

Neural correlates of audiovisual integration in music reading



Emily S. Nichols*, Jessica A. Grahn

The University of Western Ontario, The Brain and Mind Institute, Department of Psychology, The University of Western Ontario, London, Ontario, Canada N6C5B7

ARTICLE INFO

Article history:

Received 13 May 2016

Received in revised form

9 August 2016

Accepted 10 August 2016

Available online 11 August 2016

Keywords:

Audiovisual integration

Music

Reading

Event-related potentials

Expertise

MMN

P300

ABSTRACT

Integration of auditory and visual information is important to both language and music. In the linguistic domain, audiovisual integration alters event-related potentials (ERPs) at early stages of processing (the mismatch negativity (MMN)) as well as later stages (P300 (Andres et al., 2011)). However, the role of experience in audiovisual integration is unclear, as reading experience is generally confounded with developmental stage. Here we tested whether audiovisual integration of music appears similar to reading, and how musical experience altered integration. We compared brain responses in musicians and non-musicians on an auditory pitch-interval oddball task that evoked the MMN and P300, while manipulating whether visual pitch-interval information was congruent or incongruent with the auditory information. We predicted that the MMN and P300 would be largest when both auditory and visual stimuli deviated, because audiovisual integration would increase the neural response when the deviants were congruent. The results indicated that scalp topography differed between musicians and non-musicians for both the MMN and P300 response to deviants. Interestingly, musicians' musical training modulated integration of congruent deviants at both early and late stages of processing. We propose that early in the processing stream, visual information may guide interpretation of auditory information, leading to a larger MMN when auditory and visual information mismatch. At later attentional stages, integration of the auditory and visual stimuli leads to a larger P300 amplitude. Thus, experience with musical visual notation shapes the way the brain integrates abstract sound-symbol pairings, suggesting that musicians can indeed inform us about the role of experience in audiovisual integration.

© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The ability to integrate auditory and visual information is fundamental to creating a unified percept of the world. Humans routinely integrate audiovisual (AV) information from the natural environment, but we also learn to integrate arbitrary sound-sight pairings, such as when learning to read. Indeed, a strong predictor of reading success is the ability to integrate speech sounds with visual orthography (Bryant et al., 1990; Liberman et al., 1974; Mann and Liberman, 1984). Similar to reading words, humans also learn to read music, and reading of music notation also requires AV integration, in the form of written notes with musical sounds. However, compared to word reading, music reading offers a unique opportunity to study the mechanisms of learned AV integration, as identical AV stimuli can be presented to musicians, who have learned to integrate the stimuli, and non-musicians, who have not. Compared to word reading, one advantage is that non-musicians are commonly available for testing (unlike individuals who are illiterate). In addition, lack of musical

reading skill in non-musicians is not confounded with being at an earlier developmental stage or having a reading disorder.

Much research has been conducted on AV integration in reading, and AV integration in both typical and atypical reading has frequently been investigated using event-related potentials (Andres et al., 2011; Froyen et al., 2009, 2008; Herdman et al., 2006; Raji et al., 2000). The mismatch negativity (MMN) is an early (160–220 ms post stimulus onset), automatic fronto-central negativity evoked by an auditory stimulus that deviates from a standard sequence, and is thought to reflect the comparison of a short-lived memory trace with the current stimulus (Näätänen et al., 1978, 2007). The MMN can also be evoked to visual deviants, with different neural generators and topography to the auditory MMN (Pazo-Alvarez et al., 2003). The MMN has been used extensively to index the speed and automaticity of AV integration in reading (Andres et al., 2011; Froyen et al., 2008, 2009). For example, in a study by Andres et al. (2011), subjects monitored visually presented letters for vowels, while 'standard' letter names were simultaneously presented aurally. When the auditory letter name deviated from the standard, and the visual letter matched the deviant letter name (i.e., the deviant was congruent with the visual stimulus), a larger MMN was elicited than when the auditory letter name deviated but the

* Corresponding author.

E-mail address: enicho4@uwo.ca (E.S. Nichols).

visual letter did not match the deviant letter name (i.e., the deviant was not congruent with the visual stimulus). While an MMN in response to an auditory deviant reflects a violation of expectancy, the larger MMN when the visual stimulus is congruent is thought to reflect integration of the two stimuli. Because of the early and automatic nature of the MMN, the larger amplitude to congruent deviations suggests that integration of written and auditory forms of language begins pre-attentively. This is consistent with previous studies that have also shown early, automatic AV integration of linguistic stimuli (Froyen et al., 2008, 2009; Herdman et al., 2006). A similar design has been applied to studies of dyslexia, and, although results are less consistent, children with dyslexia have been found to differ from typically developing children in the latency and size of the MMN (Žarić et al., 2014) or in whether early integration occurs at all (Froyen et al., 2011).

Although most studies of AV integration in reading have focused on early integration, later, post-attentive, effects have also been found. For example, in addition to the previously discussed MMN effect, Andres et al. (2011) found that the P300 component occurred later for incongruent deviant trials than congruent deviant trials. The authors thus suggested that visual attention is influenced by auditory input, even in the absence of overt attention to the auditory stimulus, and that congruency of the auditory and visual stimuli may have reduced the effort needed to categorize the visual stimulus. Indications of later AV integration have also been found in work by Raji et al. (2000). Thus, AV integration appears to occur at both early and late stages of processing, and to consist of both automatic processes as well as post-attentive binding of auditory and visual input.

Much less is known about AV integration in music reading, although musical experience appears to shape integration of sounds and musical actions, as demonstrated by studies performed in musicians who are hearing and seeing musical instrument actions (Behne et al., 2013; Lee and Noppney, 2014; Petrini et al., 2010; Proverbio et al., 2014; Vatakis and Spence, 2006). Currently, it is unclear at what point in the processing stream AV integration in music reading occurs. That is, the temporal characteristics of AV integration of musical stimuli, and whether they are similar to reading, are unknown. However, recent studies have compared musicians and non-musicians on measures of both musical and linguistic integration (e.g., using videos of natural speech as well as videos of a musician playing an instrument; Behne et al., 2013; Lee and Noppney, 2014). By manipulating the stimulus onset asynchrony (SOA) of the auditory and visual stimuli, it was found that musicians require a shorter SOA in order to perceive the auditory and visual streams as synchronous, regardless of whether the stream was linguistic or musical. The authors suggested that musicians possess a narrower time window in which AV information is integrated than non-musicians, and that practicing music fine-tunes AV integration, which generalizes to other forms of input.

Musical training has also been shown to influence how novel AV rules are processed. In a study using magnetoencephalography (MEG) by Paraskevopoulos et al. (2012), musicians and non-musicians learned an abstract rule similar to music notation, stating that the height of a circle on the screen should be associated with the pitch of a note. Similar to previous studies, the authors modulated congruency by presenting auditory deviants, visual deviants, and AV deviants. The authors found that although both musicians and non-musicians were able to learn the rule and showed unique oddball responses to each form of deviant, musicians showed more activity in right superior frontal gyrus, right superior temporal gyrus, and right lingual gyrus, than non-musicians. From this, the authors concluded that musical training promotes plasticity in these regions, which are involved in both unisensory and multisensory experiences.

Currently, several questions remain about how auditory and written musical information is integrated. Specifically, it is unclear whether the type of arbitrary sound/symbol pairing matters for AV integration. That is, does experience influence the nature of the

integration, as might happen when reading music or with a learned novel rule, or is integration similar for individuals with and without knowledge of the relationship between the stimuli? It is also unclear at what point in the processing stream musical information is integrated, and whether this is similar to linguistic information. Here, we examined whether musicians integrate AV musical information early (i.e., before 200 ms post-stimulus) as well as late (i.e., after 250 ms post-stimulus). We also investigated whether musicians integrate AV musical information differently from non-musicians. The fact that some individuals have no musical training affords an excellent opportunity to examine whether knowledge of written music is indeed required for integration, without requiring a comparison to a population with a disorder, as is the case in reading. By examining both musician and non-musicians, we can separate the contribution of musical experience and understanding of the written notation from simple simultaneous change detection in two modalities.

To address the questions of when AV integration occurs in music reading, and whether musical AV integration in musicians differs from non-musicians, we employed an auditory pitch-interval oddball task to evoke the MMN and P300, while manipulating whether visual pitch-interval information was congruent or incongruent with the auditory information. We tested musicians and non-musicians. We hypothesized that the MMN and P300 would be larger when both the auditory and visual stimuli deviated from the standard, as AV integration would increase the evoked response when the auditory and visual deviants were congruent. We also hypothesized that this would only occur in musicians, and not in non-musicians, as understanding the information conveyed by written music would be crucial for integration. Modulation of the MMN by congruency of the deviants would indicate that integration occurs early and automatically, similarly to reading. Modulation of the P300 by congruency of the deviants would suggest that additional, post-attentive integration also occurs later in processing. A difference between musicians and non-musicians would suggest that integration of musical information is dependent on knowledge of, or experience with, the type of information to be integrated, and non-musicians are not integrating the information in the same way, as they do not have the same understanding of the information as musicians. The comparison of musicians to non-musicians is necessary, as non-musicians will still perceive the deviants in both visual and auditory stimuli. By including non-musicians, we ensure that our findings reflect integrating the information derived from musical experience, not simply the response to receiving two oddball stimuli at once, or by implicitly learning to pair two deviating stimuli.

2. Methods

2.1. Subjects

Forty right-handed neurologically healthy adults were recruited from the University of Western Ontario community. Musicians were individuals who reported having at least one year of formal musical training, and consisted of 20 subjects (13 female), ranging in age from 18 to 57, with a mean of 10.90 ($SD=4.91$) years of formal musical training. Non-musicians were individuals who reported having less than one year of formal musical training, and consisted of 20 subjects (13 female), ranging in age from 18 to 57, with a mean of .10 ($SD=.31$) years of formal musical training. Musicians and non-musicians were matched on number of years of education ($t(36.37)=1.63$, $p=.112$, *ns*) and socioeconomic status (SES; $t(36.37)=.82$, $p=.418$, *ns*). A summary of subject variables is provided in Table 1.

2.2. Behavioral music inventories

Participants completed two subscales of the Goldsmiths Musical Sophistication Index (GMSI; Müllensiefen et al., 2014). The

Table 1
Subject descriptives.

Measure	Mean (SD)		<i>t</i> (<i>df</i>)	<i>p</i>
	Musicians	Non-musicians		
Age	23.20 (8.91)	22.65 (8.59)	.20(37.95)	.844
Education (# years)	15.98 (3.41)	14.43 (2.54)	1.63(35.13)	.112
SES	4.80 (1.58)	4.37 (1.71)	.82(37)	.418
GMSI Perceptual Abilities	52.55 (7.67)	34.55 (8.94)	6.83(37.15)	< .001
GMSI musical training	32.55 (6.92)	7.85 (3.45)	14.29 (27.92)	< .001
Years of formal training	10.90 (4.91)	.10 (.31)	9.82(19.15)	< .001
Sight-reading ability	5.25 (.92)	1.25 (.72)	15.29 (35.77)	< .001

Note: Welch's *t*-test used due to unequal variances assumed between groups.

first was the Perceptual Abilities subscale, used to assess participants' self-reported musical listening skills (e.g., "I can tell when people sing or play out of tune"), and has a maximum score of 63. The second was the Musical Training subscale, used to assess the extent of participants' training and practice (e.g., "At the peak of my interest I practiced on my primary instrument including voice for ___ hours per day"), and has a maximum score of 49. Participants also answered questions about the number of years of formal training they have received, and their self-reported sight-reading ability on a Likert scale of one to seven (Table 1).

2.3. Stimuli and procedure

Auditory stimuli consisted of two 300 ms pure tones, sounded one after the other, with a sampling rate of 44,100 Hz and 10 ms rise and fall times, created using the Audacity 2.1 software program. The stimulus onset asynchrony between first and second tone was 300 ms. Visual stimuli consisted of images of two musical notes on a treble-clef staff at 700 × 311 pixels resolution. Stimuli were one of three types of musical intervals: a major second, perfect fourth, or a major sixth. A fixation cross appeared in the middle of the image; the first note of the interval was presented to the left of fixation, and the second note was presented to the right of fixation. The first auditory tone was presented simultaneously with the left visual note, and the second auditory tone was presented simultaneously with the right visual note.

Stimulus presentation is shown in Fig. 1. EEG testing was divided into four blocks of different keys (C major, G major, D major, and A major). In each block, a standard interval (e.g., a major second), in

which visual and auditory information was congruent, was presented in 85% of trials. In 15% of trials, either the visual stimulus deviated from the standard (e.g., a perfect fourth deviant was seen while a major second standard was heard), the auditory stimulus deviated from the standard (e.g., a major second standard was seen while a perfect fourth deviant was heard), or both the visual and auditory stimuli deviated from the standard (e.g., a perfect fourth deviant was seen and heard). Within each block the standard changed twice, so that the three types of intervals were each used as the standard, with one of the other two intervals used as deviants. The first note of the interval remained the same within a block. There was a five second pause between standard-interval changes, and participants were notified of the key before each block. The complete set of visual stimuli is available in Appendix A.

During testing, participants were seated comfortably in front of a CRT monitor, with speakers placed to the left and right of the monitors, set at 60 dB. Participants performed a target detection task in which they monitored the visual stimuli and pressed the "1" key on the keyboard when they saw an "X" in place of the second note in the interval, which was a congruent standard interval. Target trials consisted of 22% of the standard congruent trials, and 18.75% of total trials. Participants were instructed to ignore the auditory stimuli. Auditory and visual stimuli were presented concurrently; the first note appeared onscreen for 300 ms, after which the second note appeared for an additional 300 ms. The first note remained onscreen throughout the trial; this was done to help equalize working memory demands in the visual and auditory modalities, as visual sensory memory is temporally short (0.3–1.0 s; Averbach and Coriell, 1961; Eriksen and Collins, 1967; Sakitt, 1976; Sperling, 1960) relative to auditory sensory memory (1.5–4.0 s; Cowan, 1984; Crowder, 1982; Darwin et al., 1972). The onset of each of the two auditory tones was synchronized to the appearance of the visual tones. Each trial lasted a total of 600 ms, and the inter-trial interval was jittered from 600 to 1500 ms at 300 ms intervals. Each block consisted of 480 trials each, for a total of 1920 trials. Two presentation orders were created, one the reverse of the other, and presentation order was counterbalanced across participants.

2.4. EEG recording and preprocessing

Stimuli were presented using the E-Prime 2.0 software package (Schneider et al., 2002). Continuous EEG data was collected using Neuroscan software from 32 scalp electrodes (Fp1/2, AF3/4, F7/8, F3/4, T7/8, C3/4, CP5/6, CP1/2, P7/8, P3/4, PO3/4, O1/2, Fz, Cz, Pz, Oz) and two mastoid electrodes, and electrooculogram (EOG) was recorded from four face electrodes placed above and below the left eye and on the outer canthus of each eye using the BioSemi ActiveTwo EEG

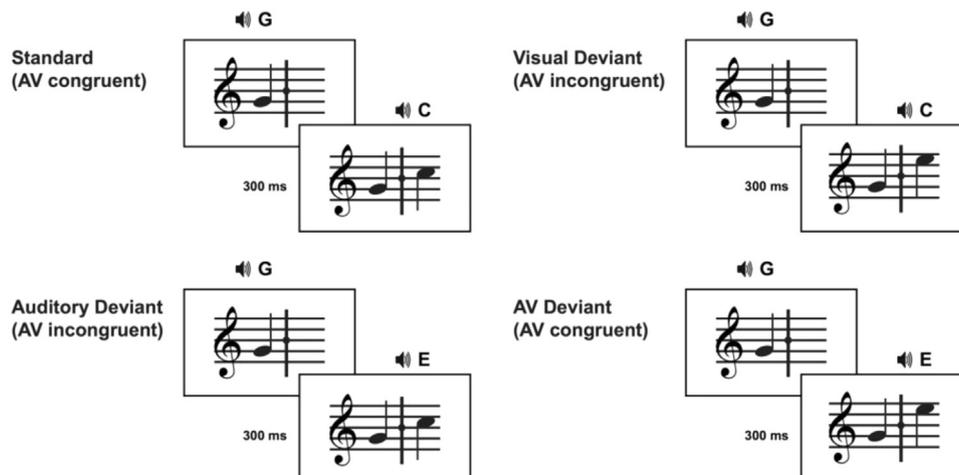


Fig. 1. Presentation of stimuli. The initial musical note and tone of the interval was presented for 300 ms, then the second note and tone was presented. The first note remained onscreen during the presentation of the second note, and tones and musical notes were presented simultaneously.

system (BioSemi Inc., Amsterdam, The Netherlands) consisting of amplifier-embedded Ag/AgCl electrodes arranged according to the International 20–30 system. A Common Mode Sense active electrode and a Driven Right Leg passive electrode were used as the ground. Data was recorded in the frequency range of .1–100 Hz at a 512 Hz sampling rate, with impedances below 20 k Ω .

ERP data was processed using EEGLAB software (Delorme and Makeig, 2004) and the ERPLAB add-on software (Markley et al., 2011). After importing the data, EEG data underwent a .1–30 Hz bandpass filter with a 60 Hz notch filter to remove line and muscle noise. EEG data was segmented into –200 to 800 ms single-trial epochs of each condition of interest (congruent standard, incongruent standard, congruent deviant, incongruent deviant), and baseline correction with a prestimulus baseline of –200 to 0 ms was applied. Eye movements and other artifacts were removed using a –100 to 100 μ V threshold.

Because the difference in signal-to-noise ratio between conditions can be an issue for traditional types of analyses, we used mixed-effects models to circumvent this problem, as they are capable of fitting unbalanced datasets. Mixed-effects models are specifically useful in cases where some trials necessarily occur more than others (e.g., an oddball paradigm). We employed a mixed-effects analysis of the data for this reason, as differences in signal-to-noise ratio are a great concern in traditional GLM analyses (Baayen et al., 2008; Tibon and Levy, 2015). The present study used the *lme4* (Bates et al., 2015) and *LMERConvenienceFunctions* (Tremblay and Ransijn, 2015) packages for R (R Core Team, 2015).

To examine the MMN, mean amplitude between 160 and 220 ms (a timeframe typical of the MMN; Näätänen and Kreegipuu, 2012) was computed over nine electrodes (F3/4, C3/4, P3/4, Fz, Cz, and Pz), to provide coverage over anterior, middle, and posterior regions over both hemispheres. To ensure that the auditory mismatch was eliciting an MMN, congruent and incongruent standard trials were averaged, and congruent and incongruent deviant trials were averaged separately. Amplitudes were then submitted to LME analysis with auditory mismatch (standard/deviant), group (musician/non-musician), and electrode position as fixed effects and participants as a random effect. To assess whether congruency of auditory and visual stimuli affected the amplitude of the MMN, standard trials were subtracted from deviant trials for each congruency condition (i.e., AV Deviants minus Standards, Auditory Deviants minus Visual Deviants). The MMN is commonly analyzed as a difference between standards and deviants, and by examining these differences waves we were able to ensure that we were isolating the congruency effect from auditory or visual mismatch effects alone. Amplitudes were then submitted to LME analysis with congruency, group (musician/non-musician), and electrode position as fixed effects and participant as a random effect.

To quantify the P300, mean amplitude between 250 and 500 ms (a typical timeframe of the P300) was computed over the same nine electrodes. Similar to the MMN, we examined auditory mismatch by first averaging congruent and incongruent standard trials and averaging congruent and incongruent deviant trials. LME was again used with mismatch (standard/deviant), group (musician/non-musician), and electrode position as fixed factors and participants modeled as a random effect. To examine the effect of congruency on the P300, we analyzed standards and deviants separately. This is because in addition to the auditory deviant, the visual oddball was also expected to influence the P300, thus congruency effects could be expected in both standards and deviants. Amplitudes were again submitted to LME analysis with congruency (congruent/incongruent), group (musician/non-musician), and electrode position as fixed factors and participant as a random effect. Satterthwaite approximation was used to estimate degrees of freedom for all analyses, and significant interactions were followed up with post-hoc paired two-tailed *t*-tests.

3. Results

3.1. Behavioral

Group measures on the GMSI Perceptual Abilities and Musical Training subscales are summarized in Table 1. Musicians reported significantly better musical listening skills than non-musicians, with musicians ranging in GMSI Perceptual Abilities score from 32 to 63, and non-musicians ranging from 15 to 42. Musicians also had significantly more musical training than non-musicians, with musicians ranging in GMSI Musical Training score from 21 to 44, and non-musicians ranging from 6 to 20. Musicians also reported having significantly better sight-reading abilities than non-musicians.

3.2. Mismatch negativity

To confirm that the auditory deviants produced an MMN, we first compared the standard and deviant trials, averaged across congruency, in the 160–220 ms time window. Deviant trials did indeed produce an MMN, with amplitude to deviants being significantly more negative than amplitude to standards (as evidenced by a main effect of mismatch, $F(1, 501)=67.56, p < .001$). MMN amplitudes were similar between musicians (Fig. 2A) and non-musicians (Fig. 2B), as supported by no significant three-way interaction ($F(8, 646)=.27, p=.975, ns$) nor group \times mismatch interaction ($F(1, 501)=.00, p=.953, ns$). However, a significant group \times electrode interaction showed that scalp topography differed between groups ($F(8, 646)=2.12, p=.032$). Bonferroni corrected post-hoc *t*-tests revealed that non-musicians had significantly more negative amplitudes than musicians at electrodes Cz ($t(604)=2.05, p=.041$) and F4 ($t(604)=2.77, p=.006$), suggesting that non-musicians had a more right-lateralized response to deviants than musicians. These results confirmed that the auditory deviant indeed produced an MMN in both congruent and incongruent deviant conditions, as expected, however scalp distributions between groups differed, as indexed by the group \times electrode interaction.

Next we examined the effect of congruency, after subtracting standard trials from deviant trials for each congruency condition (i.e., AV Deviants minus AV Standards, Auditory Deviants minus Visual Deviants). Difference waves are shown in Fig. 2C. The MMN was not modulated by congruency in either group. Although it appeared that incongruent trials led to a larger MMN, this difference failed to reach significance, as there was no main effect of congruency ($F(1, 38)=.56, p=.460, ns$). There was no significant group effect, as evidenced by a lack of three-way interaction ($F(8, 608)=.48, p=.870, ns$), and no significant two-way interactions. These results suggested that contrary to our predictions, at the group level, there was no indication of early AV integration in musicians or non-musicians.

As musical training varied widely across our musician group (see Table 1), we assessed whether GMSI musical training score modulated the MMN. We submitted MMN difference wave amplitudes for musicians to LME analysis with congruency and electrode as fixed factors and training as a continuous factor, with subjects as a random factor, and found a congruency \times training interaction ($F(1, 306)=6.03, p=.015$). Fig. 2D shows difference waves of two individual high-scoring and low-scoring musicians. Fig. 3 visualizes the congruency \times training interaction; amplitudes to congruent stimuli were less negative as training increased, but amplitudes to incongruent stimuli did not change with training. These results suggest that amongst those with knowledge of musical notation, the amount of training shapes the integration of written musical notes and auditory musical tones at early stages of processing. Specifically, more skilled musicians may be obtaining more information from the visual stimulus than less skilled musicians, leading to a smaller auditory MMN when the visual stimulus is congruent with (and therefore informative about) the auditory stimulus.

Because the effect of musical expertise was contradictory to

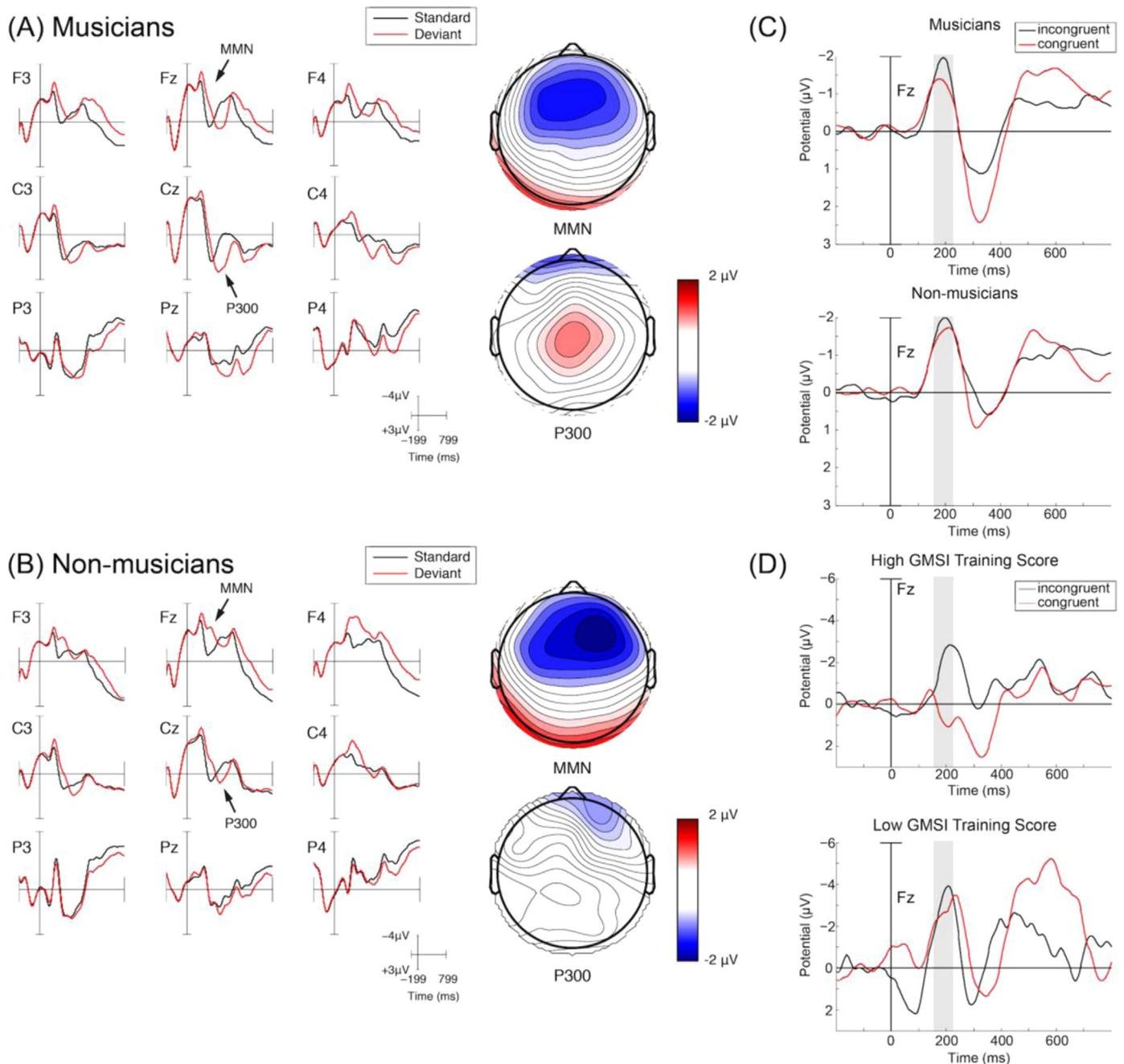


Fig. 2. Grand average waveforms ($N=20$) over nine scalp electrodes for standard and deviant trials, averaged across congruent and incongruent trials in musicians (A) and non-musicians (B). Non-musicians show a clear MMN to deviant trials, whereas musicians show a larger P300 than non-musicians to deviant trials. Topographic plots illustrate the group average subtraction of standards from deviants in mean voltage between 160 and 220 ms for the MMN, and 250–500 ms for the P300. (C) Grand average deviants-standards difference waves at electrode Fz for congruent and incongruent trials in musicians and non-musicians. The grey bar indicates the MMN time window. (D) Deviants-standards difference waves at electrode Fz for congruent and incongruent trials in a single high GMSI training score musician and a single low GMSI training score musician. The grey bar again indicates the MMN time window.

our predictions, we performed a post-hoc analysis of the visual MMN (vMMN) at occipital electrodes, using electrodes PO3, PO4, O1, Oz, and O2. As vMMN amplitudes have been shown to predict auditory MMN amplitudes (Horvath et al., 2013), we expected to see the complementary pattern, with vMMN amplitudes increasing with musical training. We submitted MMN difference wave amplitudes at occipital electrodes for musicians to LME analysis with congruency and electrode as fixed factors and training as a continuous factor, with subjects as a random factor. Results confirmed the complementary pattern, as evidenced by a significant congruency \times training interaction ($F(4, 144)=4.65, p=.033$). Fig. 4 again visualizes the

congruency \times training interaction; vMMN amplitude to AV incongruent deviants remained the same across training, while vMMN amplitude to AV congruent deviants increased (i.e., became more negative) with musical training. This effect appears to be concentrated over left occipital electrodes PO3 and O1, which is not surprising given that critical stimuli appeared in the right visual field. Amplitudes at electrodes that indexed the visual and auditory MMNs (O1 and Fz, respectively) were strongly negatively correlated ($r=-.69, p<.001$), indicating that as vMMN amplitudes increased, auditory MMN amplitudes decreased.

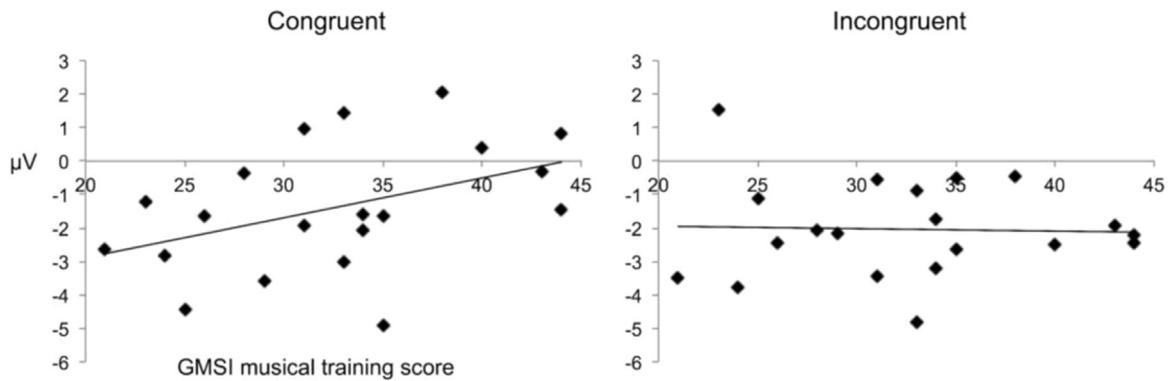


Fig. 3. Scatterplots of mean amplitudes (μV) in the MMN time window as a function of GMSI musical training score at electrode Fz in musicians, for congruent and incongruent trials.

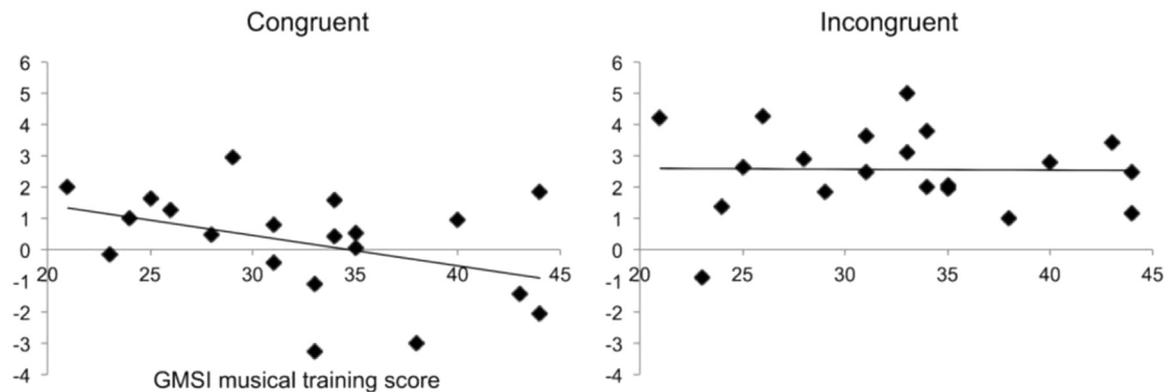


Fig. 4. Scatterplots of mean amplitudes (μV) in the MMN time window as a function of GMSI musical training score at electrode O1 in musicians, for congruent and incongruent trials.

3.3. P300

We next examined the effects of auditory mismatch on the P300. Similar to the MMN, we compared the standard and deviant trials, averaged across congruency in the 250–500 ms time window. There was a main effect of mismatch ($F(1, 202)=9.72, p=.002$), however, a significant group \times mismatch interaction revealed that the P300 was significantly larger in musicians than non-musicians ($F(1, 202)=4.09, p=.044$; Fig. 2). Although there was no three-way interaction, scalp topography did differ between groups, as shown by a significant group \times electrode interaction ($F(8, 646)=2.62, p=.008$). Bonferroni corrected post hoc t -tests revealed that musicians had significantly more positive amplitudes over electrodes C4 ($t(604)=2.94, p=.003$), Cz ($t(604)=4.01, p<.001$), and F4 ($t(604)=2.11, p=.036$). Again, these results confirmed that the oddball manipulation indeed produced a P300 in both groups, although it was larger and had different scalp topography in musicians compared to non-musicians.

When examining congruency of standard trials in the 250–500 ms time window, amplitudes to congruent trials were significantly more negative (smaller P300) than to incongruent trials ($F(1, 91)=27.37, p<.001$). This effect was similar across groups, as there was no significant three-way interaction ($F(8, 646)=.41, p=.913, ns$), and no significant two-way interactions. When examining congruency of deviant trials in the 250–500 ms time window, amplitudes to congruent trials were significantly more positive (larger P300) than to incongruent trials ($F(1, 138)=61.24, p<.001$, Fig. 5), and this effect was similar across groups as evidenced by a lack of congruency \times group interaction ($F(1, 138)=.02, p=.895, ns$). Although there was no three-way interaction, a significant group \times electrode interaction indicated that topography of the P300 differed between musicians and non-musicians ($F(8,$

$646)=3.82, p<.001$). Bonferroni corrected post hoc t -tests revealed larger amplitudes for musicians at electrodes C4 ($t(604)=2.91, p=.004$) and Cz ($t(604)=3.94, p<.001$). Thus, despite musicians showing in general larger P300s to deviants, at the group level there were no differences in *integration* between musicians and non-musicians at post-attentive stages of processing.

We next assessed whether GMSI musical training score modulated the P300. We submitted P300 amplitudes for musicians to LME analysis with congruency and electrode as fixed factors and training as a continuous factor, with subjects as a random factor. Standards and deviants were again analyzed separately. When examining standard trials, there were no significant interactions with training, however, in deviant trials there was a congruency \times training interaction ($F(1, 306)=5.89, p=.016$). Fig. 6 visualizes the congruency \times training interaction; musicians' P300 amplitudes at electrode Cz to congruent stimuli increased as musical training increased. Similar to the MMN results, these results suggest that amongst those with knowledge of musical notation, the amount of training shapes the integration between written musical notes and auditory musical tones at later stages of processing. However, in contrast to the MMN, P300 amplitude in musicians *increased*, rather than decreased, for congruent trials, suggesting that at post-attentive stages, integration may be driven by conscious perception of the two congruent, changing, stimuli.

4. Discussion

The present study used event-related potentials to identify the neural correlates of audiovisual integration in music reading. The effects of musical training were investigated by comparing brain responses in musicians and non-musicians on an auditory pitch-

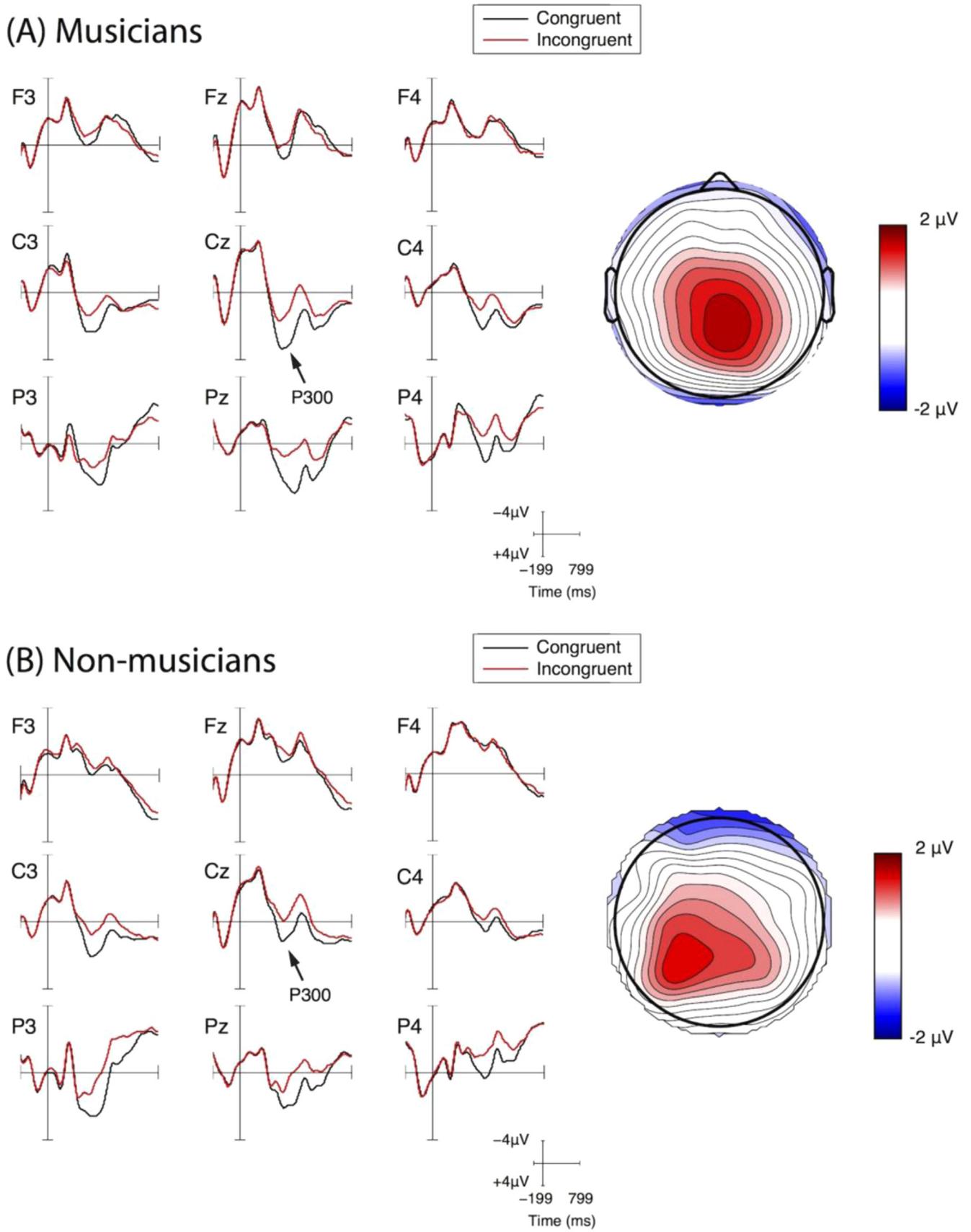


Fig. 5. Grand average waveforms ($N=20$) showing the P300 over nine scalp electrodes for congruent and incongruent deviant trials, in musicians (A) and non-musicians (B). Topographic plots illustrate the group average subtraction of incongruent from congruent in mean voltage between 250 and 500 ms.

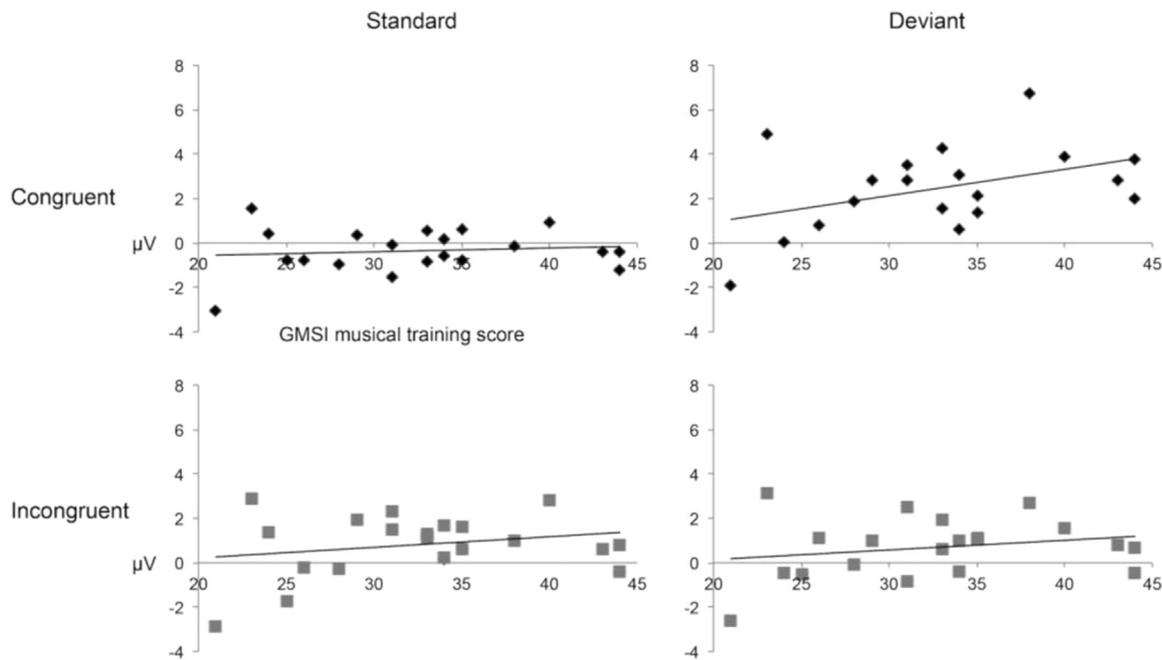


Fig. 6. Scatterplots of mean amplitudes (μV) in the P300 time window as a function of GMSI musical training score at electrode Cz for musicians.

interval oddball task that evoked the MMN and P300, while manipulating whether visual pitch-interval information was congruent or incongruent with the auditory information. Results indicated that scalp topography differed between groups in both the MMN and P300 response to deviants. Interestingly, musicians' musical training was shown to modulate integration of congruent deviants at both early and late stages of processing.

4.1. ERP markers of audiovisual integration

4.1.1. MMN

The elicited MMN in both groups by the auditory deviant condition confirms that the auditory oddball response is produced regardless of musical training. Consistent with previous research, the underlying neural generators differed between groups as evidenced by different scalp topographies; deviants elicited a more right-lateralized MMN in non-musicians compared to musicians. Similarly, musicians have shown more frontally-located MMN than non-musicians (Vuust et al., 2012), and non-musicians have been found to have a right-lateralized oddball response to deviant chords (Tervaniemi et al., 2000), including when compared to musicians (Minati et al., 2009). Additionally, musicians have shown higher levels of activation in the bilateral STG in brain responses typically associated with the MMN relative to non-musicians (Paraskevopoulos et al., 2012; 2013). Together, a consistent pattern emerges, with musicians producing fronto-central MMN topographies, and non-musicians producing right-lateralized MMN topographies.

The congruent deviant condition, when both the auditory and visual stimuli were deviant, did not elicit a change in MMN amplitude in either group. By definition, non-musicians had minimal musical training, therefore we were unable to investigate training in this group; however, the wide range of GMSI Musical Training scores in our musicians enabled us to examine the influence of musical training on the congruency effect within this group. Indeed, we found that in musicians, more training led to a smaller MMN. Previously, larger MMNs have been found in response to auditory stimuli in musicians with higher musical aptitude (Vuust et al., 2012), although the effect of musical training on AV integration has not been extensively studied. We found no effect of musical training on incongruent deviant trials; however, musicians with less musical training

did show greater MMN amplitude to congruent deviant trials, an effect that decreased with more musical training. One interpretation of these results is that visual information facilitates processing of the auditory information. Indeed, a recent study by Horvath et al. (2013) found that the visual and auditory MMN are linked, with the amplitude of the MMN in response to AV deviants being strongly correlated with subjects' visual responses. The authors suggested that there is an inhibitory interaction between the visual and auditory MMN responses, with the size of the auditory MMN depending on the size of the visual MMN. Although our results cannot distinguish whether there is a causal relationship between the auditory MMN and the vMMN, it is possible that musicians' visual response to the congruent deviant condition reduces the size of the auditory MMN, and that this effect is graded with level of musical training. If so, the auditory oddball in the congruent deviant condition would be more predictable than in the incongruent deviant condition, as it matches the visual stimulus. It is not unlikely that at ~ 170 ms, expectancy has developed; early (i.e., ~ 40 ms, 165 ms, 180 ms) AV interactions have been found, simultaneous with the onset of early visual sensory processing (Giard and Peronnet, 1999; Molholm et al., 2002) and with reading (Froyen et al., 2009; Mittag et al., 2011).

4.1.2. P300

Auditory deviants elicited a clear P300 in both musicians and non-musicians, although amplitude was significantly larger in the musician group. Previous research has shown similar effects, which may result from increased sensitivity or greater attention to auditory musical stimuli (George and Coch, 2011). Scalp topography also differed, with non-musicians showing a more left-lateralized P300 than musicians. Unlike the MMN, P300 amplitude was indeed modulated by congruency, with larger P300 amplitudes to congruent vs. incongruent stimuli. A larger P300 to congruent deviants was consistent with our predictions; integration of congruent deviant stimuli causes superadditivity of the auditory and visual oddball responses (Van Atteveldt et al., 2004). The size of the P300 was also significantly correlated with musical training. The congruency effect in the P300 with increasing musical training also confirms that the larger P300 in response to congruent deviants is not simply an effect of receiving two oddballs (auditory and visual) at once, but rather an effect of integrating the two deviants. While

both groups showed a larger P300 to congruent deviants than to incongruent deviants, this effect was modulated by experience in those with formal musical training. If the effects of congruency were simply due to the additive effects of simultaneous visual and auditory oddballs, experience would not modulate the size of the P300, suggesting that instead expertise is guiding AV integration.

Previous work has shown that both left and right superior temporal gyri are crucial in audiovisual integration of speech and letters (Van Atteveldt et al., 2004), and these areas are known to be neural generators of the P300 (Clark et al., 2000; Downar et al., 2000; Halgren et al., 1998; Horn et al., 2003; Kiehl et al., 2001; Linden et al., 1999; McCarthy et al., 1997; Menon et al., 1997; Mulert et al., 2004; Stevens et al., 2000). Van Atteveldt et al. (2004) suggested that the superadditive activity in bilateral STG indicates successful binding of cross-modal information. Thus, in our study, the graded effect of P300 amplitude with musical training suggests that training increased the binding of auditory and visual input.

Notably, the direction of the P300 interaction with musical training differs from the direction of the MMN interaction with musical training. While contrary to our predictions, each component reflects a different neural process, thus it is reasonable that the components interact differently with training. The fronto-central auditory MMN reflects the comparison of a short-lived auditory memory trace with the current stimulus, and is guided by the visual modality (Horvath et al., 2013). Accordingly, greater understanding of the visual stimulus (i.e., more musical training) leads to less conflict at this stage of processing, in turn leading to smaller MMN amplitudes. However, the P300 reflects post-attentive processing and is sensitive to both auditory and visual changes. The sensitivity to both modalities likely underlies the difference in MMN and P300 effects; both modalities evoke an oddball response, without one guiding the other, as is the case with the MMN. Instead, there is superadditivity of the response in the two modalities, and this superadditivity increases with musical training, perhaps because of stronger binding (Van Atteveldt et al., 2004). Thus, the differences in the underlying neural processes of the MMN and P300 result in the contrasting interactions with musical training.

4.2. Comparing AV integration in reading and music

We revealed several important similarities between reading linguistic information and music. In those with musical training, while the amplitude of the MMN to congruent deviants decreased as training increased, incongruent deviants produced a significant MMN, and amplitude remained the same across levels of training. In literate adults, Froyen et al. (2008) also found larger MMNs to incongruent letter-speech sound AV stimuli, in which only the auditory stimulus deviated, relative to auditory deviants alone. These results support the present findings, with both studies showing increased MMN amplitudes to incongruent deviants, suggesting that receiving incongruent visual information provides incorrect predictive knowledge.

Studies of reading development also provide insight into how AV integration in music and language differ, as the different stages of reading development can serve as a basis of comparison to various levels of musical training. In a similar task to their 2008 study, Froyen et al. (2010) studied AV integration in beginner readers of 8–9 years of age, and intermediate readers of 11 years of age. Beginner readers showed no difference in MMN amplitude between auditory deviants alone and AV deviants, similar to our non-musician group in which no congruency effect was found. Intermediate readers showed a larger MMN to the AV deviants, but only when the letter preceded the speech sound by at least 200 ms, which is earlier than required for adult readers. The authors suggested that beginner readers are not yet using visual information to inform their expectations, and that the ability to integrate letters and speech takes years to automate, explaining why they found graded integration effects in readers. Thus, reading

development shows a pattern parallel to the graded congruency effect we found for those with musical training, with amount of training affecting integration at both early and late stages of processing.

The current results suggest that integration in music notation differs from reading, in that there was not a consistent early effect of congruency in our musician group. This may be for several reasons. First, reading is an over-learned skill, beginning at ages 4–6 and heavily practiced throughout life. Although many musicians read music every day, reading words is a more frequent, crucial part of everyday life. Second, previous work on reading has studied undergraduates, who are generally typically-developing proficient readers with extensive reading experience. In contrast, our musician sample contained a wide variety of levels of musical training, with adult musicians with less training showing patterns of integration more similar to young, developing readers.

An outstanding question is whether the pattern observed in non-musicians would be comparable to non-readers. There is no readily available group in the reading literature comparable to typically-developing non-musicians; illiterate adults are difficult to find and often are not matched to literate adults on demographic variables such as socioeconomic status or level of schooling (Eme, 2011), and other non-readers are either individuals with dyslexia, which is a developmental disorder, or children who have not yet learned to read, confounding age and other developmental factors.

Additionally, although previous research supports the conclusion that visual information is indeed guiding the expectancy of the auditory stimulus (Froyen et al., 2009; Horvath et al., 2013), more research is needed to confirm this. As in studies of integration in reading, by varying the amount of time between auditory and visual stimuli, it can be determined how tightly paired a written musical note and auditory tone must be in order to be integrated, and whether this time window varies with musical training. After establishing the necessary window of integration, it could be determined whether visual information indeed affects expectancy of auditory information, by manipulating which type of information is presented first.

4.3. Conclusions

The current study has implications for our understanding of AV integration in music, specifically the effect of training on the integration of auditory and written music. While both early and late stages of AV integration seem to depend on expertise, the relationship between expertise and early/late integration differs. Early (i.e., ~170 ms) into processing, understanding the visual stimulus helps guide pre-attentive auditory perception and expectations. Later (~300 ms) into processing, there is a superadditive congruency effect, as the P300 is sensitive to both auditory and visual deviants. Training increases the congruency effect, as musicians with more training show larger P300s to congruent deviants. This suggests that the congruency of the deviants (which is meaningful to those with musical training) leads to the observed superadditivity. Taken together, continued musical training and experience with musical notation shapes the way the brain processes these abstract sound-symbol pairings, much like the development and mastery of linguistic reading.

Funding

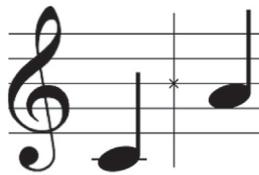
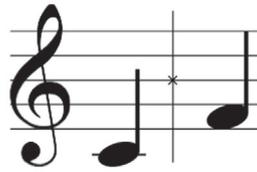
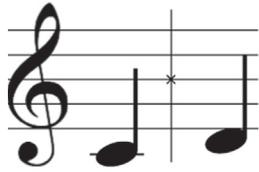
This study was supported by the James S. McDonnell Foundation and an NSERC Discovery Grant to [Grant number RGPIN/402557-201].

Acknowledgments

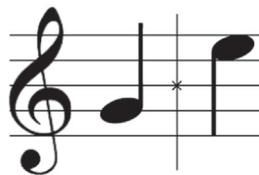
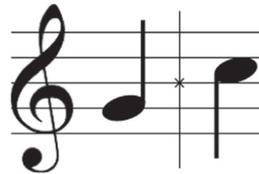
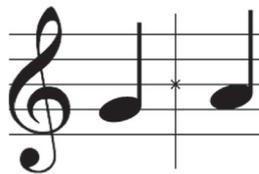
We are grateful to Alexis Harrison and Mikayla Keller for assistance with data collection.

Appendix A

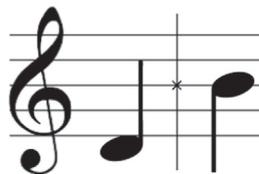
Key of C major



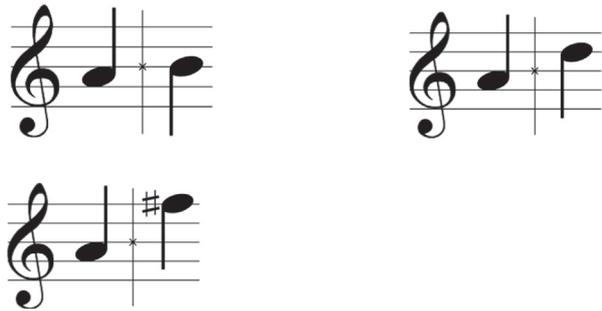
Key of G major



Key of D major



Key of A major



References

- Andres, A.J.D., Oram Cardy, J.E., Joanisse, M.F., 2011. Congruency of auditory sounds and visual letters modulates mismatch negativity and P300 event-related potentials. *Int. J. Psychophysiol.* 79 (2), 137–146.
- Averbach, E., Coriell, E., 1961. Short-term memory in vision. *Bell Syst. Tech. J.* 40 (1), 309–328.
- Baayen, R.H., Davidson, D.J., Bates, D.M., 2008. Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang.* 59 (4), 390–412.
- Bates, D., Maechler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67 (1), 1–48.
- Behne, D., Alm, M., Berg, A., Engell, T., Foy, C., Johnsen, C., Srigaran, T., Torsdottir, A. E., 2013. Effects of musical experience on perception of audiovisual synchrony for speech and music. *Proc. Meet. Acoust.* 19, 1–6.
- Bryant, P.E., Maclean, M., Bradley, L.L., Crossland, J., 1990. Rhyme and alliteration, phoneme detection, and learning to read. *Dev. Psychol.* 26 (3), 429–438.
- Clark, V.P., Fannon, S., Lai, S., Benson, R., Bauer, L., 2000. Responses to rare visual target and distractor stimuli using event-related fMRI. *J. Neurophysiol.* 83 (5), 3133–3139.
- Cowan, N., 1984. On short and long auditory stores. *Psychol. Bull.* 96 (2), 341–370.
- Crowder, R.G., 1982. Decay of auditory memory in vowel discrimination. *J. Exp. Psychol.: Learn. Mem. Cognit.* 8 (2), 153–162.
- Darwin, C.J., Turvey, M.T., Crowder, R.G., 1972. An auditory analogue of the Sperling partial report procedure: evidence for brief auditory storage. *Cognit. Psychol.* 3, 255–267.
- Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *J. Neurosci. Methods* 134, 9–21.
- Downar, J., Crawley, A.P., Mikulis, D.J., Davis, K.D., 2000. A multimodal cortical network for the detection of changes in the sensory environment. *Nat. Neurosci.* 3 (3), 277–283.
- Eme, E., 2011. Cognitive and psycholinguistic skills of adults who are functionally illiterate: current state of research and implications for adult education. *Appl. Cognit. Psychol.* 25 (5), 753–762.
- Eriksen, C.W., Collins, J.F., 1967. Some temporal characteristics of visual pattern perception. *J. Exp. Psychol.* 74, 476–484.
- Froyen, D., Willems, G., Blomert, L., 2011. Evidence for a specific cross-modal association deficit in dyslexia: an electrophysiological study of letter-speech sound processing. *Dev. Sci.* 14 (4), 635–648.
- Froyen, D., Atteveldt, N., Van, Blomert, L., 2010. Exploring the role of low level visual processing in letter-speech sound integration: a visual MMN study. *Front. Integr. Neurosci.* 4, 1–14.
- Froyen, D.J.W., van Atteveldt, N., Bonte, M., Blomert, L., 2008. Cross-modal enhancement of the MMN to speech-sounds indicates early and automatic integration of letters and speech-sounds. *Neurosci. Lett.* 430 (1), 23–28.
- Froyen, D.J.W., Bonte, M.L., van Atteveldt, N., Blomert, L., 2009. The long road to automation: neurocognitive development of letter-speech sound processing. *J. Cognit. Neurosci.* 21 (3), 567–580.
- George, E.M., Coch, D., 2011. Music training and working memory: An ERP study. *Neuropsychologia* 49 (5), 1083–1094.
- Giard, M.H., Peronnet, F., 1999. Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J. Cognit. Neurosci.* 11 (5), 473–490.
- Halgren, E., Ksenija, M., Chauvel, P., 1998. Generators of the late cognitive potentials in auditory and visual oddball tasks. *Electroencephalogr. Clin. Neurophysiol.* 106 (2), 156–164.
- Herdman, A.T., Fujioka, T., Chau, W., Ross, B., Pantev, C., Picton, T.W., 2006. Cortical oscillations related to processing congruent and incongruent grapheme-phoneme pairs. *Neurosci. Lett.* 399 (1–2), 61–66.
- Horn, H., Syed, N., Lanfermann, H., Maurer, K., Dierks, T., 2003. Cerebral networks linked to the event-related potential P300. *Eur. Arch. Psychiatry Clin. Neurosci.* 253 (3), 154–159.
- Horvath, J.C., Schilberg, L., Thomson, J., 2013. Does sight predominate sound? Electrophysiological evidence for multisensory mismatch negativity correlation. *Neurophysiology* 45 (5–6), 459–467.
- Kiehl, K., Laurens, K., Duty, T., Forster, B., Liddle, P., 2001. An event-related fMRI study of visual and auditory oddball tasks. *J. Psychophysiol.* 15 (4), 221–240.
- Lee, H., Noppeney, U., 2014. Music expertise shapes audiovisual temporal integration windows for speech, sinewave speech, and music. *Front. Psychol.* 5, 1–9.
- Liberman, I.Y., Shankweiler, D., Fischer, F.W., Carter, B., 1974. Explicit syllable and phoneme segmentation in the young child. *J. Exp. Child Psychol.* 18 (2), 201–212.
- Linden, D.E., Prvulovic, D., Formisano, E., Völlinger, M., Zanella, F.E., Goebel, R., Dierks, T., 1999. The functional neuroanatomy of target detection: an fMRI study of visual and auditory oddball tasks. *Cereb. Cortex* 9 (8), 815–823.
- Mann, V.A., Liberman, I., 1984. Phonological awareness and verbal short-term memory. *J. Learn. Disabil.* 17 (10), 592–599.
- Markley, C., Luck, S.J., Lopez-Calderon, J., 2011. ERPLAB Toolbox 2.0 User's Manual. University of California, Davis, Sacramento, CA.
- McCarthy, G., Luby, M., Gore, J., Goldman-Rakic, P., 1997. Infrequent events transiently activate human prefrontal and parietal cortex as measured by functional MRI. *J. Neurophysiol.* 77 (3), 1630–1634.
- Menon, V., Ford, J.M., Lim, K.O., Glover, G.H., Pfefferbaum, A., 1997. Combined event-related fMRI and EEG evidence for temporal-parietal cortex activation during target detection. *Neuroreport* 8 (14), 3029–3037.
- Minati, L., Rosazza, C., D'Incerti, L., Pietrocini, E., Valentini, L., Scaiola, V., Bruzzone, M. G., 2009. Functional MRI/event-related potential study of sensory consonance and dissonance in musicians and nonmusicians. *Neuroreport* 20 (1), 87–92.
- Mittag, M., Takegata, R., Kujala, T., 2011. The effects of visual material and temporal synchrony on the processing of letters and speech sounds. *Exp. Brain Res.* 211, 287–298.
- Molholm, S., Ritter, W., Murray, M.M., Javitt, D.C., Schroeder, C.E., Foxe, J.J., 2002. Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Cognit. Brain Res.* 14 (1), 115–128.
- Mulert, C., Pogarell, O., Juckel, G., Rujescu, D., Giegling, I., Rupp, D., Hegerl, U., 2004. The neural basis of the P300 potential. *Eur. Arch. Psychiatry Clin. Neurosci.* 254 (3), 190–198.
- Müllensiefen, D., Gingras, B., Musil, J., Stewart, L., 2014. The musicality of non-musicians: an index for assessing musical sophistication in the general population. *PLoS One* 9 (2), 1–23.
- Näätänen, R., Kreigipuu, K., 2012. The mismatch negativity (MMN). In: Luck, S.J., Kappenman, E.S. (Eds.), *The Oxford Handbook of Event-Related Potential Components*. Oxford University Press, New York, pp. 143–157.
- Näätänen, R., Gaillard, A.W., Mäntysalo, S., 1978. Early selective-attention effect on evoked potential reinterpreted. *Acta Psychol.* 42 (4), 313–329.
- Näätänen, R., Paavilainen, P., Rinne, T., Alho, K., 2007. The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol.* 118 (12), 2544–2590.
- Paraskevopoulos, E., Kuchenbuch, A., Herholz, S.C., Pantev, C., 2012. Musical expertise induces audiovisual integration of abstract congruency rules. *J. Neurosci.* 32 (50), 18196–18203.
- Paraskevopoulos, E., Kuchenbuch, A., Herholz, S.C., Pantev, C., 2013. Multisensory integration during short-term music reading training enhances both uni- and multisensory cortical processing. *J. Cognit. Neurosci.* 26 (10), 2224–2238.
- Pazo-Alvarez, P., Cadaveira, F., Amenedo, E., 2003. MMN in the visual modality: a review. *Biol. Psychol.* 63 (3), 199–236.
- Petrini, K., McAleer, P., Pollick, F., 2010. Audiovisual integration of emotional signals from music improvisation does not depend on temporal correspondence. *Brain Res.* 1323, 139–148.
- Proverbio, A.M., Calbi, M., Manfredi, M., Zani, A., 2014. Audio-visuomotor processing in the musicians' brain: an ERP study on professional clarinetists and violinists. *Sci. Rep.* 4, 5866–5876.
- R Core Team, 2015. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria <http://www.R-project.org/>.
- Raij, T., Uutela, K., Hari, R., 2000. Audiovisual integration of letters in the human brain. *Neuron* 28 (2), 617–625.
- Sakitt, B., 1976. Iconic memory. *Psychol. Rev.* 83 (4), 257–276.
- Schneider, W., Eschman, A., Zuccolotto, A., 2002. *E-Prime User's Guide*. Psychology Software Tools, Inc., Pittsburgh, PA.
- Sperling, G., 1960. The information available in brief visual presentation. *Psychol. Monogr.* 74 (11), 1–29.

- Stevens, A.A., Skudlarski, P., Gatenby, J.C., Gore, J.C., 2000. Event-related fMRI of auditory and visual oddball tasks. *Magn. Reson. Imaging* 18 (5), 495–502.
- Tervaniemi, M., Medvedev, S.V., Alho, K., Pakhomov, S.V., Roudas, M.S., Van Zuijen, T.L., Näätänen, R., 2000. Lateralized automatic auditory processing of phonetic versus musical information: a PET study. *Hum. Brain Mapp.* 10 (2), 74–79.
- Tibon, R., Levy, D.A., 2015. Striking a balance: analyzing unbalanced event-related potential data. *Front. Psychol.* 6, 1–4.
- Tremblay, A., Ransijn, J., 2015. *LMERConvenienceFunctions: Model Selection and Post-hoc Analysis for (G)LMER Models*. R package version 2.10.
- Van Atteveldt, N., Formisano, E., Goebel, R., Blomert, L., 2004. Integration of letters and speech sounds in the human brain. *Neuron* 43 (2), 271–282.
- Vatakis, A., Spence, C., 2006. Audiovisual synchrony perception for music, speech, and object actions. *Brain Res.* 1111 (1), 134–142.
- Vuust, P., Brattico, E., Seppänen, M., Näätänen, R., Tervaniemi, M., 2012. The sound of music: differentiating musicians using a fast, musical multi-feature mismatch negativity paradigm. *Neuropsychologia* 50 (7), 1432–1443.
- Žarić, G., Fraga González, G., Tijms, J., van der Molen, M.W., Blomert, L., Bonte, M., 2014. Reduced neural integration of letters and speech sounds in dyslexic children scales with individual differences in reading fluency. *PLoS One* 9 (10), 1–14.