

Effects of Music Reading on Motor Cortex Excitability in Pianists: A Transcranial Magnetic Stimulation Study Fabio Giovannelli, a,b Simone Rossi, c Alessandra Borgheresi, b Gioele Gavazzi, d Gaetano Zaccara, b,e Maria Pia Viggiano a and Massimo Cincotta b * aSection of Psychology – Department of Neuroscience, Psychology, Drug Research and Child's Health (NEUROFARBA), University of Florence, Firenze, Italy b Unit of Neurology of Florence, Central Tuscany Local Health Authority, Firenze, Italy c Department of Medicine, Surgery and Neuroscience, Unit of Neurology and Clinical Neurophysiology, Siena Brain Investigation and Neuromodulation Lab (Si-BIN Lab), University of Siena, Italy d IRCCS SDN, Napoli, Italy e Regional Health Agency of Tuscany, Firenze, Italy Abstract—

Neurophysiological studies suggest that music reading facilitates sensorimotor cortex. The aim of this study was to evaluate (1) whether in pianists, reading notes in bass and treble clef selectively enhances right and left primary motor cortex (M1) excitability; and (2) whether reading notes played with the thumb or little finger selectively modulates the excitability of specific muscles. Twenty musicians (11 pianists, 9 non-pianists) participated. Transcranial magnetic stimulation (TMS) was applied while subjects read the bass or the treble clef of sheets music and during the observation of a blank staff (baseline). When pianists read the treble clef, the excitability of the left M1 was higher compared to that recorded in the right M1. Moreover, in the treble clef condition motor evoked potentials (MEPs) induced by TMS of the left M1 were higher when pianists read notes to be played with the 5 finger (little finger) with respect to 1 finger (thumb) notes, whereas in the bass clef condition TMS of the right M1 induced higher MEPs for 1 finger note compared to 5 finger notes. No significant modulation was observed in non-pianists. These data support the view that music reading may induce specific inter- and intra-hemispheric modulation of the motor cortex excitability. 2020 Published by Elsevier Ltd on behalf of IBRO. Key words: motor cortex excitability, motor evoked potential, music reading, musicians, transcranial magnetic stimulation.

INTRODUCTION Music sight reading is an essential skill for musicians that involves the ability to decode symbols contained in a music score (i.e. musical notes and other notations reproduced on a staff) and to transform the visual information into coordinated movements [or motor synergies, (see Rossi et al., 2019)] consonant with the instrument to be played, and finally into a musical performance (Sloboda, 1984). Music sight reading relies on highly complex cognitive processes (e.g. pattern recognition, spatial information analysis, motor planning and sensorimotor integration, working memory, and cognitive flexibility). Although some similarities with words and numbers reading can be detected, music reading represents a distinct ability (Sloboda, 1984), that differs from written language both at a functional and anatomical level (Peretz and Zatorre, 2005). Brain lesions (Brust, 1980; Cappelletti et al., 2000; Schoen et al., 2001) and functional neuroimaging studies (Sergent et al., 1992; Stewart et al., 2003; Meister et al., 2004; Schoen et al., 2012) revealed that the neural processes underlying music reading rely on a distributed neural network mainly, but not exclusively, lateralized to the dominant hemisphere. A selective impairment in music reading (music alexia) with spared ability in playing music has been described in aphasic patients (Brust, 1980). However, a dissociation between a deficient ability to read, write and understand musical notation and preserved ability to read aloud letters, words, and numbers has also been reported (Cappelletti et al., 2000). Moreover, a study conducted in seven musicians undergoing pre-surgical brain mapping demonstrated that cortical areas involved in score reading may be partially <https://doi.org/10.1016/j.neuroscience.2020.04.022> 0306-4522/ 2020 Published by Elsevier Ltd on behalf of IBRO. *Corresponding author. Address: SOC di Neurologia Firenze, Azienda USL Toscana Centro, Ospedale San Giovanni di Dio, Via di Torregalli, 350143 Firenze, Italy. Tel: +39-055-6932055; Fax: +39-055-6932432. E-mail address: massimo.cincotta@uslcentro.toscana.it (M. Cincotta). Abbreviations: ADM, abductor digitorum minimi; APB, abductor pollicis brevis; EMG, electromyographic; M1,

primary motor cortex; MEPs, motor evoked potentials; RMT, resting motor threshold; TMS, transcranial magnetic stimulation. NEUROSCIENCE RESEARCH ARTICLE F. Giovannelli et al. / Neuroscience 437 (2020) 45–53

45 distinct from other language functions (Roux et al., 2007). Namely, interferences during score reading (reading arrest, hesitation, and paraphasia) was induced when direct cortical stimulation was applied either to the same sites producing impairment in other language tasks (i.e. word and number reading) and to sites specific to the score-reading task. The first study exploring the neural substrate of music reading was performed by Sergent et al. (1992) using positron emission tomography. The authors observed an activation of the left occipitoparietal junction when musicians read a musical score without listening or playing. Moreover, when musicians simultaneously read the musical score and listened to its performance played on a piano, activations emerged in the superior and posterior part of the left supramarginal gyrus and in the inferior parietal lobule bilaterally (Sergent et al., 1992). In a recent fMRI study, Schön et al. (2012) emphasized the contribution of the right hemisphere in music reading reporting the activation of the right occipitotemporal junction, the superior parietal lobule, and the intraparietal sulcus during score compared to verbal and numerical notation reading. A substantial body of studies using transcranial magnetic stimulation (TMS) of the primary motor cortex (M1) revealed how the excitability of the corticomotoneuronal system is selectively modulated by several cognitive tasks. For instance, the size of TMS-induced motor evoked potential (MEP) is enhanced by movement observation (Fadiga et al., 1995; Strafella and Paus, 2000) and motor imagery (Izumi et al., 1995; Rossi et al., 1998; Fadiga et al., 1999; Rossini et al., 1999), although these tasks do not imply overt movements. Similarly, it has been widely reported that linguistic tasks may influence the motor system irrespective of the engagement of the tested muscles in task performance (Tokimura et al., 1996; Meister et al., 2003; Papathanasiou et al., 2004; Sparing et al., 2007; Bracco et al., 2009; Suppa et al., 2015). Namely, in healthy subjects reading aloud facilitated the corticospinal excitability as revealed by the increase in the MEP size recorded in the muscles of the dominant hand (Tokimura et al., 1996; Meister et al., 2003; Bracco et al., 2009; Suppa et al., 2015). Lack of spinal excitability changes suggested that the linguistic task-induced MEP facilitation occurred at the cortical level (Tokimura et al., 1996). This effect was specific for the dominant hand area of the M1 and absent in the non-dominant hand (Tokimura et al., 1996; Meister et al., 2003; Bracco et al., 2009; Suppa et al., 2015) or in leg muscles (Meister et al., 2003). This bunch of evidence supports the hypothesis of a functional link between the brain areas mediating language processes and the motor system (Willems and Hagoort, 2007). Recent electroencephalographic data suggest that reading sheet music facilitates sensorimotor cortex in musicians (Behmer and Jantzen, 2011). Moreover, it has been argued that reading the notes of a musical score activates specific expectations about the type of sound to be produced and implies an automatic sensorimotor translation (Stewart, 2005; Schön and Besson, 2003, 2005). Given the strict relationship between music reading and action, it is conceivable to hypothesize that similarly to word and text reading also music reading may modulate motor cortex excitability as tested by TMS. The ‘meaning’ of the notes written on a staff depends on a reference point by which the name and pitch of each note may be determined according to their position. This reference point (i.e. the clef) is indicated by a symbol given at the beginning of the staff. In music, there are seven clefs (setticlavio) but the treble and bass clefs are by far the most common (D’Anselmo et al., 2018). Music for instruments with a higher pitch (e.g. violin or flute) are typically written in the treble clef whereas the bass clef notates for the lower register. As the piano has a wide pitch range, pianists read either the treble or the bass clef, with the two hands playing simultaneously different notes and chords: usually the left hand ‘reads’ the bass clef, while the right hand ‘reads’ the treble clef. In contrast, for most of other instruments (e.g. brass and strings) both hands simultaneously contribute in

playing the same sound. This difference between pianists and nonpianists offers the possibility to investigate possible hemispheric differences during music reading. Another basic skill required to play a musical instrument, particularly the piano, is the fingering, that is ‘choosing’ which fingers are more appropriate to play certain sequences of notes in a musical piece. Although individual differences in piano fingerings exist, mainly due to ergonomic and music-interpretive factors (Parncutt et al., 1997), most of finger sequences are defined by rules and physical constraints. The objective of the present study was twofold. First, to evaluate whether in piano players, compared to nonpianists musicians, reading notes in the bass clef (usually played with the left hand) and in the treble clef (usually played with the right hand) selectively enhances right and left M1 excitability (we defined this as ‘interhemispheric effect’). Second, to evaluate whether in pianists reading notes played with the thumb (1 finger) or with the little finger (5 finger) selectively modulates the excitability of the abductor pollicis brevis (APB) and abductor digitorum minimi (ADM) muscles (we defined this as ‘intra-hemispheric effect’). To this end, TMS has been used to record MEPs in the left and right hand from two groups of musicians (pianists and non-pianists) during reading of different music sheets without performing movements.

EXPERIMENTAL PROCEDURES

Participants

Twenty healthy musicians (11 pianists and 9 non-pianists; thirteen women; mean age 26.6 years, range 21–39 years) with no history of implanted metal devices, neurological or psychiatric diseases, and use of central nervous system drugs were included in the study. All subjects were naïve for TMS techniques and blinded to the main purpose of the study. All participants were right-handed (mean dexterity index 91.7%, range 77.8–100%) except one (index 78.9) according to the Edinburgh handedness inventory (Oldfield, 1971), and all had normal or corrected-to-normal vision. 46 F. Giovannelli et al. / Neuroscience 437 (2020) 45–53

The study was performed according to the Declaration of Helsinki and was approved by the local ethics committee. TMS parameters were in accordance with published international safety recommendations (Rossi et al., 2009). Before each experimental session all participants gave a written informed consent to participate in the study, were screened as suitable for TMS (Rossi et al., 2011), and were asked to report adverse effects experienced during or after TMS. All musicians (professional and non-professional) reported at least 4 years of formal musical training (pianists: 8.8 ± 2.9 years, range 5–15 years; nonpianists: 7.6 ± 3.3 years, range 4–15 years). The group of non-pianist musicians consisted of two violinists, two guitarists, one flautist, two clarinet players, one singer, and one percussionist. Although most of non-pianists had a basic knowledge of piano, they were specialized in another instrument and were able to fluently read both the treble clef and the bass clef.

Stimuli

Stimuli consisted of didactic piano arrangements from compositions by Johann Sebastian Bach (Allegro from Brandenburg concerto n.2 BWV 1047; Allegro from concerto BWV 1060; Arioso from cantata BWV 156; Corale from concert BWV 645; and Minuetto from suite n.2 BWV 1067 for orchestra). As a baseline condition a blank staff (without notes) was used (Fig. 1). Sheets music were created by the open source music notation software MuseScore BVBA and converted as images (Fig. 1). The software E-prime 2.0 (Psychology Software Tools, Inc., Sharpsburg, PA, USA) was used for stimuli presentation. Each sheet music consisted of two staves (Fig. 1): the first for the treble clef, the other for the bass clef. Reading was guided by a red bar moving on the score according to the note values with a tempo of 80 beats per minute (bpm). This tempo was a compromise between a reasonable and comfortable reading speed and the appropriate inter-pulse interval for TMS. In each sheet music, notes usually played with the first (thumb) and with the fifth finger (little finger) were identified. In order to verify whether the selected fingering was correct, expert musicians were consulted. Electromyographic (EMG) recordings and TMS Surface EMG was recorded simultaneously from left and right APB and ADM muscles, amplified by a Digitimer D360 (Digitimer Ltd, Welwyn Garden City, UK),

analog filtered (20–2000 Hz), and digitized (A/D rate 5000 Hz) by a micro 1401 unit and signal 2 software (Cambridge electronic design, UK). Focal TMS of the M1 was delivered using a Magstim 200 stimulator (Magstim Co., UK) connected to an 8-shaped coil. The coil was initially placed at the optimal position (hot spot) to elicit MEP in contralateral APB muscle, with the handle pointing backwards and 45° away from the midline. Then the coil was slightly moved until the site that guaranteed stable simultaneous responses from both muscles was localized. Once the coil position was established, the resting motor threshold (RMT) was measured in the left and right APB/ADM, in accordance with international recommendations (Rossini et al., 2015). Briefly, RMT was defined as the minimum stimulus intensity that produced MEP >50 µV in at least five of 10 consecutive trials during muscle relaxation. Then, the intensity of the magnetic pulse was adjusted to elicit MEP of about 0.5–1.5 mV in peak-to-peak amplitude in at least one of the contralateral relaxed muscle (APB and ADM). This stimulus intensity (ranging from 110% to 130% individual RMT) was used in the experimental session. Subjects were instructed to avoid limb muscle contraction throughout the recordings. Relaxation of the APB and ADM muscles was monitored online from high gain EMG. Offline, trials with background EMG activity (defined as involuntary EMG activity greater than 50 µV in a time window of 20 ms preceding the magnetic pulse) were discarded. Moreover, for all the experimental conditions, the root mean squares (RMS) of the pre-stimulus EMG activity were calculated off-line to confirm full relaxation. Group RMS values (mean ± SD) ranged from 0.006 ± 0.008 mV to 0.016 ± 0.007 mV throughout experiments. These low values did not show any significant difference and were within Fig. 1.

Example of music sheet used for the three experimental conditions: blank staff as a baseline condition; reading the treble clef; reading the bass clef. Reading was guided by a red bar (indicated by the arrow) moving on the score according to the note values. F. Giovannelli et al. / *Neuroscience* 437 (2020) 45–53 47 the range of previously reported data in resting muscles (Müller et al., 2009; Giovannelli et al., 2009, 2013). Experimental paradigm The experiment took place in a quiet room. Subjects were seated in a comfortable chair with their arms fully supported. They were instructed to remain fully relaxed and to read the sheets music silently without oral movements and vocalizations (the treble and the bass clef alternatively), guided by the red bar, without performing any movements. As reported in the specific scientific literature, musicians may spontaneously generate musical imagery (defined ‘notational audiation’) during silent reading of music scores (Brodsky et al., 2003, 2008). However, in the present study no specific indication was given about imagery in order to keep the reading task as ‘ecological’ as possible. Single pulse TMS was applied to either M1 while subjects alternatively read the treble or the bass clef of the sheets music. Reading the treble and the bass clef were performed as two different experimental conditions. TMS pulses were synchronized to the instant in which the red bar passed over the predefined notes selected according to the fingering (see Stimuli section). The interval between two consecutive magnetic pulses was 5 s at least. For each sheet music, 8–12 artefactfree traces (depending on the length of the score to be read) were recorded from the contralateral APB and ADM muscles. For each of the two different experimental conditions (treble and bass clef), a total of 30 MEPs were recorded (15 representing notes to be played with the 1 and 15 representing notes to be played with the 5 finger). The order of experimental conditions (treble and bass clef) and hemisphere was randomized and counterbalanced across subjects, except for the baseline condition (blank staff) that was performed twice: before and after the experimental session, in order to test stability of M1 excitability at the baseline. As paired Student’s t-test revealed no difference between the measures obtained in the two baseline blocks (TMS of the right M1: $P = 0.375$ and $P = 0.292$, for the APB and ADM muscles, respectively; TMS of the left M1: $P = 0.297$, and $P = 0.447$, for the APB and ADM muscles), data were pooled together for further analysis. At the end of the experiment, pianists were

asked to report the fingering they would use for the specific notes triggering the TMS pulses, by filling a paper score that reproduced all sheets music presented during the experiment. Cohen's kappa coefficient was used to measure the agreement between the selection of the notes and the individual fingering. The maximal duration of the overall session was 90 min, of which about 50 min served for the TMS experimental paradigm. Data analysis Separately for each of the two different experimental conditions (treble and bass clef reading) and individual, the peak-to-peak amplitude of each single MEP was measured and the mean amplitude was calculated. Moreover, the amplitude of each MEP was normalized to baseline (mean MEP amplitude recorded during the observation of the blank staff before and after the experimental session pooled together) and expressed as z-scores using the following formula: $z = (A - B) / SDB$, where A and B are the individual amplitudes in the reading and baseline conditions, respectively, while SDB is the standard deviation of the baseline amplitude (Papathanasiou et al., 2004; Bracco et al., 2009; Giovannelli et al., 2013). A first analysis has been performed to evaluate whether music reading per-se may facilitate the corticospinal excitability in musicians, regardless the clef and the muscles. To this end, the average MEP amplitude obtained during the treble and the bass clef reading in the APB and ADM muscles was calculated for either the TMS of the left and right M1 conditions. The values obtained were compared with that measured in the baseline condition (blank staff). Namely, mean MEP amplitudes were analyzed using a mixed design analysis of variance (ANOVA) with GROUP (two levels: pianists and non-pianists) as between subjects factor, and CONDITION (two levels: baseline and sheet music reading) and HEMISPHERE (two levels: TMS of the left and right M1) as within-subject factors. For the 'inter-hemispheric effect', the z-scores calculated for APB and ADM muscles were averaged in order to increase the statistical power and the resulting values were analyzed by a mixed ANOVA design with GROUP (two levels: pianists and non-pianists) as between subjects factor, and CLEF (two levels: treble and bass) and HEMISPHERE (two levels: TMS of the left and right M1) as within-subject factors. The 'intra-hemispheric effect' has been evaluated only in the group of pianists. Namely, the z-scores values, obtained from all MEPs regardless any mismatch between the fingering reported by each pianist at the end of the experiment and the selection of the notes, were analyzed by a three-way repeatedmeasure ANOVA with CLEF (two levels: treble and bass), FINGERING (two levels: notes to be played with the 1 vs notes to be played with the 5 finger), and MUSCLE (two levels: APB and ADM) as within-subject factors. This analysis was repeated after removing trials in which pianists reported a different preferred fingering, in order to evaluate whether mismatch between the fingering reported by each pianist at the end of the experiment and the selection of the notes used in the present study could have influenced the results. To correct violations of the sphericity assumption, Greenhouse-Geisser corrections were applied when necessary. All tests were two-tailed and significance was set at $P < 0.05$ and adjusted by Bonferroni correction for multiple comparisons (Bonferroni corrected alpha levels are reported in the Results). In addition, partial eta squared (η^2_p) has been calculated as effect size. Analyses were performed using the software IBM SPSS 20.0 (SPSS, Chicago, IL, USA).

48 F. Giovannelli et al. / Neuroscience 437 (2020) 45–53

RESULTS None of the participants reported adverse effects during or after the experimental procedures. The RMT (mean \pm standard deviation) measured in the left and right M1 was 41.7 ± 9.4 and 40.7 ± 8.9 for pianists and 39.7 ± 5.6 , 39.2 ± 7.0 for non-pianists, without significant differences. The mixed ANOVA conducted to evaluate whether music reading per-se may facilitate the corticospinal excitability showed that the main effect of CONDITION ($F_{1,18} = 19.420$; $P < 0.001$, $\eta^2_p = 0.519$) was significant whereas the main effects of HEMISPHERE ($F_{1,18} = 0.046$; $P = 0.832$, $\eta^2_p = 0.003$) and the interaction between these two factors were not significant ($F_{1,18} = 0.395$; $P = 0.538$, $\eta^2_p = 0.021$). Also the main effect of GROUP ($F_{1,18} = 0.388$; $P = 0.541$, $\eta^2_p = 0.021$) and the interactions

CONDITION GROUP ($F_{1,18} = 0.704$; $P = 0.412$, $gp\ 2 = 0.038$), HEMISPHERE GROUP ($F_{1,18} = 0.168$; $P = 0.687$, $gp\ 2 = 0.009$), and CONDITION HEMISPHERE GROUP ($F_{1,18} = 2.087$; $P = 0.166$, $gp\ 2 = 0.104$), were not significant. These comparisons showed that in either pianists or non-pianists the mean MEP amplitude was higher during reading the sheets music with respect to the observation of a blank staff. Inter-hemispheric effect (pianists vs non-pianists) The mixed ANOVA revealed that the main effects of the GROUP ($F_{1,18} = 0.120$; $P = 0.733$, $gp\ 2 = 0.007$) and HEMISPHERE ($F_{1,18} = 1.624$; $P = 0.219$, $gp\ 2 = 0.083$) were not significant whereas a significant effect emerged for CLEF ($F_{1,18} = 5.312$; $P = 0.033$, $gp\ 2 = 0.228$): the amplitude of the MEPs was significantly higher during treble clef reading compared to the bass clef condition, regardless of the stimulated hemisphere and the group of musicians. Significant interactions emerged for HEMISPHERE GROUP ($F_{1,18} = 4.392$; $P = 0.051$, $gp\ 2 = 0.196$) and CLEF HEMISPHERE ($F_{1,18} = 5.598$; $P = 0.029$, $gp\ 2 = 0.237$) whereas the interaction CLEF GROUP ($F_{1,18} = 0.057$; $P = 0.814$, $gp\ 2 = 0.003$) was not significant. Finally, the interaction between CLEF, HEMISPHERE, and GROUP ($F_{1,18} = 8.868$; $P = 0.008$, $gp\ 2 = 0.330$) was significant. Post-hoc comparisons (Bonferroni corrected) revealed that when pianists read the treble clef, the excitability of the left M1 was significantly higher compared to that recorded in the right M1 ($P < 0.001$) (Fig. 2A). In contrast, no significant differences emerged during reading of the bass clef ($P = 0.923$). Moreover, no significant MEP modulation was observed in non-pianists either for the treble and bass clef condition ($P = 0.497$ and $P = 0.818$, respectively) (Fig. 2B). The individual MEPs changes expressed as z-scores obtained in pianists and non pianists in the two different experimental conditions (treble and bass clef) are shown in Fig. 3. Intra-hemispheric effect (pianists) The three-way ANOVA showed that the main effects of FINGERING ($F_{1,10} = 0.084$; $P = 0.778$, $gp\ 2 = 0.008$) and MUSCLE ($F_{1,10} = 0.120$; $P = 0.737$, $gp\ 2 = 0.012$) were not significant whereas a significant effect emerged for CLEF ($F_{1,10} = 6.331$; $P = 0.031$, $gp\ 2 = 0.388$): again the MEP size was significantly higher for treble clef compared to the bass clef condition. The interaction between CLEF and FINGERING ($F_{1,10} = 15.723$; $P = 0.003$, $gp\ 2 = 0.611$) was significant. Post-hoc comparisons (Bonferroni corrected) showed that in the treble clef condition, MEPs induced by TMS of the left M1 were significantly higher when pianists read notes to be played with the 5 finger (little finger) with respect to note to be played with the 1 finger (thumb) of the right hand, regardless of the muscle (Fig. 4A). In contrast, in the bass clef condition TMS of Fig. 2. Inter-hemispheric effect: mean (\pm SE) values of MEP size obtained in the left and right hands (APB and ADM muscles were averaged) by TMS of the right and left M1 hand in the treble and bass clef conditions in (A) pianists and (B) non-pianists. Values are expressed as z-scores relative to the baseline (see Methods). Asterisk indicates a significant difference ($P < 0.05$, Bonferroni post-hoc comparison test). F. Giovannelli et al. / Neuroscience 437 (2020) 45–53 49 the right M1 induced higher MEPs for note to be played with the 1 finger compared to note to be played with the 5 finger of the left hand (Fig. 4B). The interaction CLEF MUSCLE ($F_{1,10} = 0.056$; $P = 0.817$, $gp\ 2 = 0.006$), FINGER MUSCLE ($F_{1,10} = 0.146$; $P = 0.710$, $gp\ 2 = 0.014$), and CLEF X HEMISPHERE GROUP ($F_{1,10} = 0.526$; $P = 0.485$, $gp\ 2 = 0.050$) were not significant. Finally, a good mean agreement emerged between the fingering reported by each pianist at the end of the experiment and the selection of the notes as revealed by the Cohen's kappa coefficient (treble clef: mean 0.731, range 0.415–0.878; bass clef: mean 0.704, range 0.429–0.875; total: mean 0.716, range 0.422–0.875). Results of the analysis of the intra-hemispheric effect were confirmed after removing trials in which pianists reported a different fingering. Namely, the three-way ANOVA showed that the main effects of CLEF ($F_{1,10} = 4.821$; $P = 0.053$, $gp\ 2 = 0.325$), FINGERING ($F_{1,10} = 0.125$; $P = 0.731$, $gp\ 2 = 0.012$), and MUSCLE ($F_{1,10} = 0.095$; $P = 0.764$, $gp\ 2 = 0.009$) were not significant. The interaction between CLEF and FINGERING ($F_{1,10} = 10.609$; $P = 0.009$, $gp\ 2 = 0.515$) was significant.

The interaction CLEF MUSCLE ($F_{1,10} = 0.073$; $P = 0.793$, $gp\ 2 = 0.007$), FINGER MUSCLE ($F_{1,10} = 0.325$; $P = 0.581$, $gp\ 2 = 0.031$), and CLEF HEMISPHERE GROUP ($F_{1,10} = 1.101$; $P = 0.317$, $gp\ 2 = 0.100$) were not significant. **DISCUSSION** Current findings show that music reading facilitates the corticospinal excitability in musicians, with a pattern of modulation in the two hemispheres that differs between pianists and non-pianists. Interhemispheric differences in piano players emerged during reading notes in the treble clef usually played with the right hand, as the excitability of the left M1 was significantly higher compared to that recorded in the right M1, whereas no significant differences emerged during the reading in the bass clef. In contrast, in nonpianists the effect of music reading was non-specific as this inter-hemispheric difference in MEP modulation was not observed. It should be acknowledged that the relatively small sample size in the current study did not allow to draw definite conclusions and larger studies are needed to confirm these results. However, data at individual level are consistent with the effects size measure (estimated by the partial eta squared) showing that the highest proportion of variance was the one attributable to inter-hemispheric differences in the corticospinal excitability changes observed in pianists while reading the treble clef. Overall, present findings are consistent with recent neurophysiological data revealing that reading sheet music facilitates the motor system (Behmer and Jantzen, 2011). That is, an activation of the sensorimotor cortex, as revealed by event related desynchronization of the mu-rhythm, was observed in violin and trumpet players during musical notation reading compared to nonmusicians. Authors suggested that the activation of the motor system during music reading may mediate and foster the association between visual abstract symbols (musical notes) and specific motor programs (Behmer and Jantzen, 2011). Reading the notes of a sheet music activates expectations about the type of sound to be produced (Stewart, 2005; Schön and Besson, 2003, 2005) and, according to our results, is linked to a specific action Fig. 3. Inter-hemispheric effect: Individual values of motor evoked potentials (MEPs) size expressed as z-score (see Methods) obtained in pianists and non pianists in the two different experimental conditions (treble and bass clef). 50 F. Giovannelli et al. / *Neuroscience* 437 (2020) 45–53 that depends on the instrument used to produce the sound. These data expand previous neuroimaging and neurophysiological studies showing that the activity of the hand area of the motor cortex in piano players is increased during listening of musical pieces (Haueisen and Knösche, 2001; Bangert et al., 2006; D’Ausilio et al., 2006), during the observation of fingering errors (Candidi et al., 2014), or during motor imagery of chords (Rossi et al., 2019). Moreover, we demonstrated that similarly to word reading (Tokimura et al., 1996; Meister et al., 2003; Papathanasiou et al., 2004; Sparing et al., 2007; Bracco et al., 2009; Suppa et al., 2015) also music reading modulates the excitability of the motor system at rest. A complex pattern of hemispheric asymmetries during music reading was observed in a behavioral study (D’Anselmo et al., 2015), in which a group of pianists was required to read single notes or chords presented in the left or right visual field and to play them on a keyboard with the left, right, or both hands. The hemispheric asymmetries, as revealed by differences in response times and accuracy, emerged depending on the motor output, i.e. whether the responses were given with two hands or one hand. Namely, higher accuracy and lower response times emerged when pianists played with one hand (left or right) the notes presented in the right visual field, suggesting a left hemisphere advantage. Conversely, a right hemisphere advantage was observed when both hands were used to perform the task (D’Anselmo et al., 2015). Interestingly, also the clef of the musical notations was a crucial factor influencing the performance. Indeed, the typical Simon effect (i.e. faster and more accurate responses when stimulus and motor output are on the same side) was observed only for responses given with the right hand whereas no facilitation was induced for notes and chords played with the left hand with stimuli presented in the left visual field. Authors argued that reading in the bass clef may be a function lateralized in the

left hemisphere due to an ‘inability’ of the right hemisphere in reading the bass clef (D’Anselmo et al., 2015). This detriment effect appears paradoxical, as pianists typically play the bass clef with the left hand. However, it should be noted that, while reading and performing only in the treble clef is not unusual, reading the bass clef alone and playing only with the left hand is an atypical condition, except for particular musical pieces (e.g. the piano concerto for the left hand by Maurice Ravel). However, pianists usually play using both hands and read both clefs simultaneously. Moreover, it has been suggested that the bass clef is decoded having as a reference the treble clef by a visuo-spatial processing (Sloboda, 1984, D’Anselmo et al., 2018). The lack of inter-hemispheric differences in corticospinal excitability in pianists emerged in the present study, is in keeping with behavioral data suggesting a peculiarity of the bass clef reading (D’Anselmo et al., 2015). Differences in the treble and bass clef processing is also supported by neuropsychological evidence. Namely, a dissociation between a deficit in naming notes in the bass clef with a spared ability to read the treble clef was described in a professional musician with amusia following a left temporoparietal ischemia (Schoen et al., 2001). Regarding the evaluation of the ‘intra-hemispheric effect’ in pianists, that is to test the effect of fingering, we expected to record higher MEPs in the APB muscle when subjects read note to be played with the 1 finger (thumb), and higher MEPs in the ADM for notes to be played with the 5 finger (little finger). Contrary to our hypothesis, MEPs induced by TMS of the left M1 were significantly higher when pianists read notes to be played with the 5 finger in the treble clef condition whereas TMS of the right M1 induced higher MEPs for Fig. 4. Intra-hemispheric effect (group of pianists): mean (\pm SE) values of MEP size obtained in the APB and ADM muscles when pianists read notes to be played with the thumb (1 finger) or with the little finger (5 finger) for (A) treble clef (right hand) and (B) bass clef (left hand). Values are expressed as z-scores relative to the baseline (see Methods). Asterisk indicates a significant difference ($P < 0.05$, Bonferroni post-hoc comparison test). F. Giovannelli et al. / Neuroscience 437 (2020) 45–53

51 note to be played with the 1 finger in the bass clef condition, regardless the muscle. At present, we can only suggest a conjecture to interpret these data. Indeed, it should be noted that the thumb for the left hand and the little finger for the right hand usually are related to hand position characterized by higher extension, for example in playing chords or notes at the end of ascendant scales. In a recent TMS study, Rossi et al. (2019) demonstrated that when professional pianists were engaged in mental imagery of triad chords, corticospinal output progressively increased along with task demands in terms of mental representation of hand extension. Namely, the corticospinal excitability evaluated during motor imagery of chords was greater at the APB and ADM muscles (i.e. muscles always engaged in the mental execution of each triad chord) compared to the extensor communis digitorum and flexor digitorum superficialis. Moreover, the MEP amplitude was higher during imagery of the 9th interval chord (i.e. the most demanding in terms of hand extension possible) compared to chord with smaller intervals (Rossi et al., 2019). The hypothesis that a motor imagery process triggered by the music reading task may account for the effect observed in our study is unlikely due to the lack of muscle specificity in the MEP facilitation (Izumi et al., 1995; Rossi et al., 1998; Fadiga et al., 1999; Rossini et al., 1999). However, in consideration of the data by Rossi et al. (2019), our ‘intra-hemispheric effect’ findings could reflect the fact that the main determinant of M1 excitability changes during music reading may be how the action needed to produce the related sound is ‘challenging’ and not the selective activation of the cortical representation of single muscles. Further studies are needed to corroborate this hypothesis.

CONFLICT OF INTEREST STATEMENT None of the authors have potential conflicts of interest to be disclosed.

ACKNOWLEDGEMENT A grant from “Ente Cassa di Risparmio di Firenze”, Florence, Italy supported this project. We are grateful to Dr. Loretta De Nigris for her kind help in selecting the music scores used in this study.

REFERENCES

Bangert M, Peschel T, Schlaug G, Rotte M, Drescher D, Hinrichs H, et al. (2006) Shared networks for auditory and motor processing in professional pianists: evidence from fMRI conjunction. *Neuroimage* 30:917–926. Behmer Jr LP, Jantzen KJ (2011) Reading sheet music facilitates sensorimotor mu-desynchronization in musicians. *Clin Neurophysiol* 122:1342–1347. Bracco L, Giovannelli F, Bessi V, Borgheresi A, Di Tullio A, Sorbi S, Zaccara G, Cincotta M (2009) Mild cognitive impairment: loss of linguistic task-induced changes in motor cortex excitability. *Neurology* 72:928–934. Brodsky W, Henik A, Rubinstein BS, Zorman M (2003) Auditory imagery from musical notation in expert musicians. *Percept Psychophys* 65:602–612. Brodsky W, Kessler Y, Rubinstein BS, Ginsborg J, Henik A (2008) The mental representation of music notation: notational audiation. *J Exp Psychol Hum Percept Perform* 34:427–445. Brust JC (1980) Music and language: musical alexia and agraphia. *Brain* 103:367–392. Candidi M, Sacheli LM, Mega I, Aglioti SM (2014) Somatotopic mapping of piano fingering errors in sensorimotor experts: TMS studies in pianists and visually trained musically naïves. *Cereb Cortex* 24:435–443. Cappelletti M, Waley-cohen H, Butterworth B, Kopelman M (2000) A selective loss of the ability to read and to write music. *Neurocase* 6:321–332. D’Anselmo A, Giuliani F, Marzoli D, Tommasi L, Brancucci A (2015) Perceptual and motor laterality effects in pianists during music sight-reading. *Neuropsychologia* 71:119–125. D’Anselmo A, Giuliani F, Campopiano F, Carta E, Brancucci A (2018) Hemispheric asymmetries in setticlavio reading. *Neuropsychology* 32:337–343. D’Ausilio A, Altenmüller E, Olivetti Belardinelli M, Lotze M (2006) Cross-modal plasticity of the motor cortex while listening to a rehearsed musical piece. *Eur J Neurosci* 24:955–958. Fadiga L, Buccino G, Craighero L, Fogassi L, Gallese V, Pavesi G (1999) Corticospinal excitability is specifically modulated by motor imagery: a magnetic stimulation study. *Neuropsychologia* 37:147–158. Fadiga L, Fogassi L, Pavesi G, Rizzolatti G (1995) Motor facilitation during action observation: a magnetic stimulation study. *J Neurophysiol* 73:2608–2611. Giovannelli F, Borgheresi A, Balestrieri F, Zaccara G, Viggiano MP, Cincotta M, et al. (2009) Modulation of interhemispheric inhibition by volitional motor activity: an ipsilateral silent period study. *J Physiol.* 587:5393–5410. Giovannelli F, Banfi C, Borgheresi A, Fiori E, Innocenti I, Rossi S, et al. (2013) The effect of music on corticospinal excitability is related to the perceived emotion: a transcranial magnetic stimulation study. *Cortex* 49:702–710. Haueisen J, Knoösche TR (2001) Involuntary motor activity in pianists evoked by music perception. *J Cogn Neurosci* 13:786–792. Izumi S, Findley TW, Ikai T, Andrews S, Daum M, Chino N (1995) Facilitatory effects of thinking about movement on motor-evoked potential transcranial magnetic stimulation of the brain. *Am J Phys Med Rehab* 74:207–213. Meister IG, Boroojerdi B, Foltys H, Sparing R, Huber W, Topper R (2003) Motor cortex hand area and speech: implications for the development of language. *Neuropsychologia* 41:401–406. Meister IG, Krings T, Foltys H, Boroojerdi B, Müller M, Topper R, et al. (2004) Playing piano in the mind—an fMRI study on music imagery and performance in pianists. *Brain Res Cogn Brain Res* 19:219–228. Müller C, Arai N, Lücke J, Ziemann U (2009) Hysteresis effects on the input-output curve of motor evoked potentials. *Clin Neurophysiol* 120:1003–1008. Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychology* 9:97–113. Papathanasiou I, Filipovic SR, Whurr R, Rothwell JC, Jahanshahi M (2004) Changes in corticospinal motor excitability induced by nonmotor linguistic tasks. *Exp Brain Res* 154:218–225. Parncutt R, Sloboda J, Clarke E, Raekallio M, Desain P (1997) An ergonomic model of keyboard fingering for melodic fragments. *Music Perception* 14:341–382. Peretz I, Zatorre RJ (2005) Brain organization for music processing. *Annu Rev Psychol* 56:89–114. Rossi S, Hallett M, Rossini PM, Pascual-Leone A (2011) Screening questionnaire before TMS: an update. *Clin Neurophysiol* 122:1686. Rossi S, Pasqualetti P, Tecchio F, Pauri F, Rossini PM (1998) Corticospinal excitability modulation during mental simulation of wrist

movements in human subjects. *Neurosci Lett* 243:147–151. Rossi S, Spada D, Emanuele M, Olivelli M, Santarnecchi E, Fadiga L, et al. (2019) Cross-modal audiovisual modulation of corticospinal 52 F. Giovannelli et al. / *Neuroscience* 437 (2020) 45–53 motor synergies in professional piano players: A TMS study during motor imagery. *Neural Plast* 1328453. Rossi S, Hallett M, Rossini PM, Pascual-Leone A, Safety of TMS Consensus Group (2009) Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clin Neurophysiol* 120:2008–2039. Rossini PM, Burke D, Chen R, Cohen LG, Daskalakis Z, Di Iorio R, et al. (2015) Non-invasive electrical and magnetic stimulation of the brain, spinal cord, roots and peripheral nerves: Basic principles and procedures for routine clinical and research application. An updated report from an I.F.C.N. Committee. *Clin Neurophysiol* 126:1071–1107. Rossini PM, Rossi S, Pasqualetti P, Tecchio F (1999) Corticospinal excitability modulation to hand muscles during movement imagery. *Cereb Cortex* 9:161–167. Roux FE, Lubrano V, Lotterie JA, Giussani C, Pierroux C, De´monet JF (2007) When “abegg” is read and (“A, B, E, G, G”) is not: a cortical stimulation study of musical score reading. *J Neurosurg* 106:1017–1027. Scho´n D, Anton JL, Roth M, Besson M (2012) An fMRI study of music sight-reading. *NeuroReport* 13:2285–2289. Scho´n D, Besson M (2003) Audiovisual interactions in music reading. A reaction times and event-related potentials study. *Ann N Y Acad Sci* 999:193–198. Scho´n D, Besson M (2005) Visually induced auditory expectancy in music reading: a behavioral and electrophysiological study. *J Cogn Neurosci* 17:694–705. Scho´n D, Semenza C, Denes G (2001) Naming of musical notes: a selective deficit in one musical clef. *Cortex* 37:407–421. Sergent J, Zuck E, Terriah S, MacDonald B (1992) Distributed neural network underlying musical sight-reading and keyboard performance. *Science* 257:106–109. Sloboda JA (1984) Experimental studies of music reading: a review. *Music Percept* 2:222–236. Sparing R, Meister IG, Wienemann M, Buelte D, Staedtgen M, Boroojerdi B (2007) Task-dependent modulation of functional connectivity between hand motor cortices and neuronal networks underlying language and music: a transcranial magnetic stimulation study in humans. *Eur J Neurosci* 25:319–323. Stewart L, Henson R, Kampe K, Walsh V, Turner R, Frith U (2003) Brain changes after learning to read and play music. *Neuroimage* 20:71–83. Stewart L (2005) A neurocognitive approach to music reading. *Ann N Y Acad Sci* 1060:377–386. Strafella AP, Paus T (2000) Modulation of cortical excitability during action observation: a transcranial magnetic stimulation study. *NeuroReport* 11:2289–2292. Suppa A, Marsili L, Giovannelli F, Di Stasio F, Rocchi L, Upadhyay N, et al. (2015) Abnormal motor cortex excitability during linguistic tasks in adductor-type spasmodic dysphonia. *Eur J Neurosci* 42:2051–2060. Tokimura H, Tokimura Y, Oliviero A, Asakura T, Rothwell JC (1996) Speech-induced changes in corticospinal excitability. *Ann Neurol* 40:628–634. Willems RM, Hagoort P (2007) Neural evidence for the interplay between language, gesture, and action: a review. *Brain Lang* 101:278–289. (Received 27 January 2020, Accepted 16 April 2020) (Available online 23 April 2020) F. Giovannelli et al. / *Neuroscience* 437 (2020) 45–53 53

Perception & Psychophysics 2004, 66 (2), 183–195 Introspecting on the process of sight-reading, pianist Boris Goldovsky says: I have come to the conclusion that it [sight-reading] has something to do with the speed at which the visual image is converted into a muscular act; and with some people the transformation is so rapid that a great deal of what happens escapes awareness—it just happens. (from Wolf, 1976) Just as reading written text appears to be automatic for literate individuals, performing music from a written score appears to be automatic for those who have a reasonable degree of musical literacy, enabling them to play complex pieces of music at first sight. However, to date, the presumed automaticity of music reading has not been empirically demonstrated. Automaticity of word reading, in contrast, has been shown repeatedly by the classic Stroop task, in which printed names of colors

interfere with naming of the ink color (eg., the word RED written in blue ink), while ink color does not interfere with color word reading (MacLeod, 1991; Stroop, 1935). Conversely, facilitation occurs when both dimensions of the stimulus specify the same response (e.g., the word RED written in red ink) (Glaser & Glaser, 1982; MacLeod & Dunbar, 1988). The two main theoretical accounts of the Stroop effect are the relative speed account and the automaticity account. Briefly, the relative speed account (Morton & Chambers, 1973; Posner & Snyder, 1975) assumes that interference occurs at the response stage and that the direction of interference is determined by the relative speeds of each of the potential responses (one from the ink color, one from the color word). Since word reading is faster than color naming (Cattell, 1886), it is argued that word reading will always interfere with color naming, but not vice versa. The automaticity account focuses on the extent to which each process demands attention. Naming the ink color is assumed to draw more heavily on attentional processes than word reading does. Reading the word is seen to be obligatory, whereas naming the ink color is not. The speed of processing and automaticity accounts are not mutually exclusive; speed of processing and automaticity of a process will often, but not always, correlate, and both will be determined by experience. One model (Cohen, McClelland, & Dunbar, 1990), however, incorporates aspects of both the relative speed and the automaticity accounts. Its emphasis is on the strength, and not the speed, of processing. Each dimension of the stimulus (the color word and the ink color) provides input to a common response layer, and task demand units specify the dimension on which the response should be based (e.g., “respond according to ink color” or “respond according to color word”). When the information from both dimensions is incongruent (e.g., the word RED written in blue ink), response selection is delayed. When information from both dimensions is congruent (e.g., the word RED written in red ink), response selection is facilitated (but see MacLeod & MacDonald, 2000, for an alternative explanation of facilitation effects). In order to ask whether the presence of irrelevant musical notation has a measurable effect on the speed at which pianists execute a sequence of number to finger mappings, we used a novel musical Stroop task that was based on the 183 Copyright 2004 Psychonomic Society, Inc. This work was supported by Medical Research Council Grant G9617036; UF. V. W. is supported by a Royal Society University Research Fellowship. We thank John Morton, Sam Gilbert, Lucy Mills, Steven Dakin, William Thompson, and three anonymous reviewers. Correspondence may be addressed to L. Stewart, Institute of Cognitive Neuroscience, Alexandra House, 17 Queen Square, London WC1N 3AR, England (e-mail: l.stewart@ucl.ac.uk). Reading music modifies spatial mapping in pianists LAUREN STEWART, VINCENT WALSH, and UTA FRITH University College London, London, England We used a novel musical Stroop task to demonstrate that musical notation is automatically processed in trained pianists. Numbers were superimposed onto musical notes, and participants played five-note sequences by mapping from numbers to fingers instead of from notes to fingers. Pianists’ reaction times were significantly affected by the congruence of the note/number pairing. Nonmusicians were unaffected. In a nonmusical analogue of the task, pianists and nonmusicians showed a qualitative difference on performance of a vertical-to-horizontal stimulus–response mapping task. Pianists were faster when stimuli specifying a leftward response were presented in vertically lower locations and stimuli specifying a rightward response were presented in vertically higher locations. Nonmusicians showed the reverse pattern. No group differences were found on a task that required horizontal-to-horizontal mappings. We suggest that, as a result of learning to read and play keyboard music, pianists acquire vertical-to-horizontal visuomotor mappings that generalize outside the musical context. 184 STEWART, WALSH, AND FRITH classic language Stroop task. In our musical Stroop task, numbers, referring to the fingers of the right hand, are superimposed onto written musical notes, and participants execute a sequence of keypresses by mapping from numbers to fingers. Thus the numbers are

the analogue of the ink color, the relevant dimension, in the language Stroop task, and the notes, the analogue of the word, the “to-beignored” dimension. The hypothesis is that pianists possess two mappings: the mapping between musical notes and fingers, and the mapping between numbers and fingers. The former will be characterized by a greater strength of processing than will the latter, and it will interfere at the response selection stage. Nonmusicians, having only the number to finger mapping, will not show interference from the musical notation. Note that it is possible that the number-to-finger mapping will be stronger for the pianists than for the nonmusicians, since pianists often use numbers to indicate fingering in a piece of music. However, this would suggest that in pianists, the two types of mapping would be closer together in terms of their strength of processing, arguing against the hypothesis that musical notation will interfere. In Experiment 2, we used a nonmusical analogue of the musical Stroop task to test a hypothesis concerning the nature of the representation of musical notation. Musical notation is a system that maps pitch height to vertically organized spatial locations (specified by a set of horizontal lines called a staff), whereby high-frequency pitches are represented higher up on the staff than low-frequency pitches. Time is mapped horizontally, so that a sequence of musical events is mapped from left (earlier events) to right (later events) on the staff. For instruments such as the piano, the vertical position on the score also maps onto the horizontal location of the note on the keyboard, so that notes that are visually higher (and higher in pitch) are produced toward the right of the keyboard and notes that are visually lower (and lower in pitch) are produced toward the left. We hypothesized that the representations used during music reading would be characterized by a set of vertical-to-horizontal stimulus–response mappings. If pianists have developed such spatial mappings, they may be evident on tasks that involve stimulus–response mappings similar to those used in music reading but that lack any surface resemblance to a music reading situation.

EXPERIMENT 1 Musical Stroop Task

The effect of irrelevant musical notation on performance of the task was measured by comparing response times for executing the motor sequence under different conditions of number/note congruence.

Method

Participants

Two participant groups were used: 12 pianists (10 female, average age 26), recruited from the Royal Academy of Music and 14 nonmusicians (10 female, average age 22), recruited from within University College London. On average, the participants in the pianist group had been playing piano for 20 years (range: 13–34 years). The participants in the nonmusician group had no experience of reading or playing music. All participants were healthy and right-handed and gave their informed consent. The participants were naive to our experimental hypotheses.

Stimuli

Stimuli were generated using the music notation software package Sibelius (Sibelius Group, <http://www.sibelius.com>). Each stimulus was a notated bar of five crotchets (quarter notes) on which numbers were superimposed. Each note (G, A, B, C, D) and each number (1, 2, 3, 4, 5) appeared once in every stimulus. Five types of stimuli were used (Figure 1). Baseline stimuli contained no musical notation and consisted of a row of five white numbers against a background black strip. Congruent stimuli were musically congruent and therefore also spatially systematic: G, the lowest note used, would be labeled “1”; A, the next lowest note used, would be labeled “2”; and so on. This mapping is similar to that used by keyboard players reading a score; notes extending from the bottom to the top of the staff map respectively onto digits extending from the left to the right of the hand. Incongruent (random) stimuli were musically incongruent and spatially unsystematic: B might be labeled “4”; G might be labeled “5”; C might be labeled “1”; and so on. The relationship between notes and numbers was not consistent across stimuli of this type, so the participants could not learn an arbitrary number/note pairing. The incongruent (random) stimuli were designed to exhibit an approximately equal degree of note/number incongruence across the set. The method of determining note/number incongruence was as follows: If the specified number to finger mapping was, for example, 13542, each pair of

consecutive elements of the sequence was considered separately (e.g., 1–3, 3–5, 5–4, 4–2) and the direction and magnitude of each difference was calculated. The first pair, 1–3, shows a positive difference with a magnitude of 2; thus, these numbers would be superimposed onto a pair of notes opposite in their direction (proceed from high to low on the staff) and separated by a magnitude other than 2. The possible note pairs were calculated for each of the number pairs, yielding several possibilities. The choice of note pairs for an entire sequence was made so that each of the notes G to D was used only once in combination. Incongruent (systematic) stimuli were musically incongruent but spatially systematic: the number/note relationship was the inverse of that used in the congruent trials: G, the lowest note used, would be labeled “5”; A, the next lowest note used, would be labeled “4”; and so on. Since stimuli were either completely congruent or completely incongruent, we used catch trials to deter participants from using a note-reading strategy on congruent trials. These stimuli were part congruent and part incongruent (random): Two of the last three elements would be unpredictably incongruent. A music-reading strategy on these trials would therefore be maladaptive. Twelve motor sequences were used, each of which could be presented as baseline, congruent, incongruent (random), incongruent (systematic), or catch trials. Each trial was presented twice, making a total of 120 trials.

Task. Participants sat with the fingers of the right hand resting over five adjacent keys of a laptop keyboard. A computer keyboard was used rather than a piano keyboard, in order to equate familiarity with the response interface across the two groups. The stimuli were presented and the responses recorded using a program written in MATLAB (The Mathworks, <http://www.mathworks.com>). Each trial consisted of one stimulus presentation. Before each trial, a central fixation point appeared for 1 sec. Following this, the stimulus was presented in the center of the screen for 3 sec or 4 sec for the pianist and nonmusician groups, respectively (timings were based on a pilot study), after which time the fixation point reappeared for another second before the second stimulus appeared. The participants were instructed to ignore the musical notation and use only the numbers to perform the task. The number “1” inside a note indicated that the participant should press the key beneath the thumb, the number “2” indicated a keypress beneath the index finger, and so on. The participants were instructed to read from left to right, mapping the number information presented inside each note to the correct key as **READING MUSIC MODIFIES SPATIAL MAPPING IN PIANISTS** 185 quickly and as accurately as possible. It was stressed that the hand should remain over the same five keys throughout the experiment. A practice session consisting of five trials was given to familiarize the participants with the task. Twelve blocks, each comprising 10 stimuli, were presented. Trials were pseudorandomly ordered across motor sequence and trial type (baseline/congruent/incongruent(random)/incongruent(systematic)/ catch). First and second presentations of each trial were also pseudorandomly ordered. Thus the participants did not know in advance what trial type or motor sequence to expect.

Predictions The pianists’ response times were predicted to be shorter on the congruent trials than on the baseline trials. In the case of baseline trials, the response could only be arrived at by using the number-to-finger mapping. In the case of congruent trials, both the number-to-finger mapping and the note-to-finger mapping specified the same response. Although instructed to ignore musical notation, if pianists automatically read musical notation, they would derive facilitation from the congruent notation at the response stage. The pianists’ response times were predicted to be longer on both types of incongruent trials than on the baseline trials. The number-to-finger mapping and the note-to-finger mapping used in the incongruent trials specified different responses. Again, although the instruction was to ignore musical notation, if music reading is automatic, these two conflicting stimulus–response mappings would result in interference at the response stage. Nonmusicians were predicted to be unaffected by the trial type (baseline/congruent/incongruent/catch). Nonmusicians have no existing pathway from notes to fingers; thus,

regardless of the congruence of the number-to-finger mapping and note-to-fingermapping, only the number-to-finger mapping would be used in selection of a response. The pianists' response times were predicted to be longer on the incongruent (random) trials than on the incongruent (systematic) trials. The incongruent (systematic) trials were characterized by a rule-based relationship between the number and its location on the staff; the incongruent (random) trials were not. The spatial systematicity governing where a number would appear could therefore be used, in addition to the number information, for selection of the appropriate response. The congruent trials were also characterized by a spatial systematicity that governed where the numbers appeared. However, the kind of spatial systematicity contained in the congruent trials, in contrast to that in the incongruent (systematic) trials, corresponded to that used in musical notation. Since nonmusicians have no existing note-to-finger mapping, we predicted that the effect of spatial systematicity in the irrelevant dimension would be negligible.

Statistical Analysis Errors. Because, in the cumulative analysis (see below), each five-element sequence was treated as a single trial, a trial was discarded if a participant made one or more errors out of a possible five. Response time. Keypress identities and onsets were collected and used for two types of statistical analysis; cumulative, in which the time taken to execute each entire sequence was calculated, and itemized, in which the response times for each of the responses (first to fifth) in a sequence were used. Cumulative analysis. Response time data was processed for each participant separately. Trials on which the participant produced an incorrect sequence were discarded. Since each stimulus was presented twice, the cumulative response time for each trial was averaged across these two presentations. In cases where one of the two trials was discarded because of production of an incorrect sequence, the cumulative response time was taken from the single remaining trial. The cumulative response times were collated into motorically matched sets, where a set comprised each trial type of 1 of the 12 motor sequences. If any set of data was incomplete—for instance, because a participant had produced an error on both presentations of a particular trial type for that sequence—the entire set of data was removed. Although this resulted in loss of data, it was necessary to ensure that any eventual difference.

Figure 1. Each motor sequence (12 in total) could appear with five different number/note arrangements. Here each trial type is shown for one motor sequence (1, 4, 2, 5, 3): baseline, congruent, incongruent (random), incongruent (systematic), and catch. Motor sequences and trial types were pseudorandomly interspersed.

186 STEWART, WALSH, AND FRITH ences in cumulative response time as a function of trial type (baseline/congruent/incongruent(random)/incongruent(systematic)/catch) would not merely reflect the frequency distribution of certain motor sequences. If, after this process, only 50% or fewer sets of data remained, the participant was eliminated from further analysis. For the remaining sets, cumulative response times were averaged across each trial type. To exclude overall differences in response time as a result of pianists' simply being faster to make the keypresses, cumulative response times for each participant were expressed as a proportion of that participant's cumulative response time for baseline trials. Itemized analysis. The data corresponding to all the first keypresses of every sequence were considered separately from the data corresponding to all the second keypresses of every sequence, and so on. In each case, the data were subjected to the same procedures as those described for the cumulative analysis: For each participant, response times were averaged across the two trials and collated into motorically matched sets according to trial type; incomplete sets of data were discarded and response times were averaged across each trial type for each participant separately. To exclude overall differences in response time, response times for each keypress and for each participant were expressed as a proportion of the cumulative response time for that participant's baseline trials.

Results Errors. Descriptive data are shown for each group in Table 1. Errors shown are out of a possible 24. There was no significant effect of trial type on

the number of errors produced for either group. Participants were no more error prone on catch trials than on congruent trials, from which we can infer that they were using a number-reading strategy for both of these trial types. Cumulative response time. Two participants (1 from each group) were excluded because the number and pattern of their errors resulted in the removal of more than 50% of the data sets. Descriptive data and corresponding graphs for the remaining subjects are shown in Table 1 and Figure 2. A repeated measures analysis of variance (ANOVA) of mixed design with trial type (4 levels: congruent, incongruent [random], incongruent [systematic], catch, all as a proportion of baseline) as a within-subjects factor and group (2 levels: pianist, nonmusician) as the between-subjects factor revealed a significant interaction [$F(2.3, 51.6) = 23.96, p = .001$]. Two separate repeated measures ANOVAs, one for each group, with trial type as a within-subjects factor (4 levels) revealed a significant effect of trial type for pianists only [$F(3, 30) = 39.84, p = .001$]. It was necessary to ascertain that any potential systematic bias in the removal of data sets from each group could not account for the group \times trial type interaction described. Although the stimuli were designed so that the degree of congruence/incongruence was equated across all sequences, it was possible that a group \times trial type interaction might have resulted if, after removal of the data sets containing erroneous responses, the data sets remaining for the pianists were more susceptible to the effects of trial type than those remaining for the nonmusicians. Thus we tested whether the set of motor sequences represented in the pianists' data was more affected by trial type in comparison with the motor sequences represented in the nonmusicians. Because, on occasion, errors were made on both presentations of a given trial, the data set for some participants was incomplete. If a participant was missing data pertaining to more than 3/12 sets of stimuli, the participant was excluded from that specific analysis (0 pianists, 3 nonmusicians). For each of the remaining participants, missing data were replaced with the average value for all the trials of that type. A mixed-design repeated measures ANOVA with sequence (12 levels: each motor sequence) and congruency [4 levels: congruent, incongruent (random), incongruent (systematic), catch, all as a proportion of cumulative baseline response] as within-subjects factors and group (2 levels: pianist, nonmusician) as the between-subjects factor revealed no significant interaction. Neither of the within-subjects ANOVAs performed separately for pianists and nonmusicians revealed an interaction of sequence \times trial type. The original trial type \times group analysis was repeated, excluding the 3 participants who had been removed from the sequence \times trial type \times group interaction because of missing data. The trial type \times group interaction was replicated [$F(3, 36) = 8.02, p = .001$], precluding the possibility that it had resulted from the removal of different data sets from each group. To explore the interaction further, the pianists' data were subjected to five planned paired sample tests. Significant differences were found between baseline and congruent trials [$t(10) = 2.06, p = .03$], baseline and incongruent random trials [$t(10) = 29.65, p = .001$], baseline and incongruent (systematic) trials [$t(10) = 27.49, p = .001$], incongruent (random) and incongruent (systematic) trials [$t(10) = 2.22, p = 0.02$], and congruent and catch trials [$t(10) = 25.11, p = .001$]. Itemized: Main effect of sequential position (Figure 3, Table 2). In order to test for possible serial order effects, we looked at response times for each position (first Table 1 Musical Stroop Task: Mean Numbers of Errors (Out of 24) and Cumulative Response Times (in Milliseconds) Errors Response Time Subjects M SD M SD Pianists Baseline 3.08 2.15 1,559 271 Congruent 2.08 2.68 1,507 225 Incongruent (Random) 3.00 2.83 1,811 335 Incongruent (Systematic) 2.75 2.42 1,760 331 Catch 3.17 2.08 1,644 252 Nonmusicians Baseline 4.33 3.28 2,725 387 Congruent 3.50 2.35 2,766 409 Incongruent (Random) 4.75 3.89 2,758 385 Incongruent (Systematic) 4.42 3.23 2,730 392 Catch 4.67 3.37 2,721 353

READING MUSIC MODIFIES SPATIAL MAPPING IN PIANISTS 187 to fifth) in the sequence, regardless of trial type. Response times were averaged across baseline, congruent, incongruent (systematic), incongruent (random), and

catch trials. A mixed-design repeated measures ANOVA with sequential position (5 levels: first, second, third, fourth, fifth) as the within-subjects factor and group (2 levels: pianist, nonmusician) as the between-subjects factor revealed an interaction [$F(1.3, 26.02) = 27.80, p = .001$], which appeared to arise from the greater drop in response time between the first and subsequent keypresses for pianists than for nonmusicians. Separate repeated measures ANOVAs, with sequential position (5 levels: first, second, third, fourth, fifth) as the only within-subjects factor, revealed a significant effect of sequential position for both groups: pianists [$F(1.23, 12.27) = 116.61, p = .001$]; nonmusicians [$F(1.66, 16.62) = 136.26, p = .001$]. Although this effect appears to have been mainly due to the relatively long response time taken for the first keypress relative to subsequent keypresses, a significant serial order effect was also seen when the response times were entered from only the second, third, fourth and fifth sequential positions (leaving out the first). This effect was significant in the nonmusicians [$F(3, 30) = 23.17, p = .001$] and approached significance in pianists [$F(1.70, 16.98) = 3.123, p = .07$].

Itemized: Effect of trial type on sequential position (Figure 4, Table 3). To investigate whether the trial type 3 group interaction that we previously demonstrated in the cumulative analysis still held for data corresponding to individual responses of the sequence, we performed a mixed-design repeated measures ANOVA for each sequential position separately, using trial type as the within-subjects factor (5 levels: baseline, congruent, incongruent [random], incongruent [systematic], catch, all as a proportion of cumulative baseline response) and group (2 levels: pianists and nonmusicians) as the between-subjects factor. A significant trial type 3 group interaction was found for each sequential position: first [$F(4, 80) = 12.72, p = .001$], second [$F(4, 80) = 3.48, p = .011$], third [$F(2.39, 47.8) = 5.34, p = .005$], fourth [$F(4, 80) = 2.73, p = .034$], and fifth [$F(4, 80) = 2.55, p = .046$]. Separate within-subjects ANOVAs for each group and at each position revealed that nonmusicians showed an effect of trial type (Figure 2). Cumulative response time (time from stimulus onset to final keypress of the sequence) according to trial type [baseline/congruent/incongruent (random)/incongruent (systematic)/catch] for pianists (white) and nonmusicians (black). Data are normalized to the cumulative response time data from baseline trials. Error bars represent 61 standard error of the mean. Table 2 Musical Stroop Task: Mean Itemized Response Times, Regardless of Trial Type, as a Proportion of Baseline Cumulative Response Times

Subjects	M	SD	Keypress 1	Keypress 2	Keypress 3	Keypress 4	Keypress 5
Pianists	.61	.12	.12	.04	.13	.05	.1
Nonmusicians	.39	.05	.16	.02	.18	.02	.16

188 STEWART, WALSH, AND FRITH type (5 levels) for the first sequential position only. A paired samples t test (baseline vs. incongruent [random]) showed that this was carried by a selective increase in response time for incongruent (random) trials [$t(11) = 23.873, p = .003$], suggesting that nonmusicians were sensitive to the lack of spatial systematicity in these trials. Nevertheless, this effect was not seen in the cumulative analysis (see above). The pianists showed a significant effect of trial type (5 levels) at every sequential position: first [$F(4, 40) = 28.03, p = .001$], second [$F(1.71, 17.13) = 4.57, p = .03$], third [$F(1.46, 14.59) = 7.61, p = .009$], fourth [$F(2.19, 21.93) = 9.72, p = .001$], and fifth [$F(4, 40) = 4.10, p = .007$].

Itemized: Sequential position 3 trial type. We assessed whether the effect of trial type differed according to the sequential position of the response. First, we asked whether any potential interaction differed between groups. A mixed-design repeated measures ANOVA with trial type (5 levels: baseline, congruent, incongruent [random], incongruent [systematic], catch, all as a proportion of cumulative baseline response) and sequential position (5 levels: first, second, third, fourth, fifth) as within-subjects factors and group (2 levels: pianists, nonmusicians) as the between-subjects factor revealed a significant interaction [$F(5.67, 113.5) = 4.36, p = .001$]. Separate within-subjects ANOVAs for each group revealed a significant trial type 3 sequential position interaction for pianists only [$F(3.33, 33.28) = 11.80, p = .001$]. The same analysis carried out in

pianists for only the second, third, fourth, and fifth sequential positions (leaving out the first) failed to result in a significant interaction, suggesting that the interaction was carried by a larger effect of trial type on the first response than on subsequent responses. Discussion This experiment showed that, for musically literate pianists, the presence of irrelevant musical notation affects the speed at which they can produce a motor sequence, Figure 3. Response times itemized according to sequential position (first, second, third, fourth, fifth), averaged across all trial types. Response time for first sequential position represents the time elapsed between stimulus onset and the first keypress of the sequence. Response time for second sequential position represents the time elapsed between the first and second keypresses, and so on. Pianists are depicted by the dark line; nonmusicians, by the light line. The data are expressed as a proportion of the cumulative response time (stimulus onset to fifth response) for baseline trials. Error bars represent 61 standard error of the mean. **READING MUSIC MODIFIES SPATIAL MAPPING IN PIANISTS** 189 using a number-to-fingermapping. In much the same way that decoding from a written word to a spoken output seems to be obligatory (Stroop, 1935), musical notation appears to represent a highly salient and overlearned stimulus for this group of individuals. Nonmusicians, in contrast, are unaffected by the presence of musical notation. Note that this effect is seen despite pianists' potentially greater experience of mapping numbers to fingers, in comparison with nonmusicians, which would work to minimize any interference effects from the musical notation. This differential effect of musical notation cannot be explained by the removal of a different set of sequences (due to errors) for pianists versus nonmusicians. A comparison of response times for congruent trials and response times for incongruent (systematic) trials revealed a relative interference effect for incongruent trials of the order of 300 msec. Critically, although these two trial types were equated on the degree of inherent spatial systematicity, only the congruent trials made use of the particular spatial systematicity that occurs in musical notation. A separate comparison of congruent and incongruent (systematic) trials, each with baseline, revealed that the congruent 2 incongruent (systematic) difference comprised both facilitation and interference. This pattern of data, with interference substantially greater than the facilitation, is typical in Stroop-like tasks (Glaser & Glaser, 1982; MacLeod & MacDonald, 2000). It has been argued that facilitation and interference may not represent opposite sides of the same coin, but rather that facilitation may be an artefact of response error (MacLeod, 1998). Although our data do not pertain to this question directly, the inclusion of catch trials makes it unlikely that the facilitation observed was the result of response error. When the data were broken down so that we could look at the response times for individual keypresses in the sequence, the overall pattern of response times (averaged across all trial types) showed evidence of response preparation, since time taken to make the first keypress was considerably longer than time taken to make subsequent keypresses. This is consistent with previous work on speech production, which has shown that latency increases linearly with the number of items in the sequence (Monsell & Sternberg, 1981). In addition, there was evidence of a serial order effect for the subsequent keypresses, which was qualitatively similar across the two groups. Notably, response time was increased for the third keypress. Serial order effects have also been reported in Figure 4. Response times, according to trial type, for each sequential position. Pianists are represented by the white bars; nonmusicians, by the black bars. The data are expressed as a proportion of the cumulative response time (stimulus onset to Keypress 5) for baseline trials. Error bars represent 61 standard error of the mean. 190 STEWART, WALSH, AND FRITH other transcription tasks such as typing (Sternberg, Monsell, Knoll, & Wright, 1978). The effect of trial type that was seen in the cumulative analysis was also seen at the level of the individual response times for each position in the sequence: for Sequential Positions 1 to 5, the effect of musical notation on response time was greater for pianists than for nonmusicians. The pianists also showed an

interaction of trial type with sequential position, which appeared to arise because the effect of trial type was stronger at Keypress 1 than at the other positions. However, the effect of trial type was not restricted to the first keypress, suggesting that number/note congruency made a contribution, not only during the initial response preparation period but also during the on-line execution of the sequence. EXPERIMENT 2 Nonmusical Stroop Task Having demonstrated, in Experiment 1, that musical notation is automatically processed in musically literate pianists, in Experiment 2 we investigated the nature of the representation of musical notation. One of the major differences between music reading and text reading is that music reading almost always involves a response component, whereas text reading can, and normally is, performed without explicit mapping to a response system. Thus the automaticity demonstrated in Experiment 1 is likely to act at the interface between perception and action—that is, at the stimulus–response level. The more compatible a stimulus–response pairing, based on the degree of stimulus–response overlap on some physical or representational dimension, the greater the response time difference that will exist when this stimulus–response pairing is violated (Kornblum, Hasbroucq & Osman, 1990). For instance, it is easy to map a stimulus appearing on the right onto a rightward response (e.g., the right hand) and likewise a stimulus appearing on the left to a leftward response, but there is a reaction time cost associated with mapping a stimulus appearing on the right to a leftward response and vice versa. In contrast, stimuli that appear at different vertical locations (e.g., high/low) do not naturally map onto a particular response in the horizontal meridian. For instance, there does not appear to be a natural correspondence between high and right, low and left, or vice versa. However, the high-right, low-left mapping is precisely that used by pianists when reading music, since musical notes vary in their vertical location on the staff and systematically map onto a horizontal response system (the keyboard) in this way. In this experiment, we tested the hypothesis that music reading, for keyboard playing, can be characterized by a set of vertical-to-horizontal stimulus–response mappings. A vertical-to-horizontal stimulus–response mapping task was used to measure response time benefits and costs to making a response that was spatially compatible or incompatible with the learned stimulus–response mappings used by pianists reading music. A horizontal-to-horizontal stimulus–response mapping task served as a control task. Method Participants. Eight pianists (7 female; average age, 25) and 14 nonmusicians (10 female; average age, 22) took part in the experiment. All participants had also taken part in Experiment 1, and all remained naive to our experimental hypotheses. Stimuli. In this experiment, we used a nonmusical analogue of the musical Stroop task. The stimuli were numbers presented at different spatial locations. A standard graphics package was used to draw the stimuli. Two versions were used: a vertical-to-horizontal stimulus–response task and a horizontal-to-horizontal stimulus–response task (Figures 5 and 6). In both tasks, five numbers were presented consecutively in five different locations. The stimulus types were the following: baseline, congruent, and incongruent (systematic). Incongruent (random) stimuli were not used because Experiment 1 had already confirmed that spatial systematicity (regardless of musical congruence) affects response time. Thus the comparison of interest occurred between trials equated on spatial systematicity, such as the congruent and incongruent (systematic trials). The number/location relationships were based on those in the analogous stimuli of Experiment 1. In the horizontal-to-horizontal task, congruent trials were characterized by a “1” appearing in the leftmost position, a “2” appearing in the position second from the left, and so on. In the vertical version, congruent trials were characterized by a “1” appearing in the bottommost position, a “2” appearing in the position second from the bottom, and so on. The stimulus–response mapping used for congruent trials in the vertical

Table 3 Musical Stroop Task: Mean Itemized Response Time, According to Trial Type, as a Proportion of Baseline Cumulative Response Time

Pianists	Nonmusicians	Position in Sequence	M	SD	M	SD
First	Baseline	.57	.10	.38	.04	

Congruent .57 .11 .39 .05 Incongruent (random) .67 .12 .40 .06 Incongruent (systematic) .64
 .13 .39 .05 Catch .59 .13 .38 .05 Second Baseline .11 .04 .17 .03 Congruent .11 .03 .16 .02
 Incongruent (random) .12 .05 .16 .02 Incongruent (systematic) .13 .05 .16 .02 Catch .12 .04
 .16 .02 Third Baseline .12 .04 .18 .02 Congruent .11 .03 .18 .02 Incongruent (random) .15
 .06 .18 .02 Incongruent (systematic) .14 .06 .18 .02 Catch .13 .06 .19 .02 Fourth Baseline
 .09 .02 .16 .02 Congruent .09 .02 .16 .01 Incongruent (random) .11 .03 .16 .02 Incongruent
 (systematic) .11 .03 .16 .02 Catch .11 .03 .16 .02 Fifth Baseline .10 .02 .12 .02 Congruent
 .10 .02 .12 .02 Incongruent (random) .11 .02 .12 .02 Incongruent (systematic) .10 .02 .12 .02
 Catch .10 .03 .12 .02 READING MUSIC MODIFIES SPATIAL MAPPING IN PIANISTS

191 to-horizontal version of the task paralleled the mapping used in the reading of keyboard music, in which notes ascending from the bottom to the top of the staff map onto fingers extending from the left to the right of the hand. The incongruent (systematic) stimuli exhibited the inverse relationship to that used in the congruent trials. Whereas the stimuli in the musical Stroop task consisted of five notes presented simultaneously (requiring the participant to proceed serially from left to right), the trials in Experiment 2 consisted of five nonmusical stimuli presented one by one. A keypress made in response to the first stimulus triggered the appearance of the next. This constraint prevented any strategic differences' occurring between the two groups (perceptual or oculomotor) that might account for a performance difference. Task. In both tasks, the participants placed the right hand over five adjacent keys of a computer laptop keyboard, as in the musical Stroop task. They were told that a number would appear in one of five locations, and they were instructed to map the number presented onto the appropriate finger and make the required keypress. It was stressed that participants should maintain their hand position above the designated keys throughout the experiment, ignoring any information provided by the spatial location at which the number appeared. The stimuli were presented and the results recorded using a program written in MATLAB. The stimuli appeared one at a time on the computer monitor. As soon as a response was made, that stimulus disappeared from the screen and the next stimulus appeared. After five responses had been made, by each finger once, there was a 1-sec pause before the first stimulus of the next trial appeared. Both the vertical-to-horizontal and horizontal-to-horizontal versions were run, counterbalanced in order across all participants to control for transfer effects between experiments. Figure 5. Horizontal-to-horizontal task: Each trial consisted of the sequential presentation of five numbers, within a horizontal array of boxes. Location was to be ignored, and participants performed a simple number-to-finger mapping task. Each keypress triggered the onset of the subsequent number. Here each trial type is shown for one motor sequence (1, 4, 2, 5, 3): baseline, congruent, and incongruent (systematic). Motor sequences and trial types were pseudorandomly interspersed. 192 STEWART, WALSH, AND FRITH Predictions Horizontal-to-horizontal task. It was predicted that both pianists and nonmusicians would have shorter response times on congruent trials and longer response times on incongruent trials than on baseline trials because of the spatial overlap between the location of the stimulus and the required response. Vertical-to-horizontal task. It was predicted that pianists would have shorter response times on congruent trials and longer response times on incongruent trials than on baseline trials because of their learned association between the vertical position of a stimulus and its horizontal response. It was predicted that nonmusicians would have similar response times on congruent trials and incongruent trials because, unlike pianists, they would have had no reason to acquire an association between vertical stimuli and horizontal responses. Statistical Analysis Response time data were processed for each participant separately. Each set of five consecutive stimuli was treated as a single trial, and each trial was analyzed for cumulative response time. An itemized analysis was not performed. The data were treated in the same way as for Experiment 1: Trials on which the participant produced an Figure 6. Vertical-to-

horizontal task: Each trial consisted of the sequential presentation of five numbers, within a vertical array of boxes. The location at which the number was presented was to be ignored, and participants performed a simple number-to-finger mapping task. Each keypress triggered the onset of the subsequent number. Here each trial type is shown for one motor sequence (1, 4, 2, 5, 3): baseline, congruent, incongruent (systematic). **READING MUSIC MODIFIES SPATIAL MAPPING IN PIANISTS** 193 incorrect sequence were discarded, the cumulative response time for each trial was averaged across both presentations, response times were collated into motorically matched sets comprising one of each trial type, and incomplete data sets were removed. For the remaining sets, cumulative response times were averaged across each trial type for each participant separately. This data processing was carried out separately for each task version (vertical to horizontal and horizontal to horizontal). In order to exclude overall differences in response time, response times for each participant were expressed as a proportion of the cumulative response time for baseline trials. Results Errors. Descriptive data are shown for each group in Table 4. Errors shown are out of a possible 24 for each trial type. Wilcoxon signed ranks tests revealed that the pianists showed no significant effect of trial type on error rate for either version of the task. The nonmusicians, however, made significantly more errors on the incongruent (systematic) trials than on the congruent trials in the horizontal-to-horizontal task [$Z(22.64)$, $p = .008$]; they did not differ in the number of errors produced in the vertical-to-horizontal task. Response time. Descriptive data and the corresponding graphs are shown for each group in Figure 7 and Table 4. A repeated measures ANOVA of mixed design with task version (2 levels: horizontal to horizontal, vertical to horizontal) and trial type [2 levels: congruent, incongruent (systematic), both as a proportion of baseline] as within-subjects factors and group (2 levels: pianists, nonmusicians) as the between-subjects factor revealed a significant interaction of task version \times trial type \times group [$F(1,20) = 13.30$, $p = .002$]. A repeated measures ANOVA with task version (2 levels) and trial type (2 levels) as within-subjects factors revealed a significant interaction for nonmusicians alone [$F(1,13) = 31.43$, $p = .001$]. A repeated measures ANOVA for pianists alone with task version (2 levels) and trial type (2 levels) as within-subjects factors and order of experiments (2 levels: Experiment 1 first, Experiment 2 first) as a between-subjects factor did not reveal a significant effect of order. Thus no evidence was found to suggest that the pattern of results could be accounted for by transfer effects from Experiment 1 to Experiment 2.

Discussion The results of Experiment 2 support the hypothesis that musically literate pianists, through extensive practice in reading and playing keyboard music, acquire a set of vertical-to-horizontal stimulus–response mappings that generalize outside of the musical context. Pianists and nonmusicians both showed the expected interference on the horizontal-to-horizontal task, but pianists and nonmusicians showed the opposite pattern of results on the vertical-to-horizontal task. Pianists, as predicted, were facilitated when numbers specifying leftward responses were presented at vertically lower locations and numbers specifying rightward responses were presented at vertically higher locations, the same stimulus–response mapping that is used when a pianist reads music. Nonmusicians, on the other hand, were facilitated when numbers specifying leftward responses were presented at vertically higher locations and numbers specifying rightward responses were presented at vertically lower locations, opposite to the stimulus–response mapping used when pianists read music. Although the pattern of results seen in the nonmusicians was not predicted, it can be explained if we consider the way in which numbers are normally represented in vertical space. When numbers are listed, “1” will usually appear at the top, and subsequent numbers will occupy vertically lower positions. If individuals associate the number “1” with the thumb, the number “2” with the index finger, and so on, their experience of numerical lists may cause them to be quicker to map from a “1” presented in the highest vertical position as opposed to the lowest position. This is to suggest that the effect of trial type seen in the

nonmusicians arose because of a learned mapping between space (the experience of numbers ordered vertically from high to low) and digits. If nonmusicians performed the experiment with the left hand, the prediction would be that their responses would be facilitated when small numbers appear in vertically higher locations and larger numbers appear in vertically lower locations. Spatially, this means that, for a nonmusician using the right hand, a leftward response (e.g., pressing the key beneath the thumb [1]) would be facilitated by the stimulus appearing at a vertically higher location, whereas for a nonmusician using the left hand, a leftward response (e.g., pressing the key beneath the little finger [5]) would be facilitated by the stimulus appearing at a vertically lower location. The trial type effect seen in pianists, on the other hand, we argue, arises because of a learned mapping between visual space (experience of notes that vary vertically) and response space (the layout of the keyboard). Thus when pianists perform with either the right or the left hand, leftward responses will be facilitated when the response-eliciting stimulus appears in a vertically higher location.

Table 4
Nonmusical Stroop Task: Mean Numbers of Errors (out of 24) and Mean Cumulative Response Times (in Milliseconds)

Subjects	M	SD	M	SD
Pianists				
Hz to Hz, Congruent	2.00	1.93	2,654	364
Hz to Hz, Incongruent	2.88	3.04	2,792	408
Vt to Hz, Congruent	3.25	2.38	2,785	382
Vt to Hz, Incongruent	4.38	3.20	2,897	419
Nonmusicians				
Hz to Hz, Congruent	2.64	1.60	3,592	475
Hz to Hz, Incongruent	5.36	3.46	3,887	533
Vt to Hz, Congruent	4.36	2.90	3,970	538
Vt to Hz, Incongruent	5.57	2.10	3,814	496

STEWART, WALSH, AND FRITH location and rightward responses will be facilitated when the response-eliciting stimulus appears in a vertically lower location. The particular vertical-to-horizontal stimulus–response mappings are likely to be specific to pianists. Other instrumentalists who produce notes by covering certain holes and depressing other keys (woodwind) or by combining fingering with contraction of lip and diaphragm muscles (brass), having had no reason to acquire vertical-to-horizontal stimulus–response mappings, would be unlikely to show equivalent interference on the vertical-to-horizontal Stroop task. However, most conservatory musicians who are not first-study pianists nevertheless possess good keyboard skills, making this a difficult prediction to test.

GENERAL DISCUSSION

Experiment 1 showed that, for musically literate pianists, the presence of irrelevant musical notation can affect the speed at which numbers are converted into a sequence of keypresses. Experiment 2 showed that pianists possess a set of stimulus–response mappings that correspond to the stimulus–response mappings required for one to read and play music. Nonmusicians showed evidence of a different set of stimulus–response mapping, which we attribute to a possible learned association between the vertical organization of numbers in lists and specific digits. The pianists’ results from Experiment 2 suggest that musical notation for keyboard performance is represented as a set of vertical-to-horizontal stimulus–response mappings that persist even outside of a musical context. Violation of these learned mappings is therefore the likely source of interference in the musical Stroop task seen in Experiment 1. The musical Stroop effect found in Experiment 1 suggests that music reading, like text reading, is automatic. However, as Cohen’s model emphasizes (Cohen et al., 1990), automaticity is not an all-or-none phenomenon and is better thought of as a continuum. Automaticity, and thus the interference effects observed in Stroop-like tasks, is subject to the effects of practice. The data on which Cohen’s model is based show that 20 h of training on a simple shape-naming task can reverse the direction of interference in a Stroop-like situation (MacLeod & Dunbar, 1988). The pianists in our study had received extensive practice on music reading, averaging 20 years. However, interference effects on the musical Stroop task have also been demonstrated in a group of novice pianists who had been playing the keyboard for only 3 months (Stewart et al., 2003). Whether the automaticity of music reading seen in these novices is as robust as that seen in the highly trained pianists used in our study is a question that remains to be addressed. In addition, longitudinal studies

will be required so that one can ascertain the precise point at which interference effects first develop and at what point they reach a maximum level. The results of Experiment 2 suggest that the acquisition of keyboard skills modifies spatial mapping. Within the music performance domain, knowledge accrued through practice has been shown to transfer at a motoric and/or conceptual level (Palmer & Meyer, 2000), but the effects of music have also been claimed to transfer outside the Figure 7. Cumulative response time (time from stimulus onset to fifth response for horizontal-to-horizontal task and vertical-to-horizontal task. Pianists are represented by the dark lines; nonmusicians, by the light lines. Response times are expressed as a proportion of the response for baseline trials. Error bars represent 61 standard error of the mean.

READING MUSIC MODIFIES SPATIAL MAPPING IN PIANISTS

195 musical domain (Chan, Ho, & Cheung, 1998; Graziano, Peterson, & Shaw, 1999; see Schellenberg, 2001, for a review). Reports that music can affect spatial processing emanate from two different types of study: investigations of the nonmusical consequences of listening to music (i.e., the “Mozart Effect”; Rauscher & Shaw, 1998; Rauscher, Shaw, & Ky, 1995; Rauscher et al., 1997), and of the nonmusical consequences of musical instruction (reviewed by Hetland, 2000). Studies of the latter kind have reported effects of musical instruction on a wide variety of spatial tasks: spatial-temporal tasks, tasks requiring spatial memory, spatial recognition, mental rotation, and spatial visualization. However, numerous differences among the studies in terms of type and duration of musical instruction, whether or not a control training program was used, the age of the children, and the spatial tasks themselves make it difficult to draw any firm conclusions regarding precisely what is transferred from musical instruction and what conditions are necessary for this transfer to take place. Two kinds of theories have been proposed. “Neural connection” theories such as the “trion” theory (Shaw, 2000) suggest that musical and spatial processing centers in the brain are proximal or overlapping in such a way that the development of certain kinds of musical abilities and that of certain kinds of spatial abilities are related. “Near transfer” theories concentrate on the cognitive similarities between music making and spatial processing. It is proposed that visuospatial intelligence is required for understanding musical notation and spatial relations such as those depicted on keyboards (Hetland, 2000). Thus it is argued that training in music will enhance visuospatial abilities of many different kinds. A meta-analysis of 15 studies of the transfer of musical training to spatial processing (Hetland, 2000) revealed that musical training that included the use of standard notation led to bigger improvements in spatial skills than did musical training that did not. However, even musical programs that did not use notation were found to yield a modest effect on spatial skills. All but two of the studies that used notation also used piano training, making it impossible to say whether either of these variables makes an independent contribution. In the present experiments, we also investigated the effect of musical training on spatial processing. We did not, however, predict general improvements in spatial abilities; instead we predicted a specific kind of spatial transfer, based on the particular spatial relations used in music reading and keyboard performance. Our prediction that pianists who had learned to read music and play keyboard instruments would possess a set of stimulus–response mappings was supported by the results of Experiment 2. Such a finding suggests that representation of musical notation, at least for pianists, is characterized by specific vertical-to-horizontal stimulus–response mappings. The acquisition of skill in music reading and playing thus involves the building of spatial representations that, once in place, extend outside a musical context.

REFERENCES

Cattell, J. M. (1886). The time it takes to see and name objects. *Mind*, 2, 63-85. Chan, A. S., Ho, Y. C., & Cheung, M. C. (1998). Music training improves verbal memory. *Nature*, 396, 128. Cohen, J. D., McClelland, J. L., & Dunbar, K. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, 97, 332-361. Glaser, M. O., & Glaser, W. R. (1982). Time course analysis of the

Stroop phenomenon. *Journal of Experimental Psychology: Human Perception & Performance*, 8, 875-894. Graziano, A. B., Peterson, M., & Shaw, G. L. (1999). Enhanced learning of proportional math through music training and spatial-temporal reasoning. *Neurological Research*, 21, 139-152. Hetland, L. (2000). Learning to make music enhances spatial reasoning. *Journal of Aesthetic Education*, 34, 179-238. Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus-response compatibility—A model and taxonomy. *Psychological Review*, 97, 253-270. MacLeod, C. M. (1991). Half a century of research on the Stroop effect. *Psychological Bulletin*, 109, 163-203. MacLeod, C. M. (1998). Training on integrated versus separated Stroop tasks: The progression of interference and facilitation. *Memory & Cognition*, 26, 201-211. MacLeod, C. M., & Dunbar, K. (1988). Training and Stroop-like interference: Evidence for a continuum of automaticity. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 14, 126-135. MacLeod, C. M., & MacDonald, P. A. (2000). Interdimensional interference in the Stroop effect: Uncovering the cognitive and neural anatomy of attention. *Trends in Cognitive Sciences*, 4, 383-391. Monsell, S., & Sternberg, S. (1981). Speech programming: A critical review, a new experimental approach and a model of the timing of rapid utterances: Part 1 (Unpublished technical memorandum). Murray Hill, NJ: AT&T Bell Laboratories. Morton, J., & Chambers, S. M. (1973). Selective attention to words and colours. *Quarterly Journal of Experimental Psychology*, 25, 387-397. Palmer, C., & Meyer, R. (2000). Conceptual and motor learning in music performance. *Psychological Science*, 11, 63-69. Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 55-85). Hillsdale, NJ: Erlbaum. Rauscher, F. H., & Shaw, G. L. (1998). Key components of the Mozart effect. *Perceptual & Motor Skills*, 86, 835-841. Rauscher, F. H., Shaw, G. L., & Ky, K. N. (1995). Listening to Mozart enhances spatial-temporal reasoning: Towards a neurophysiological basis. *Neuroscience Letters*, 185, 44-47. Rauscher, F. H., Shaw, G. L., Levine, L. J., Wright, E. L., Dennis, W. R., & Newcomb, R. L. (1997). Music training causes long-term enhancement of preschool children's spatial-temporal reasoning. *Neurological Research*, 19, 2-8. Schellenberg, E. G. (2001). Music and nonmusical abilities. In R. J. Zatorre & I. Peretz (Eds.), *The biological foundations of music* (*Annals of the New York Academy of Sciences*, Vol. 930, pp. 355-371). New York: New York Academy of Sciences. Shaw, G. (2000). *Keeping Mozart in mind*. San Diego: Academic Press. Sternberg, S., Monsell, S., Knoll, R. L., & Wright, C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G. E. Stelmach (Ed.), *Information processing in motor control and learning* (pp. 117-152). New York: Academic Press. Stewart, L., Henson, R., Kampe, K., Walsh, V., Turner, R., & Frith, U. (2003). Becoming a pianist: Brain changes associated with learning to read and play music. *NeuroImage*, 20, 71-83. Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643-662. Wolf, T. A. (1976). A cognitive model of sight-reading. *Journal of Psycholinguistic Research*, 5, 143-171. (Manuscript received April 24, 2002; revision accepted for publication February 14, 2003.)

Visually Induced Auditory Expectancy in Music Reading: A Behavioral and Electrophysiological Study Daniele Schön and Mireille Besson
Abstract & The general aim of this experiment was to investigate the processes involved in reading musical notation and to study the relationship between written music and its auditory representation. It was of main interest to determine whether musicians are able to develop expectancies for specific tonal or atonal auditory events based on visual score alone. Can musicians expect an “atonal” event or will it always sound odd? Moreover, it was of interest to determine whether the modulations in amplitude of a late positive component (P600) described in previous studies

are linked to a general mismatch detection process or to specific musical expectancies. Results showed clearly that musicians are able to expect tonal auditory endings based on visual information and are also able to do so for atonal endings, although to a smaller extent. Strong interactions seem to exist between visual and auditory musical codes and visual information seems to influence auditory processing as early as 100 msec. These results are directly relevant for the question of whether music reading is actually music perception. & INTRODUCTION In the study of reading music, as for reading words, an important question is to understand what types of representations are used. In reading music, the representations may be phonologic (“la”), graphemic (), motoric (second finger), and acoustic (sound). Indeed, musicians most often read a score and play it at the same time. For instance, a string player sight-reading a score needs to identify the signs on the score to produce the appropriate motor action, to check whether the played notes are well tuned, and, sometimes, even add a vibrato depending upon the musical importance of each note. This latter point of musical interpretation is of great importance, not only in musical praxis, but also in psychological theory. To interpret the piece musically, the musician needs to assign a “musical weight” to each note, thus computing a set of complex operations. These operations comprise the understanding and updating of the metric, rhythmic, melodic, and harmonic structures. Therefore, as noted by Sloboda (1976, 1978, 1984), music reading is not just a visuomotor coding task but rather belongs to the realm of music perception. It has been shown that, in a music-reading task, musicians can better memorize “good” (well structured) than “bad” musical sequences (Halpern & Bower, 1982). The fact that memory is sensitive to musical structure implies that some musical representations are at work while reading. However, this does not provide precise information about the characteristics of these representations. It may be the case that the representations built from music reading are based on complex relevant visual patterns, on the rules governing music perception in auditory modality, or on an explicit knowledge of musical structure (e.g., the knowledge that in B[major, there are 2 flats). Note that these possibilities are not mutually exclusive and it is likely that these different types of representations coexist and interact (Schoen & Besson, 2002a; Schoen, Semenza, & Denes, 2001). An interesting possibility is that musicians have, to a certain extent, an auditory-like representation of the written music, before they actually play it. Although the visuomotor coding can be sufficient to play the written notes, an anticipated auditory representation of the written music is important for an expressive performance that assigns a different “weight” to each single note, depending upon its melodic/harmonic and metric/rhythmic importance. Note also that the sole visuomotor strategy cannot be used with all instruments. For example, string instrument players need to check that the note played is indeed the written note, and this verification process most likely relies on a match/mismatch comparison with their anticipated auditory representation. Indeed, no keys or frets are present on the violin “fingerboard.” To test the hypothesis that musicians are able to use an auditory-like representation of written music, we investigated the relationship between reading and listening to music when both are coupled within the same experimental task. We recorded both behavioral (reaction INCM-CNRS, Marseille, France times [RTs] and error rates) and electrophysiological D 2005 Massachusetts Institute of Technology Journal of Cognitive Neuroscience 17:4, pp. 694–705 data (event-related brain potentials [ERPs]) to examine performance and the processes involved in music reading and listening as they unfold in time. The general hypothesis is that if written music induces musical auditory expectancy (i.e., not just motor expectancy), this may influence the way music is perceived. From studies in other domains, it is well known that a sensory system may, in some cases, degrade, enhance, or modify the perception of another system (Sekuler, Sekuler, & Lau, 1997; McGurk & McDonald, 1976). Musical notes can be categorized according to the degree to which they are expected. In a given context, some notes are highly

stable (Krumhansl, 1990) and plausible, whereas others are unstable and implausible, thus creating musical tension or surprise. Let us take an example. The familiar rhythm and melody snippet “Shave and a Haircut, Two Bits,” whose ending is so irresistible to Roger Rabbit, is represented in Figure 1A. If we play the same tune again, raising the last note by half a tone (Figure 1B), the ending will be perceived as highly implausible and surprising. The issue addressed in this article is to determine whether knowing in advance by means of a written score (Figure 1B) that such an odd note is going to be played does influence the expectancy for this note. In other words, is the expectancy of an auditory event in a given musical context influenced by visual information? If yes, at which stage of processing? Musicians were asked to judge whether the last note of a 5-note auditory musical sequence matched or mismatched the information provided on a score presented on a computer screen (Figure 2). There were two types of visual stimuli (all comprising 5 notes): In one type, the last written note was stable, whereas in the other type, the last note was unstable with respect to the previous tonal context.¹ The sequence of auditory events always matched the sequence of written notes, except for the final notes that could either match or mismatch with the last visual stable or unstable notes. In the visual–auditory mismatching conditions, two types of mismatches were introduced: Although the played note was always different from the one written on the score (i.e., mismatch), it could be either plausible (2 and 5) or implausible (3 and 6) with respect to the previous tonal context. To summarize, the experimental design comprised 6 experimental conditions with two types of visual sequences (ending with stable or unstable notes relative to the previous context), each being associated to three types of auditory sequences (ending with notes that match the visual sequence, that mismatch but are still musically plausible, and that mismatch and are musically implausible). The first aim was to determine whether musicians would be able to anticipate unstable auditory events based on visual information? If so, the two matching conditions (1 and 4) should be processed similarly, that is, independently of whether the final auditory event matching the score is stable (henceforth, stable match, Condition 1) or unstable (henceforth, unstable match, Condition 4). By contrast, if musicians are unable to anticipate unstable events, comparison of the stable and unstable matching conditions should show differences in behavioral data (longer RTs and more errors) and in the electrophysiological components (larger amplitude and longer latency, see below) described in the literature as correlated with the violation of musical expectancy (e.g., Schön & Besson, 2002b; Koelsch, Gunter, Friederici, & Schroger, 2000; Besson & Macar, 1987). The second aim was to further track the functional significance of the late positive component (LPC, or P600), peaking between 300 and 600 msec, that, in a Figure 1. Example of expected/stable (A) and unexpected/unstable endings (B). A is the tonic, the most expected note in this musical context, and B is the minor second, completely out of context. Figure 2. Illustration of the experimental conditions. Schön and Besson 695 melody, are associated to incongruous/unexpected endings (e.g., Besson & Macar, 1987). The amplitude and the latency of these late positivities have been shown to be sensitive to musical expertise, to the familiarity of the musical piece, and to the degree of incongruity (Besson & Fàlta, 1995). However, in previous studies (e.g., Regnault, Bigand, & Besson, 2001; Patel, Gibson, Ratner, Besson, & Holcomb, 1998; Besson & Fàlta, 1995; Besson & Macar, 1987), the congruous stimulus was always a match, whereas the incongruous stimulus was always a mismatch, because the notes/chords were not appropriate in the musical context. Moreover, insofar as the task, in most cases, was to decide whether the note/ chord was expected or not within the previous context, the mismatching aspect of the incongruous notes/ chords was clearly task-relevant. Therefore, the extent to which the occurrence of the LPCs reflects a general mismatch process or a more specific musical (tonal) mismatch process remains an open question. In the present experiment, the use of two types of auditory mismatches (plausible and implausible) allowed us to address this question. If the amplitude of the LPC is mainly

sensitive to a general mismatching process, auditory mismatching events should be processed similarly, regardless of whether they are plausible or implausible within the previous tonal context. By contrast, if this component is sensitive to more specific musical expectancy processes, governed by the rules of musical “syntax,” one should find differences between the two types of mismatching conditions, with the implausible mismatches showing larger LPCs than the plausible mismatches. The third aim was to study the relationship between the expectancy created by the visual score (stable vs. unstable endings) and the tonal dimension of the auditory stimuli in the 2 mismatching conditions (plausible vs. implausible). Indeed, musicians may perceive the same event (e.g., plausible or implausible mismatch) differently when they expect a stable ending from when they expect an unstable ending, based on visual information.

RESULTS Behavioral Results Error Rate A 2-way repeated measures ANOVA was carried out, including visual expectancy (stable/unstable) and auditory match/mismatch dimensions as factors (Figure 3). The main effect of visual expectancy was significant, $F(1,13) = 14.2$, $p = .002$, with more errors in the unstable-ending conditions (9.3% vs. 5.6%). Moreover, the visual expectancy by auditory match/mismatch interaction was also significant, $F(2,26) = 9.1$, $p = .001$. Post hoc comparisons showed that only the plausible mismatch in the visual unstable-ending condition was significantly different from all the other conditions (always $p < .008$) that did not differ from one another (always $p > .47$).

Reaction Times Again, a 2-way repeated measures ANOVA was carried out, including the same factors as above (Figure 3). The main effect of visual expectancy was significant, $F(1,13) = 4.7$, $p = .049$, with slower RTs in the unstable-ending conditions (682 vs. 639 msec). The auditory match/ mismatch main effect was also significant, $F(2,26) = 12.8$, $p = .0001$, with slower RTs to mismatching than matching auditory events. Moreover, the visual expectancy by auditory match/mismatch interaction was also significant, $F(2,26) = 3.7$, $p = .04$. Two further analyses were conducted: *t* tests revealed that RTs in the matching conditions were faster with stable than unstable visual endings (573 vs. 632 msec, $p = .04$). By contrast, results of a 2-way repeated measures ANOVA, including visual expectancy (stable/unstable) and the 2 mismatching conditions only (plausible/implausible mismatch) as factors, showed no main effect of visual expectancy ($p > .05$). However, plausible mismatches were associated with longer RTs than implausible mismatches (709 vs. 679 msec), $F(1,13) = 17.1$, $p = .001$, especially when the visual endings were unstable [visual expectancy by mismatch type interaction: $F(1,13) = 4.6$, $p = .05$].

Discussion Overall, accuracy and RTs showed a consistent pattern of results (i.e., no speed–accuracy trade-off). First, the matching conditions were performed faster and with fewer errors overall than the mismatching conditions. Second, although the 2 matching conditions did not differ in accuracy as a function of the stability of the musical sequence, RTs were shorter to stable than to unstable endings. Finally, the differences in accuracy and RTs between plausible and implausible mismatches were only significant when the visual endings were unstable. The finding that RTs were shorter and accuracy higher for matching than mismatching auditory events is in line with many results in the literature. In the present experiment, a target note with a specific pitch was expected based on visual information. Thus, although participants could anticipate matching targets, and therefore be ready to respond, they could not anticipate mismatching targets, thus explaining why RTs were slower in this latter case. Most importantly, this was true for both visual endings, showing that, upon visual information, musicians were able not only to anticipate stable, but also unstable endings. Such a tentative conclusion needs, however, to be further considered in light of the results in the matching conditions (1 vs. 4). Indeed, RTs were shorter for stable than unstable matching endings. Two possible interpretations may account for this result. It may be that it is more difficult to anticipate unstable than stable matching endings. Alternatively, musicians may be able to anticipate visual unstable endings

to the same degree as stable endings, but the pitch of the unstable endings may remain somewhat surprising (it is a “wrong” note after all), thus causing a delay in the response. This latter explanation builds on the hypothesis that a wrong note remains incongruous even when one knows it in advance (see also general discussion). Finally, the finding that plausible mismatches were associated with more errors and slower RTs than implausible mismatches, when the visual endings were unstable, may reflect some inhibition process. Indeed, within unstable-ending sequences, a plausible mismatch provides a better musical closure than the matching ending. Therefore, musicians may need to inhibit a positive response (i.e., match) when presented with a plausible mismatch, thus increasing RTs and error rates.

Event-related Brain Potential Results As can be seen in Figure 4, the visuoauditory music task used in this experiment elicited several ERP components in the different experimental conditions. A N1 component is clearly evident at all electrode sites, peaking around 110 msec. A rapid positive (190 msec)–negative (220 msec) complex (P2–N2) is then elicited, with larger negative components in the mismatching conditions. Although a positivity then develops at posterior sites (P3), a negative-going component is present at frontal sites (N5). To analyze in detail how these components were modulated by the independent variables manipulated in this experiment, we first computed general Figure 3. Error rates (in percentage) and RTs (in milliseconds). M = match; MMP = plausible mismatch; MMI = implausible mismatch.

Schoen and Besson 697 analyses including all conditions and then separate analyses for the matching and mismatching conditions. In line with the general aims of this experiment, and in order not to complicate further the presentation of the results, we describe in detail only the results of separate analyses for the matching and mismatching conditions. For the global analyses, it is sufficient to note that they showed a main effect of visual expectancy in the 350- to 800-msec latency range, $F(1,13) = 6.9$, $p = .02$, a main effect of auditory match/mismatching in the 100- to 300-msec latency range, $F(1,13) = 9.0$, $p = .002$, then interacting with the anterior/posterior dimension in the 300- to 600-msec window, $F(2,26) = 30.5$, $p < .0001$, and finally, a visual by auditory match/mismatch by anterior/posterior interaction in the 250- to 400-msec latency range, $F(2,26) = 9.6$, $p = .007$.

Matching Events An ANOVA was computed with visual expectancy (2), hemisphere (2), anterior/posterior (2), and electrodes (3) as factors. A visual expectancy by anterior/posterior Figure 4. Illustration of the variations in brain electrical activity time-locked to the final note when expecting a stable ending (top, A) or an unstable ending (bottom, B) for matching (blue trace) or plausible (green trace) and implausible mismatching (red trace) conditions. Each trace represents an average of electrophysiological data recorded from 14 musicians. Although EEG was recorded from 28 electrodes, only the most representative clusters of electrodes (12 electrodes) were analyzed using ANOVAs and selected traces from 9 electrodes are presented. On this figure, as on the following ones, the amplitude (in microvolts) is plotted in ordinate (negative up) and the time (in milliseconds) is in abscissa.

698 *Journal of Cognitive Neuroscience* Volume 17, Number 4 interaction was found significant in both the 200- to 350-msec and 500- to 700-msec latency ranges [$F(1,13) = 8.0$, $p = .01$, and $F(1,13) = 5.7$, $p = .03$, respectively]. Precise analysis of the time course of these effects (50-msec windows) revealed that the amplitude of early (N2: 200–300 msec latency range) and late (N5: 500- to 700-msec latency range) anteriorly distributed negative components was larger for unstable than stable endings (Figure 5). Moreover, a posteriorly distributed positive component (P3: 250- to 350-msec latency range) was also larger in unstable than stable matching endings (Figure 5).

Mismatching Events An ANOVA was computed with visual expectancy (2), mismatch type (2), hemisphere (2), anterior/posterior (2), and electrodes (3) as factors. The main effect of visual expectancy was significant between 300 and 450 msec, $F(1,13) = 10.7$, $p = .006$, and interacted with the anterior/posterior dimension, $F(1,13) = 19.6$, $p = .0007$: The amplitude of posteriorly distributed positivities was larger for visual stable than unstable

endings. The main effect of mismatch type was also significant in the 100- to 150-msec (N1), 200- to 350-msec, and 400- to 700-msec latency bands [$F(1,13) = 9.4, p = .009$; $F(1,13) = 10.2, p = .007$; and $F(1,13) = 14.2, p = .002$, respectively]: The N1, the P3a, and a later positive component were larger in implausible than in plausible mismatches (see Figure 6). Interestingly, the visual expectancy by auditory mismatch interaction was significant between 450 and 550 msec, $F(1,13) = 4.6, p = .05$: The positivity elicited by the implausible mismatches was larger when the visual stimuli were musically stable endings (see Figure 7).

Discussion Overall, results showed that the mismatching conditions differed from the matching conditions both when the visual stimuli had stable and unstable endings. Although this speaks in favor of the capacity of musicians to anticipate or prepare themselves for unstable endings, detailed analyses of our results show that such a strong conclusion cannot be arrived at without precautions. In the following discussion, we will first consider the results in the auditory matching conditions and then those in the auditory mismatching conditions.

Matching. When the auditory stimuli match the note expected based on the visual score (matching conditions), the amplitude of early and late negative components (N2 and N5), as well as the amplitude of a positive component (P3), is larger in the unstable than in the stable-ending condition. Thus, the answer to the first question asked in the Introduction (i.e., “Are musicians able to anticipate an unstable matching auditory event based on the visual score?”) is rather negative because unstable matching events seem to be processed differently from stable matching events. Figure 5. Illustration of the variations in brain electrical activity time-locked to the final note when it matches either the visual stable endings (solid line; Condition 1) or the visual unstable endings (dashed line; Condition 4). A larger late positivity (P3) is visible when the subjects expect an unstable ending. Schön and Besson 699 In other words, the expectancy that musicians develop based on the visual score is not as strong for unstable than for stable-ending events. As mentioned previously, musicians can anticipate a specific expected (i.e., stable) event, but cannot fully anticipate unexpected (i.e., unstable) events, may be because of the partly automatic manner in which the music structure is parsed (Besson & Schön, 2001; Jackendoff, 1991). The larger positive component (P3) to unstable matching endings seems to indicate that, independently of the fact that musicians are “expecting” an unstable (i.e., odd) ending based on the visual stimulus, the auditory ending remains somewhat unexpected. Although our hypothesis was mainly based upon potential modulations in the amplitude of LPCs, results showed that the amplitude of an early negative component (150- to 250-msec latency band) was also sensitive to stable versus unstable tonal endings. Interestingly, the analysis of the scalp distribution of this early negativity revealed a right frontal lateralization (Figure 8) that is reminiscent of the right anterior temporal negativity (RATN) described by Patel et al. (1998) and of the early right anterior negativity (ERA) described by Koelsch and colleagues (Koelsch, Schroger, & Gunter, 2002; Koelsch, Gunter, Schroger, et al., 2001; Koelsch et al., 2000). More precisely, the negative component described here Figure 6. Zoom of the variations in brain electrical activity time-locked to the final note for both plausible mismatches (solid line, average of Conditions 2 and 5) and both implausible mismatches (dashed line, average of Conditions 3 and 6). The N1 and P3 are larger in implausible endings than in plausible endings. Figure 7. Illustration of the variations in brain electrical activity time-locked to the implausible mismatches in visual stable endings (solid line, Condition 3) and in visual unstable endings (dashed line, Condition 6). A larger late positivity is visible when the score contains a stable ending.

700 Journal of Cognitive Neuroscience Volume 17, Number 4 resembles the ERAN more closely than the RATN, because the latter peaks around 350 msec. The ERAN has been shown to be elicited by deviant chords relative to a harmonic context, maximal around 150–250 msec. It has been taken to reflect the violation of the listener musical expectancy and, more generally, music-syntactic processing (Maess, Koelsch, Gunter, & Friederici, 2001). So

far, this component has been described in studies using an auditory presentation of chords, but not when melodies were presented. It may be that a harmonic context creates stronger expectancies than a purely melodic context, because the rules governing harmony are less flexible than those governing melody. In our study, however, an ERAN may have developed because melodic expectancies were possibly enhanced by the presentation of the visual score. Note also that in previous studies using simple melodies, such early differences, which are usually small in amplitude, were possibly masked by the occurrence of larger LPCs. As the authors were generally mainly interested in the LPCs, such early differences may have gone unnoticed. Finally, the amplitude of a late negative component (N5) was also larger in the unstable matching condition than in the stable matching conditions. This negativity, peaking around 500 msec, resembles the N5 described by Koelsch et al. (2000, 2002). The authors proposed that this N5, which was larger in Neapolitan chords compared with tonic chords, reflects musical integration processes. Indeed, Neapolitan chords are more difficult to integrate into the preceding harmonic context than are tonic chords. Thus, according to this interpretation, unstable matching endings in the present experiment could be analogous to the Neapolitan chords, in that they are more difficult to integrate. As hypothesized for the ERAN, the fact that no such a late negative component was described in previous studies using melodies rather than chord sequences may be because of an enhancement of melodic expectancy by the visual score (see also Gunter, Schmidt, & Besson, 2003, for a discussion of the ERAN in music reading). Further research will be needed to directly compare the late negativities elicited by chords and melodies.

Mismatch. When the auditory stimulus mismatches the note expected based on the visual score (mismatching conditions), several interesting results emerge. First, independently of whether musicians are anticipating stable- or unstable-ending notes, implausible auditory mismatches are associated with larger N1, P3, and LPCs than plausible auditory mismatches (Figure 6). Thus, these results are in line with previous results in the ERPs and music literature (Hashimoto, Hirata, & Kuriki, 2000; Besson & Macar, 1987; Schön, Magne, & Besson, 2004) when purely auditory musical stimuli are used. However, in these studies, expected (matching) events were always compared with unexpected (mismatching) events. Therefore, it was not possible to determine whether the effects reported on the N1, P3, and late positivities were linked to a general mismatch process or to the violation of specific musical expectancies. By contrast, in the present study, the comparison is between 2 mismatching events: one that is implausible within the melodic context and the other that is highly plausible. Therefore, the fact that we found differential effects, with larger differences for implausible than plausible auditory mismatches, clearly show that these different ERP components are sensitive to specific musical expectancies. Note, furthermore, that in previous studies, the number of trials included in the averages for congruous and incongruous trials were often different, thus possibly influencing the amplitude of the early components such as the N1. It should be mentioned that the N1 results differ from the results of a previous experiment (Regnault et al., 2001), in which the N1 component was found to be larger in consonant than in dissonant chords. Insofar as notes were used in the present experiment, although chords were presented in the study of Regnault et al. (2001), it may be that these differences reflect the fact that harmonic processing differs to a certain extent from Figure 8. Topographic maps computed as an integration of mean amplitude values across time in the 170- to 210-msec time windows, from the difference waves (unstable–stable match).

Schön and Besson 701 melodic processing. Further research is needed to directly compare single notes and chords and to understand the effect of the hierarchic organization of tonal music on the different musical processing stages. Second, and very interestingly, the amplitude of the late positivity associated to implausible auditory mismatches was larger when musicians expected stable endings than when they expected unstable endings (Figure 7). This result supports the view that visual stable endings induce a

strong musical expectancy for a plausible event, thereby increasing the auditory mismatch effect of the implausible ending. By contrast, unstable endings create expectancy for an implausible event, thereby decreasing the auditory mismatch effect of the implausible ending. Third, an early negative component, peaking around 200 msec, was elicited by both plausible and implausible mismatches, although it was somewhat larger for plausible mismatches (especially in the unstable-ending condition, see Figure 4). Because this negative component is followed by a clear positivity at frontal sites, we tend to interpret this succession of negative positive components as an N2–P3 complex, which would develop in response to surprising unexpected events. However, the finding that the amplitude of this N2-like component is overall larger in plausible than implausible mismatches seems to indicate that surprise is not the only factor influencing the results. As proposed in the discussion of the RT results, it might be that some inhibition processes are at play when plausible mismatches are presented at the end of unstable sequences because they match the preceding tonal context. Following this line of reasoning, the increased N2 amplitude in plausible mismatches (most evident in the unstable conditions) would reflect some inhibition linked with the suppression of a “match response.” This interpretation is in line with the increased N2 amplitude to no-go responses reported in the literature (Simson, Vaughan, & Ritter, 1977). Finally, it is interesting to note that the distribution of this negativity presents some differences with respect to the one found for the ERAN. In particular, although the ERAN cannot be seen at posterior sites and has a quite localized right frontal distribution, the negativity to mismatching endings can clearly be seen at fronto-central and parietal electrodes. It might be that the greater saliency of the mismatch and inhibition effects hides the smaller effect sensitive to musical stability (syntax).

GENERAL DISCUSSION

The general aim of this experiment was to investigate the processes involved in reading musical notation and, more precisely, to study the relationship between written music and its auditory representation. It was of main interest to determine whether musicians can develop expectancies for specific stable or unstable auditory events based on the visual score alone. Taken together, results showed that this seems to be the case, at least to some extent. Indeed, musicians are clearly able to anticipate stable endings based on visual information and are also able to do so for unstable endings, although to a smaller extent. Moreover, error rates, RTs, and electrophysiological data showed that the auditory mismatch effects found in the present experiment were clearly modulated by the musical expectations built from the score. Thus, strong interactions seem to exist between visual and auditory musical codes, in that the representations built from visual stimuli influence the way auditory musical sequences are perceived. Indeed, although N1 amplitude was significantly different between plausible and implausible mismatches (i.e., between notes that differed in terms of musical expectancy but that could not be, being mismatches, anticipated based on visual information, see Figure 6), such an N1 difference was not significant between stable and unstable matches (i.e., between notes that differed in terms of musical expectancy in the same way as plausible and implausible mismatches, but that, by contrast, could be anticipated based on visual information, see Figure 5). Therefore, we interpret this lack of difference on the N1 as an influence of visual processing on auditory processing, because, what distinguishes the matches from the mismatches is that the matches are expected based on visual information. In other words, being able to anticipate a precise target note based on visual information seems to influence auditory processing as early as 100 msec. In support of the hypothesis of top-down influences on the N1, some studies have shown that the auditory cortex (where the N1 would be generated, Pantev et al., 1995; Liegeois-Chauvel, Musolino, Badier, Marquis, & Chauvel, 1994) is sometimes involved in imagery tasks. For example, the auditory association areas (left) are active during a simple auditory imagery task in which musicians had to imagine the sound of a single note presented visually (Schürmann et al., 2002). In

addition, imagery for familiar tunes showed activation in the right auditory association cortex of subjects imaging the continuation of a tune cued by its first few notes (Halpern & Zatorre, 1999). In the present experiments, insofar as subjects had to imagine in advance the sound of the last note of a visual sequence, this may have reduced the effect of syntactic incongruity. However, note that differences on later components (ERAN and P3, N5) are still visible (see Figure 5). It may be the case that although the N1 difference is a rather small difference, and thus when attenuated it is no longer significant, the differences on the ERAN and on the P300 remain significant. In other words, although there is an early effect of visual information on auditory processing, the effect of syntactic incongruity is only reduced and does not disappear. This is in line with the hypothesis proposed by Jackendoff (1991) that a processor uses the rules of music to analyze musical surface structure and 702 *Journal of Cognitive Neuroscience* Volume 17, Number 4 operates in a rather autonomous manner, with a limited access to musical memory: “No matter how well one knows a piece, expectation, suspense, satisfaction, and surprise still occurs within the parser” (p. 228). Such activity of deriving in real time the detailed and complex structure of music, with its points of tension and instability, can partly explain why a piece of music can create strong emotions despite repeated hearings. Further experiments using fMRI will help in defining to what extent and in which areas the visual presentation of notes can modulate their auditory perception. From the present findings, this modulation should be present in auditory areas and also in the (right) prefrontal regions considered to be important for musical syntax processing (Maess et al., 2001) and pitch retention. Another interesting result of the present study concerns the functional interpretation of the P300 (P600). Several studies, in the literature on music processing and ERPs, have demonstrated that, in the auditory modality, the amplitude of LPCs, also called the P300 or the P600, is larger in unexpected than expected endings. Thus, the aim was to determine whether the modulations in amplitude of this LPC are because of the call into play of a general mismatch detection process or are also linked with specific musical expectancies (stability/plausibility). In the present experiment, insofar as we used two types of mismatching conditions, plausible and implausible, we hypothesized that the differences between the two would not be due to a mismatching process, but rather to musical expectancy. Results clearly showed that the amplitude and latency of the LPC were sensitive to specific musical expectations. However, it is interesting to note that, overall, we found smaller differences between conditions than expected based on previous findings (e.g., Besson & Faita, 1995). Therefore, it may be that in previous studies on music processing, the amplitude (and latency) of the positivity were modulated by several factors such as a general mismatch detection process, task-relevant expectations (probability of occurrence), and expectations based on the musical structure. It is therefore important, in future studies on musical expectancy, to consider that these different factors may be at work simultaneously.

METHODS Participants Eighteen volunteer musicians were tested in the experiment. However, because of artifacts in the ERP data of 4 participants, only 14 (years of training: $M = 17$; range = 11–30) were retained for analyses. All were right-handed, neurologically normal, had normal or corrected-tonormal vision, normal audition, and were native French speakers (age: $M = 25$ years, 6 women). None of them had absolute pitch. All participants were paid for their participation to the experiment. Stimuli Three hundred twenty pairs of visual/auditory stimuli were used in the experiment. All stimuli comprised 5 notes of equal duration (quarter notes of 750 msec). White visual stimuli on a dark background were presented at the center of a computer screen (placed 60 cm in front of the participants). Auditory stimuli were generated using the piano sound produced by a general MIDI (Korg, X5DR) under computerized control. The visual 5 notes were presented all at once on the screen, and the stimulus onset was synchronized with the first note of the 5-note auditory sequence. The visual stimulus remained on the screen for the entire duration of the auditory

stimulus (i.e., 3750 msec). A row of X was presented 1250 msec after auditory final note offset for 2000 msec to allow blinking. A further 1000 msec interval followed the row of X before the next trial (see Figure 9). Visual stimuli either had stable or unstable endings (see Figure 2). To increase variability, several keys were used (major keys: C, F, B \flat , E \flat , G, D, A, B; minor keys: D, F, F \sharp , G, A, B). The 40 stable stimuli were composed of 5 notes within the same key. The last note was always an in-key note (tonic, supertonic, mediant, subdominant, dominant, submediant, or leading tone). The 40 unstable stimuli were similar to the stable-ending stimuli in that they were built by transposing them (the original stable-ending stimulus could be transposed by a minor or major third, by a fourth and by a fifth, above or below). After transposition, the last note was modified (by a chromatic semitone or by a tone) to become an unstable ending (mostly out-of-key tones, very rarely an unstable degree of the key). Auditory stimuli were of 3 types. The final note of the auditory stimulus either matched (Conditions 1 and 4) or mismatched with the last note of the visual stimulus. Mismatches were either plausible or implausible endings. Whereas in the plausible mismatch conditions the last (auditory) note was always a diatonic note (Conditions 2 and 5), in the implausible mismatch conditions, Figure 9. Illustration of the experimental design. Schön and Besson 703 the last note was most often out of key and, in some rare cases, corresponded to an unstable degree of the key (Conditions 3 and 6). All mismatches did, however, preserve the contour of the visual sequence, and the change was only in the size of the interval between the fourth and the fifth note. To test for the psychological validity of the stimuli, we asked 6 musicians, naïve with respect to the aims of the experiment, to judge on a 5-point scale how much the last note they heard was stable/expected. Results showed that ending notes in Conditions 1, 2, and 5 (stable/plausible) were judged as more stable than ending notes in Conditions 3, 4, and 6 (unstable/implausible), $F = 145.35$, $p < .0001$. By contrast, comparisons between conditions to test “stability” within the stable/plausible conditions (e.g., comparing visual stable match with plausible/stable mismatch, i.e., Condition 1 vs. Condition 2) and unstable/implausible conditions (comparing visual unstable match with implausible/unstable mismatch, i.e., Condition 4 vs. Condition 6) showed no significant differences. To summarize, 6 conditions were used in the experiment (see Figure 2): visual stable endings with (1) auditory (stable) match, (2) auditory plausible mismatch, (3) auditory implausible mismatch; visual unstable endings with (4) auditory (unstable) match, (5) auditory plausible mismatch, and (6) auditory implausible mismatch. Each condition comprised 40 items. Moreover, to balance the proportion of matches and mismatches, 40 fillers were added in each of the two matching conditions (expected and unexpected). Stimuli were presented in a pseudorandom order. Procedure Participants, comfortably seated in a Faraday box, read a 5-note score while a computer simultaneously played the notes. They were instructed to decide, as quickly and as accurately as possible, whether the final auditory note matched or mismatched with the final visual note on the score. The association between hand side (left or right) and responses (yes or no) was balanced across participants. Data Acquisition and Analysis Electroencephalogram (EEG) was recorded for 5200 msec starting 200 msec before the onset of the stimuli (baseline), from 28 scalp electrodes located at standard left and right hemisphere positions over frontal, central, parietal, occipital, and temporal areas (International 10/20 system sites: Fz, Cz, Pz, Oz, Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6, Fc5, Fc1, Fc2, Fc6, Cp5, Cp1, Cp2, and Cp6). These recording sites plus an electrode placed on the right mastoid were referenced to the left mastoid electrode; the data were then rereferenced offline to the algebraic average of the left and right mastoids. Impedances of these electrodes never exceeded 3 k Ω . The horizontal electrooculogram (EOG) was recorded from bipolar electrodes placed 1 cm to the left and right of the external canthi, and the vertical EOG was recorded from an electrode beneath the right eye, referenced to the left mastoid, to detect blinks and vertical eye movements. Trials

containing ocular artifacts, movement artifacts, or amplifier saturation were excluded from the averaged ERP waveforms. The EEG and EOG were amplified by an SA Instrumentation amplifier with a band-pass of 0.01–30 Hz and were digitized at 250 Hz. ERP data were analyzed by computing the mean amplitude, starting 200 msec before the last note was played and ending 2000 msec after. Only trials in which performance was correct were taken into account in the analyses. Analysis of variance (ANOVA) was used for statistical evaluation. To test the distribution of the effects, 4 regions of interest (ROIs) were selected for statistical evaluation: left anterior (F3, Fc1, and Fc5), left posterior (Cp1, Cp5, and P3), right anterior (F4, Fc2, and Fc6), right posterior (Cp2, Cp6, and P4). Note that ANOVAs that included the midline electrodes were also performed. However, because no major differences were found between these two types of analyses, we only report those including the ROIs. We used latency windows of 50 msec in the 0- to 500-msec range and windows of 100 msec in the 500- to 900-msec range. All *p* values reported below were adjusted with the Greenhouse–Geisser epsilon correction for nonsphericity, when appropriate. Tukey tests were used in post hoc comparisons unless specified otherwise.

Acknowledgments This work was completed while Daniele Schön was a postdoctoral student. This work was first supported by a grant from the International Foundation for Music Research to Mireille Besson (RPA 194) and later by a grant from the Human Frontier Science Program (RG P0053/2002). We thank Andras Semjen and 2 anonymous reviewers for helpful comments on a previous version of the manuscript. Reprint requests should be sent to Daniele Schön, INCMCNRS, 31 Ch J. Aiguier, 13420 Marseille, France, or via e-mail: schon@incm.cnrs-mrs.fr.

Note 1. Note that to distinguish visually presented notes from auditory presented notes, we will use the terms stable/ unstable for visual presentation and plausible/implausible for auditory presentation. However, the meaning of these terms is the same.

REFERENCES Besson, M., & Fariña, F. (1995). An event-related potential (ERP) study of musical expectancy: Comparisons of musicians with non-musicians. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 1278–1296. Besson, M., & Macar, F. (1987). An event-related potential analysis of incongruity in music and other non-linguistic contexts. *Psychophysiology*, 24, 14–25. Besson, M., & Schön, D. (2001). Comparison between language and music. *Annals of the New York Academy of Science*, 930, 232–258. Gunter, T. C., Schmidt, B. H., & Besson, M. (2003). Let's face the music: A behavioral and electrophysiological exploration of score reading. *Psychophysiology*, 40, 742–751. Halpern, A., & Bower, G. (1982). Musical expertise and melodic structure in memory for musical notation. *American Journal of Psychology*, 95, 31–50. Halpern, A., & Zatorre, R. J. (1999). When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cerebral Cortex*, 9, 697–704. Hashimoto, T., Hirata, Y., & Kuriki, S. (2000). Auditory cortex responds in 100 ms to incongruity of melody. *NeuroReport*, 11, 2799–2801. Jackendoff, R. (1991). Musical parsing and musical affect. *Music Perception*, 9, 199–230. Koelsch, S., Gunter, T., Friederici, A. D., & Schroger, E. (2000). Brain indices of music processing: “Nonmusicians” are musical. *Journal of Cognitive Neuroscience*, 12, 520–541. Koelsch, S., Gunter, T. C., Schroger, E., Tervaniemi, M., Sammler, D., & Friederici, A. D. (2001). Differentiating ERAN and MMN: An ERP study. *NeuroReport*, 12, 1385–1389. Koelsch, S., Schroger, E., & Gunter, T. C. (2002). Music matters: Preattentive musicality of the human brain. *Psychophysiology*, 39, 38–48. Krumhansl, C. (1990). *Cognitive foundation of musical pitch*. New York: Oxford University Press. Liegeois-Chauvel, C., Musolino, A., Badier, J. M., Marquis, P., & Chauvel, P. (1994). Evoked potentials recorded from the auditory cortex in man: Evaluation and topography of the middle latency components. *Electroencephalography and Clinical Neurophysiology*, 92, 204–214. Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: An MEG study. *Nature*

Neuroscience, 4, 540–545. McGurk, H., & McDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264, 746–748. Pantev, C., Bertrand, O., Eulitz, C., Verkindt, C., Hampson, S., Schuierer, G., & Elbert, T. (1995). Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. *Electroencephalography and Clinical Neurophysiology*, 94, 26–40. Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience*, 10, 717–733. Regnault, P., Bigand, E., & Besson, M. (2001). Different brain mechanisms mediate sensitivity to sensory consonance and harmonic context: Evidence from auditory event related brain potentials. *Journal of Cognitive Neuroscience*, 13, 241–255. Schoen, D., & Besson, M. (2002a). Processing pitch and duration in music reading: A RTs and ERPs study. *Neuropsychologia*, 40, 868–878. Schoen, D., & Besson, M. (2002b). Musica, maestro! In L. Baekman & C. von Hofsten (Eds.), *Psychology at the turn of the millennium* (pp. 119–152). Hove, UK: Psychology Press. Schoen, D., Magne, C., & Besson, M. (2004). The music of speech. ARTs and ERPs study of pitch contour in prosody and music. *Psychophysiology*, 41, 341–349. Schoen, D., Semenza, C., & Denes, G. (2001). Naming of musical notes: A selective deficit in one musical clef. *Cortex*, 37, 407–421. Schürmann, M., Raij, T., Fujiki, N., & Hari, R. (2002). Mind's ear in a musician: Where and when in the brain. *Neuroimage*, 16, 434–440. Sekuler, R., Sekuler, A. B., & Lau, R. (1997). Sound alters visual motion perception. *Nature*, 385, 308. Simson, R., Vaughan, H. G., & Ritter, W. (1977). The scalp topography of potentials in auditory and visual Go/NoGo tasks. *Electroencephalography and Clinical Neurophysiology*, 43, 864–875. Sloboda, J. A. (1976). Visual perception of musical notation: Registering pitch symbols in memory. *Quarterly Journal of Experimental Psychology*, 28, 1–16. Sloboda, J. A. (1978). Psychology of music reading. *Psychology of Music*, 6, 3–20. Sloboda, J. A. (1984). Experimental studies of musical reading: A review. *Music Perception*, 2, 222–236. Schoen and Besson 705