene, S. A neuronal model of predictive coding accounting for the mismatch negativity. *J. Neurosci.* **32**, 3665–3678 (2012).

47. Daneman, M. & Carpenter, P. A. Individual differences in working memory and reading. *J. Verbal Learn. Verbal Behav.* **19**, 450–466 (1980).

48. Corsi, P. *Human memory and the medial temporal region of the brain.* (McGill University, 1972).

Acknowledgements

We thank Esther De Loof for useful discussion, and Jean-Philippe Van Dijk for advice on the short-term memory tasks. The project was funded by Ghent University GOA grant BOF08/GOA011 and by the Ghent University Multidisciplinary Research Partnership “The integrative neuroscience of behavioral control”.

Author Contributions

All authors developed the study concept and contributed to the experimental design. E.V., K.K. and T.V. drafted the manuscript. J.L. conducted data collection. E.V. and J.L. analyzed the data. All authors approved the final version of the manuscript for submission.

Additional Information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Vassena, E. *et al.* Unimodal and cross-modal prediction is enhanced in musicians. *Sci. Rep.* **6**, 25225; doi: 10.1038/srep25225 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>