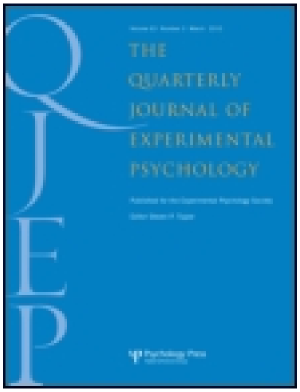

The Quarterly Journal of Experimental Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pqje20

Action-perception coupling in pianists: Learned mappings or spatial musical association of response codes (SMARC) effect?

Lauren Stewart [a], Rinus G. Verdonschot [b c d], Patrick Nasralla [e] & Jennifer Lanipekun [f]

[a] Department of Psychology, Goldsmiths, University of London, London, UK
[b] Leiden University Centre for Linguistics, Leiden, The Netherlands
[c] Leiden University Institute for Psychological Research, Leiden, The Netherlands
[d] Leiden Institute for Brain and Cognition, Leiden, The Netherlands
[e] Department of Anatomy, University College London, London, UK
[f] Music Department, Institute of Education, London, UK

Accepted author version posted online: 27 Apr 2012.Version of record first published: 19 Jun 2012.

To cite this article: Lauren Stewart , Rinus G. Verdonschot , Patrick Nasralla & Jennifer Lanipekun (2013): Action-perception coupling in pianists: Learned mappings or spatial musical association of response codes (SMARC) effect?, The Quarterly Journal of Experimental Psychology, 66:1, 37-50

To link to this article: http://dx.doi.org/10.1080/17470218.2012.687385

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 2013
Vol. 66, No. 1, 37–50, http://dx.doi.org/10.1080/17470218.2012.687385

Action–perception coupling in pianists: Learned mappings or spatial musical association of response codes (SMARC) effect?

Lauren Stewart[1], Rinus G. Verdonschot[2,3,4], Patrick Nasralla[5], and Jennifer Lanipekun[6]

[1]Department of Psychology, Goldsmiths, University of London, London, UK
[2]Leiden University Centre for Linguistics, Leiden, The Netherlands
[3]Leiden University Institute for Psychological Research, Leiden, The Netherlands
[4]Leiden Institute for Brain and Cognition, Leiden, The Netherlands
[5]Department of Anatomy, University College London, London, UK
[6]Music Department, Institute of Education, London, UK

The principle of common coding suggests that a joint representation is formed when actions are repeatedly paired with a specific perceptual event. Musicians are occupationally specialized with regard to the coupling between actions and their auditory effects. In the present study, we employed a novel paradigm to demonstrate automatic action–effect associations in pianists. Pianists and nonmusicians pressed keys according to aurally presented number sequences. Numbers were presented at pitches that were neutral, congruent, or incongruent with respect to pitches that would normally be produced by such actions. Response time differences were seen between congruent and incongruent sequences in pianists alone. A second experiment was conducted to determine whether these effects could be attributed to the existence of previously documented spatial/pitch compatibility effects. In a "stretched" version of the task, the pitch distance over which the numbers were presented was enlarged to a range that could not be produced by the hand span used in Experiment 1. The finding of a larger response time difference between congruent and incongruent trials in the original, standard, version compared with the stretched version, in pianists, but not in nonmusicians, indicates that the effects obtained are, at least partially, attributable to learned action effects.

Keywords: Learning; Spatial; Pianists; Automatic; Stroop.

When an action is repeatedly experienced in the context of previously unfamiliar perceptual consequences of that action, there is potential for the establishment of a joint cognitive representation of the action and its effects. Once this has occurred, the perceptual consequences of an action may be sufficient to elicit the representation of the action itself (Hommel, Müsseler, Aschersleben, & Prinz, 2001; Prinz, 1990). For instance, Elsner and Hommel (2001) demonstrated that if individuals learn to associate left- and right-hand responses with high and low tones, respectively (or vice versa), the speed of their response on a subsequent unrelated task is

Correspondence should be addressed to Lauren Stewart, Department of Psychology, Goldsmiths, University of London, New Cross, London, SE14 6NW, UK. E-mail: l.stewart@gold.ac.uk

We thank Emily Duggan and Joseph Sparks for assistance with data collection. Rinus G. Verdonschot's research is supported by a Canon Foundation in Europe fellowship. Lauren Stewart is supported by a grant from the Economic and Social Research Council.

affected in a systematic way by the presence of these tones, even when instructions are given to ignore them.

One group for whom the execution of certain actions is intimately linked with their auditory consequences is the musically trained. A successful musical performance demands a continual and rapid interplay between the auditory and motor domains, in particular, to allow for online error monitoring. A recent study revealed an electrophysiological component that was present before an erroneous keypress was even made (Maidhof, Rieger, Prinz, & Koelsch, 2009). The mechanism for such early error detection was suggested to result from the detection of a mismatch between the predicted sensory consequence of the prepared (erroneous) action and the desired (correct) action goal. The close coupling between auditory effects and actions in musicians, once established, appears to be bidirectional, such that auditory events can prime motor plans and vice versa. Two functional magnetic resonance imaging (fMRI) studies compared activation when musicians (a) listened to a piece of music without playing it versus (b) played a piece of music without auditory feedback (Bangert et al., 2006; Baumann et al., 2007). Both studies showed areas of overlap between the two conditions, whereby premotor cortex, supplementary motor area, and planum temporale were activated in both conditions. Another study, using magnetoencephalography (MEG), showed that merely listening to music that was within the listener's repertoire (without playing it) resulted in a response within the primary motor cortex (Haueisen & Knösche, 2001). Moreover, a dissociation in the brain surface current density was seen between those notes that would have been played by the thumb and the little finger. Similarly, when pianists listened to a piano piece that they had practised, they demonstrated higher motor cortex excitability than when they listened to a structurally similar piece on which they had not been trained (D'Ausilio, Altenmüller, Olivetti Belardinelli, & Lotze, 2006).

Such studies are suggestive of close links between perception and action but, on their own, are insufficient to establish that such associations are both functional and automatic. Behavioural approaches employing interference methods hold special promise here. A pair of studies by Drost, Rieger, Brass, Gunter, and Prinz (2005a, 2005b) demonstrated effects on response time and, in certain cases, errors, when the association between actions and their perceptual effects was manipulated: Pianists were slower to play a musical chord or two-note sequences (indicated via various types of visually presented imperative stimuli) when hearing a chord that was incongruent with the required response. A subsequent study suggested a degree of instrumental specificity to this association: Pianists were slowed when they heard incongruent chords played on a keyboard instrument (piano/organ), but not a guitar; the reverse was true for guitarists (Drost, Rieger, & Prinz, 2007).

Drost and colleagues (2005a) were careful to consider the possible locus of interference in their studies. For instance, while the interference was hypothesized to arise from learned action–effect associations, the use of musical notation as the imperative stimulus prevented them from ruling out interference between the heard stimulus and the seen imperative stimulus, as opposed to between the heard stimulus and the required response. However, effects on response time were also seen when colour patches were used as the imperative stimulus instead of musical notation, thereby eliminating any possible association with the heard stimuli. Together with the absence of any interference effects in a nonmusician group, this allowed Drost and colleagues (2005a) to interpret the interference effects as evidence for learned action–effect associations in pianists.

Since the behavioural studies by Drost and colleagues (2005a, 2005b) were published, two independent groups have published findings that bear on the interpretation of the previously mentioned results (Lidji, Kolinsky, Lochy, & Morais, 2007; Rusconi, Kwan, Giordano, Umiltà, & Butterworth, 2006). Using both explicit and implicit speeded response time tasks, both groups showed that the mental representation of pitch has a spatially horizontal, as well as vertical, component. With respect to the horizontal component, responses were facilitated for pairings of low pitch with left response location and high pitch with right response location. With respect to the vertical component, responses were facilitated for pairings of low pitch with low response location and high pitch with high response location. These effects have been termed the spatial musical association of response codes (SMARC) (Rusconi et al., 2006) or the spatial pitch association of response codes (SPARC; Lidji et al., 2007). While the horizontal SMARC/SPARC effect was seen in musicians for both explicit and implicit tasks, nonmusicians showed this effect only in an explicit task, suggesting a weaker association of horizontal spatial and pitch-based codes. The most parsimonious explanation for the enhanced horizontal SMARC/SPARC effect in musicians would be that they have learned piano keyboard-based action–effect associations. This explanation is problematic, however, since musicians who are not keyboard players also show the same effect. Although an explanation for the enhanced horizontal SMARC/SPARC effect has yet to be offered, its existence presents a difficulty for the previously mentioned studies of action–effect associations, where congruent and incongruent trials differ not only in terms of the extent to which they conform to the associations that would be learned in the course of keyboard training, but also in terms of the extent to which they exhibit SMARC/SPARC mapping of low pitch–leftward response and high pitch–rightward response.

Our aim, in the present experiment, was to measure hypothesized action–effect associations in trained pianists using a paradigm that (a) incorporated both the imperative and the irrelevant dimensions of the task in a single auditory stimulus; (b) was suitable for use with both pianists and nonmusicians; (c) permitted investigation of potential action–effect associations in the context of the production of a sequence (thus building on the previous studies of chord and two-tone production); and (d) allowed us to explore the potential contribution of the SMARC/SPARC effect to any effects we obtained.

EXPERIMENT 1

In this experiment, we use a novel paradigm to compare the effect of manipulating pitch/number congruency (where "number" corresponds to a digit and defines the location of the response in horizontal space) in a group of skilled pianists and matched control nonmusicians. Our choice of paradigm in this experiment was motivated by a series of interference tasks that indicated the existence of involuntary visuomotor mappings between musical notation and its associated responses in pianists (Stewart, Walsh, & Frith, 2004). In these experiments, pianists positioned their right hand over five adjacent keys and made a sequence of key presses based on the presence of a sequence of numbers (1 = thumb, 2 = index finger, 3 = middle finger, etc.). Key presses did not produce any sound. The speed of their key presses was affected by the presence of to-be-ignored musical notation on which the numbers were superimposed. For example, if the five numbers 1, 2, 3, 4, and 5 were superimposed onto notes that occupied successively higher vertical positions on the musical stave, the pianists' responses would be faster than in a situation in which these same numbers were superimposed onto notes that occupied successively lower positions on the musical staff, since, in the former case, both the number sequence and the notation specify the same sequence of finger movements. Musically illiterate individuals showed no such effect of number/note congruency.

By adapting the above-mentioned notational Stroop task, it is possible to test the hypothesis that action–effect associations exist in pianists as a result of learned couplings between the motor and auditory systems. Instead of presenting sequences of numbers visually against a background of musical notation (as in the notational task), the present experiment involves the presentation of sequences of numbers that are presented aurally at pitches that do or do not correspond to the pitches that would normally be associated with such a sequence of finger movements (congruent and incongruent, respectively). As in the notational Stroop task, participants are required to make speeded key press responses to each number as soon as it is presented, and no sounds are produced. Assuming a response is to be made by the right hand, the numbers 1, 2, 3, 4, and 5, pitched at successively higher frequencies, would be an example of a congruent sequence, while the same numbers presented at successively lower frequencies would be an example of an incongruent sequence. We predicted that there would be a systematic effect of pitch/number congruency manipulation on pianists' performance, with congruent sequences producing faster responses than incongruent sequences. We predicted that nonmusicians, lacking the association between actions in horizontal space and their perceptual consequences, would be unaffected by the correspondence between number sequences and the pitch at which these numbers are presented.

Method

Participants

Two participant groups were used: 16 pianists (8 female, 3 left-handed, average age 27 years; $SD = 11.4$) and 16 nonmusicians (10 female, 1 left-handed, average age 33 years; $SD = 8.5$). On average, the participants in the pianist group had been playing piano for 20 years ($SD = 11.1$). The participants in the nonmusician group reported an absence of musical training, with the exception of three who had received training in childhood but had given up within two years. All the nonmusicians were musically illiterate. All participants gave informed consent. They were not informed about our experimental hypotheses.

Stimuli

A male voice was used to record spoken versions of the numbers "1", "2", "3", "4", and "5". These samples were manipulated using the sound analysis/resynthesis program STRAIGHT (Kawahara, Masuda-Katsuse, & de Cheveigné, 1999). A fast Fourier analysis was performed on the recorded sound samples. The fundamental frequency curve was set to a specific, static frequency, and the sound sample was resynthesized to create versions of the original, spoken sample at particular fundamental frequencies. Each of the original samples was resynthesized at fundamental frequencies of 98, 110, 123, 130, and 147 Hz, corresponding to a perceived pitch of G2, A2, B2, C3, and D3. The attack and decay envelope of each sample remained constant. Thirty-two different number sequences were generated. Each sequence comprised five numbers, drawn from the set {1, 2, 3, 4, 5}. Sequences were randomly generated, with the constraint that there could, but did not have to be, repetition of only one number within the sequence. Example sequences were: 4, 1, 3, 1, 5 and 5, 1, 2, 3, 4. Thus it was not necessary for each number to appear in the sequence, and the repetition of a single number in some sequences was intentional to prevent subjects predicting which numbers would appear (if every number appeared once and once only, then the appearance of 1, 3, 2 would allow prediction of the remaining two possibilities: 4 then 5, or vice versa). Having generated this set of number sequences, stimuli were constructed using the recorded sound samples. Each number sequence appeared in several different versions. In baseline sequences, number sequences were presented using the original recordings of natural speech. In congruent sequences, each number within the sequence was heard at a corresponding pitch (e.g., the number 1 at the lowest pitch, the number 2 at the second lowest pitch, and so forth). In incongruent sequences, each number within the sequence was heard at a noncorresponding pitch. However, the association between number and pitch in this condition was not random but the inverse of that in the congruent condition. In this way, both congruent and incongruent sequences exhibited the same degree of systematicity between the numbers and their associated pitches, but only the congruent sequences exploited the presumed association that pianists had learned through reciprocal auditory–motor coupling. As with the notational Stroop experiment (Stewart et al., 2004), we included "catch" sequences to prevent participants from adopting a strategy for the congruent sequences, whereby they would make a response based only on the pitch of the numbers. Catch sequences started congruently, but the final two or three numbers were incongruent. For all sequences, numbers were 500 ms in duration, with 1,500 ms between the start of each consecutive number. There was an interval of 3 s between consecutive sequences. Example stimuli can be downloaded at http://cms.gold.ac.uk/music-mind-brain/auditory-stroop/

Procedure

Participants were instructed to place their right hand over the following keys on a laptop keyboard: "spacebar", "U", "I", "O", and ";". These keys were chosen because they lie comfortably beneath the fingers. Participants were instructed to make responses based on the numbers that were presented over the headphones. If the number 1 was heard, they were required to press the key beneath their thumb, if the number 2 was heard, they were required to press the key beneath their index finger, and so on. Participants were instructed to make their responses immediately after the presentation of each number, as opposed to waiting until all numbers had been presented. Participants maintained visual fixation on a centrally presented cross-hair throughout the experiment. There were a total of 112 trials (32 each of baseline, congruent, and incongruent trials and 16 catch trials), which were presented in a fixed random order, with seven blocks of 16 trials each, lasting approximately 20 min in total.

An initial practice block was conducted to familiarize the participant with the number to finger mappings that would be used in the actual experiment. The sequences within the familiarization block were, like the baseline sequences described above, the original, natural spoken recordings of the numbers 1 to 5. The precise number sequences used were different from those used in the actual experiment in order to prevent sequence-specific learning effects. Twenty aurally presented number sequences were presented. An adaptive thresholding algorithm was used to alter the presentation rate of numbers within a sequence according to the participants' level of accuracy. Initially, the interval between consecutive numbers within a sequence was 2,000 ms. Accurate performance resulted in a reduction of this interval by 500 ms. If any errors were made for a particular sequence, this interval would increase by the same amount. By the end of these 20 sequences, all participants were performing accurately with an interval of 1,000 ms between consecutive numbers within a sequence. Following this familiarization procedure, the experiment began.

Analysis

Each key press was automatically logged and scored for response time and accuracy. Response times were averaged for every trial, excluding erroneous responses. A grand average response time and a total error score were calculated for each sequence type (baseline, congruent, incongruent, and catch).

Results and discussion

Prior to the main analysis, we inspected the response times and error rates across all sequence types, with particular focus on how pianists performed on catch sequences, relative to the other sequence types. Catch sequences were included to ensure that participants were responding on the basis of number information, not pitch. In the case of congruent trials, number and pitch provide redundant information, and it was conceivable that pianist participants could become aware of this at the beginning of the trial and respond purely on the basis of pitch information for the remainder of the trial. Since catch trials started congruently and finished incongruently, such a strategy would result in a higher error rate for these trials than for the other sequence types. Response times for catch trials were found to be intermediate between response times for congruent and incongruent trials [mean (*SD*): baseline, 530 ms (87); congruent, 525 ms (89); incongruent, 552 ms (81); catch, 534 ms (86)], and error rates were not significantly different across the four sequence types [mean (*SD*): baseline, 3% (2.28); congruent, 2% (1.78); incongruent, 3% (3.22); catch, 3% (2.29)], $\chi^2(3) = 0.11$, $p = .99$. The absence of an elevated error rate for catch trials suggested that participants were, as intended, responding on the basis of number information, not pitch, when there was redundancy in the information provided by these two cues. Since catch trials were simply included to minimize this strategy, response times and error rates for this sequence type were therefore not considered in subsequent analyses.

Mean response times and error rates for baseline, congruent, and incongruent sequence types and for both groups can be seen in Figure 1. The pianists in the present study made significantly faster responses than did the nonmusicians, across all sequences types. Such a finding has precedence in the literature (e.g., Brochard, Dufour, & Després, 2004; Hughes & Franz, 2007; Patston, Hogg, & Tippett, 2007) and can be attributed to an overall enhancement in sensorimotor processing that accompanies instrumental skill learning.

We compared error rates across baseline, congruent, and incongruent sequences for both groups separately. This revealed no significant difference in error rates according to sequence type: pianists, $\chi^2(2) = 0.04$, $p = .98$; nonmusicians, $\chi^2(2) = 1.48$, $p = .48$.

Turning to response times, a repeated measures analysis of variance (ANOVA) with a single within-subjects variable of sequence type (3 levels: baseline, congruent, incongruent) and a between-subjects variable of group (2 levels: pianists, nonmusicians) demonstrated a main effect of sequence type, $F(2, 60) = 10.22$, $MSE = 235.96$, $p < .001$, a main effect of group, $F(1, 30) = 38.74$, $MSE = 23{,}051.88$, $p < .001$, and an interaction between sequence type and group, $F(2, 60) = 4.427$, $MSE = 235.96$, $p = .02$.[1] Subsequent ANOVAs were conducted for both groups separately. For pianists, but not nonmusicians, there was a main effect of sequence type: pianists, $F(2, 30) = 11.68$, $MSE = 283.26$, $p < .001$; nonmusicians, $F(2, 30) = 0.78$, $MSE = 188.65$, $p = .47$. Pairwise comparisons (Bonferroni adjusted $\alpha = .017$) showed that pianists' response times for incongruent sequences were significantly slower than those for baseline sequences, $t(15) = 3.93$, $p < .001$, and congruent sequences, $t(15) = 3.76$, $p = .01$. There was no significant difference in response time for baseline versus congruent sequence types. As a supplementary analysis, we focused solely on data from the congruent and incongruent trials and computed the ratio of response times for incongruent relative to congruent trials in both groups. A comparison of these ratios using an independent-sample t test revealed significantly higher ratios in the pianist group, $t(30) = 2.68$, $p = .01$.[2]

In Experiment 1, we used a novel interference task to test the hypothesis that pianists possess action–effect associations as a result of learned coupling between actions and their perceptual consequences. Our hypothesis was supported: Only the pianist group showed a significant effect of the pitch/number congruency manipulation, such that they were faster when the imperative number stimuli were presented at pitches that were congruent with the anticipated action effects than when the same stimuli were presented at pitches that were incongruent with the anticipated action effects. Nonmusicians showed no such effect. While consistent with an explanation in terms of learned coupling between actions and their perceptual consequences, the interference effect seen in the pianist group is also consistent with the finding of an enhanced horizontal SMARC/SPARC effect in musicians relative to nonmusicians (Lidji et al., 2007; Rusconi et al., 2006). Experiment 2 was therefore conducted to test the contribution of an enhanced horizontal SMARC/SPARC effect to the effects seen in Experiment 1.

EXPERIMENT 2

The rationale for this experiment was to compare the effect of pitch/number congruency manipulation in pianists and nonmusicians, using two versions of the

[1] An analysis of the effect of sequence type for slow versus fast responders in each group demonstrated that this Group × Sequence Type interaction could not be accounted for by overall differences in response time between the two groups.

[2] Ratios for congruent relative to baseline trials were also compared between groups. No significant difference was found, $t(30) = 0.86$, $p = .4$.

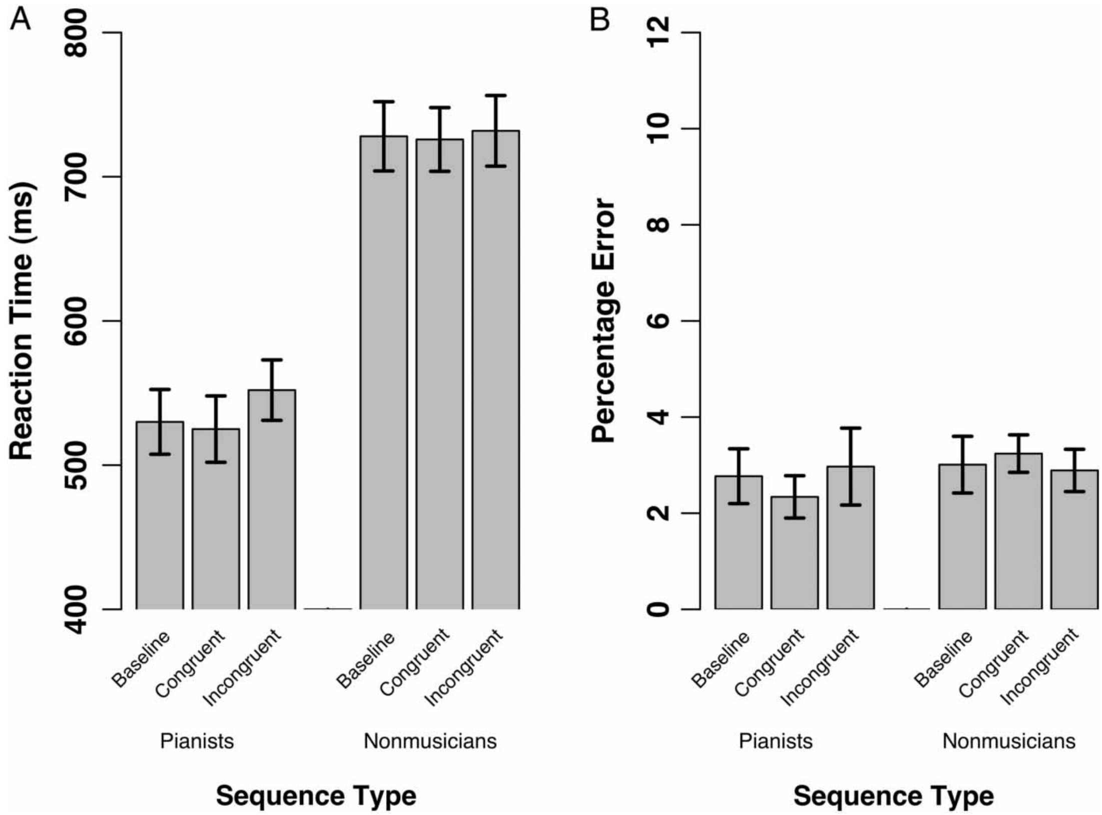

Figure 1. *Mean response times (A) and percentage errors (B) for Experiment 1.*

interference paradigm used in Experiment 1: the original ("standard") version (as in Experiment 1) and a modified ("stretched") version. In the stretched version, the pitch range over which the imperative stimuli were presented was increased, from a little more than half an octave (in the standard version) to a two-octave span (in the stretched version). While, in the standard version, the imperative stimuli were presented at frequencies that spanned a pitch range that would naturally lie beneath the corresponding fingers, this was not the case in the stretched version, where the pitch range of the imperative stimuli was far in excess of this. Importantly, both versions exploited the association between low pitch/left space and high pitch/right space that is the basis for the horizontal SMARC/SPARC effects (Lidji et al., 2007; Rusconi et al., 2006). The modified version may be considered to have even greater potential for SMARC/SPARC effects, owing to the exaggerated pitch range used. However, the important difference between the two tasks was that only the standard version used pairings of imperative stimuli and pitch that reflect auditory–motor couplings that would be expected to arise from piano training. If the effect of the pitch/number congruency manipulation seen in pianists in Experiment 1 can be accounted for by a musician-specific enhanced horizontal SMARC/SPARC effect (low pitch/left space and high pitch/right space), the effect of sequence type should be similar or even greater in the stretched compared with the standard version of the task. If, on the other hand, the effect of the pitch/number congruency manipulation seen in pianists in Experiment 1 is due to learned action–effect associations, the effect of sequence type should be greater in the standard version than in the stretched version. Nonmusicians were not expected to show an effect of pitch/number congruency manipulation in either version.

Method

Participants

Two new participant groups were used: 15 pianists (9 female, 2 left-handed, average age 26 years; $SD = 8.6$) and 15 nonmusicians (5 female, 2 left-handed, average age 23 years; $SD = 1.82$). On average, the participants in the pianist group had been playing piano for 17 years ($SD = 4.2$). The participants in the nonmusician group reported an absence of musical training, apart from 4 who had received training in childhood but had given up within two years. All the nonmusicians were musically illiterate. All participants gave informed consent. They were not informed about our experimental hypotheses.

Stimuli

The original ("standard") version involved the same stimuli and procedure as those described for Experiment 1. Stimuli in the modified ("stretched") version were generated in a similar way to that described for Experiment 1, but each of the original spoken number samples was resynthesized to one of the following frequencies: 62, 88, 124, 165, and 234 Hz, corresponding to B1, F2, B2, E3, and A#3. This resulted in a pitch range spanning approximately two octaves, more than three times the pitch range used in Experiment 1, though incorporating the same region of the keyboard as that used in Experiment 1. As for the standard version, there were a total of 112 trials (32 each of baseline, congruent, and incongruent trials and 16 catch trials), which were presented in a fixed random order, with seven blocks of 16 trials each. Example stimuli can be downloaded at http://www.gold.ac.uk/music-mind-brain/auditory_stroop

Procedure

Each participant took part in both versions of the interference tasks (standard and stretched), run in separate blocks within the same sitting, with order counterbalanced across participants. The procedure was identical to that described for Experiment 1.

Analysis

As for Experiment 1, mean response times for correct key presses were collated, and errors were logged according to sequence type.

Results

Mean response times and error rates for baseline, congruent, and incongruent sequence types can be seen in Figures 2A and 2B (standard version) and 2C and 2D (stretched version). As a first step, we analysed the data from the standard version to determine whether we had replicated the pianist-specific effect of congruency manipulation seen in Experiment 1 in an independent cohort of participants. As in Experiment 1, we first inspected the response times and error rates across all sequence types, with particular focus on how pianists performed on catch sequences, relative to the other sequence types. As in Experiment 1, response times for catch trials were found to be intermediate between response times for congruent and incongruent trials [mean (*SD*): baseline, 507 ms (72); congruent, 509 ms (57); incongruent, 555 ms (87); catch, 539 ms (71)], and error rates did not differ significantly with sequence type [mean (*SD*): baseline, 3% (2.26); congruent, 3% (1.95); incongruent, 3% (2.2); catch, 3% (2.56)], $\chi^2(3) = 1.46$, $p = .69$. As before, this demonstrated that participants were responding, as intended, on the basis of number, rather than pitch. Data corresponding to catch trials were therefore not considered in subsequent analyses.

The remaining data were first analysed for an effect of the order in which the standard and stretched version had been presented. There was no main effect of order and no interaction with any other variable so the data were subsequently analysed independently of presentation order.

In terms of the response time data relating to the standard version, we used a repeated measures ANOVA with one within-subjects variable,

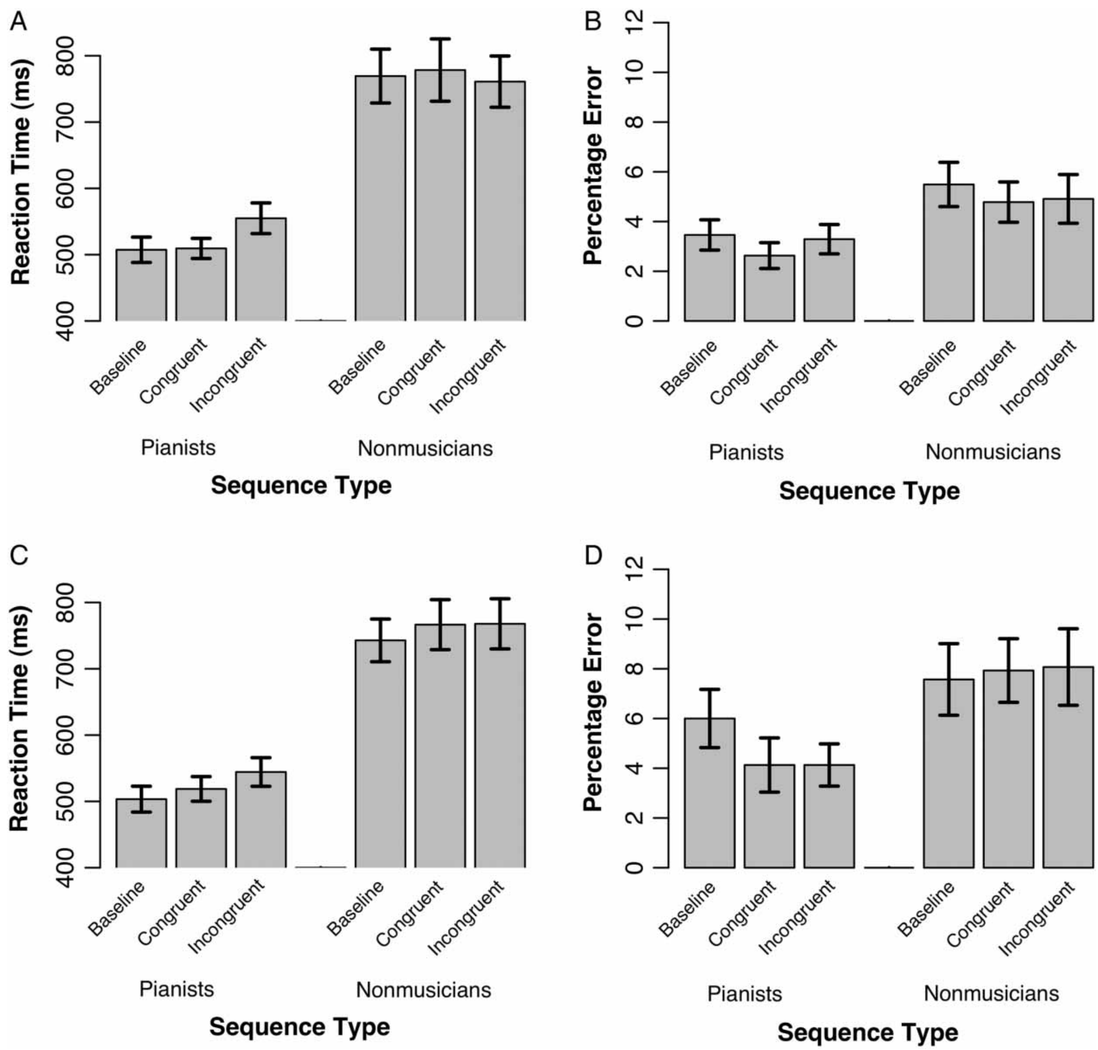

Figure 2. *Mean response times and percentage errors for Experiment 2, standard (A and B) and Experiment 2, stretched (C and D).*

sequence type (3 levels: baseline, congruent, incongruent) and one between-subjects variable, group (2 levels: pianist, nonmusician). This revealed a marginally significant effect of sequence type, $F(2, 56) = 2.94$, $MSE = 1{,}035.61$, $p = .06$, a main effect of group, $F(1, 28) = 869.89$, $MSE = 43{,}274.58$, $p < .001$, and an interaction between sequence type and group, $F(2, 56) = 8.6$, $MSE = 1{,}035.61$, $p = .001$.[3] For pianists but not nonmusicians, there was a main effect of sequence type: pianists, $F(2, 28) = 16.83$, $MSE = 642.45$, $p < .001$; nonmusicians, $F(2, 28) = 0.80$; $MSE = 1{,}428.77$, $p = .46$. Pairwise comparisons (Bonferroni adjusted $\alpha = .017$) showed that pianists' response times for incongruent sequences were significantly slower than those for baseline sequences, $t(14) = -4.91$,

[3] As in Experiment 1, an analysis of the effect of sequence type for slow versus fast responders in each group demonstrated that this Group × Sequence Type interaction could not be accounted for by overall differences in response time between the two groups.

$p < .001$, and congruent sequences, $t(14) = -3.99$, $p = .001$. There was no significant difference in response time for baseline versus congruent sequence types.

We also compared error rates across baseline, congruent, and incongruent sequences for both groups separately. This revealed no significant difference in error rates according to sequence type: pianists, $\chi^2(2) = 2.35$, $p = .31$; nonmusicians, $\chi^2(2) = 0.28$, $p = .87$.

Having replicated the Group × Sequence Type interaction seen in the standard version in an independent cohort of participants, we next ran a similar analysis with the data from the stretched version. This revealed a main effect of sequence type, $F(2, 56) = 6.63$, $MSE = 1{,}036.59$, $p = .003$, a main effect of group, $F(1, 28) = 35.30$, $MSE = 35{,}027.48$, $p < .001$, but no interaction between sequence type and group, $F(2, 56) = 1.15$, $MSE = 1{,}036.59$, $p = .32$. For pianists, but not nonmusicians, there was a main effect of sequence type: pianists, $F(2, 28) = 34.5$, $MSE = 186.15$, $p < .0001$; nonmusicians, $F(2, 28) = 2.6$, $MSE = 1{,}140.39$, $p = .09$. Pairwise comparisons (Bonferroni adjusted $\alpha = .017$) showed that pianists' response times for incongruent sequences were significantly slower than those for baseline sequences, $t(14) = -6.8$, $p < .0001$, and congruent sequences, $t(14) = -4.86$, $p < .0001$. In addition, response times for congruent sequences were significantly slower than those for baseline trials, $t(14) = -4.69$, $p < .0001$. For nonmusicians, there were no significant differences between any of the sequence types.

We compared error rates across baseline, congruent, and incongruent sequence types for both groups separately. Pianists, but not nonmusicians, differed in error rates according to sequence type: pianists, $\chi^2(2) = 8.4$, $p = .02$; nonmusicians, $\chi^2(2) = 0.28$, $p = .87$. This reflected a higher rate of errors for the baseline sequences than for the congruent sequences, $t(14) = 2.77$, $p = .02$, as well as relative to the incongruent sequences, $t(14) = 3.34$, $p = .01$. This increase in errors for baseline sequences relative to the other sequence types is seen in conjunction with the relatively faster response times reported above and constitutes a speed/accuracy trade-off for this sequence type in the pianist group.

To summarize thus far, we found a significant effect of sequence type for pianists in both versions (standard; stretched) while no significant effect of sequence type was found for nonmusicians in either version. Of critical interest was a comparison of the effect of sequence type as a function of version and group. Experiments 1 and 2 had both found a robust effect of sequence type that was driven by a relative slowing for incongruent versus congruent trials, while congruent trials showed no facilitation relative to baseline. This pattern of results is not uncommon in the Stroop paradigm literature, where interference effects tend to dominate (Glaser & Glaser, 1982). In the context of the present study, the lack of facilitation for congruent trials relative to baseline may be explained by the acoustic differences between baseline trials, on the one hand, and both congruent and incongruent trials on the other. While baseline trials were spoken, incorporating a range of frequencies, congruent and incongruent trials were pitched at specific frequencies. The relative infrequency of baseline trials appears to have resulted in a slight attentional orienting, as remarked upon by several participants. This effect appears to have been particularly evident in the stretched condition, presumably due to the more exaggerated use of pitching, resulting in an even greater discrepancy in the timbre of baseline trials on the one hand and congruent/incongruent trials on the other, such that response times for baseline trials were faster than even congruent trials in both groups (which may explain the trend towards a main effect of sequence type in the nonmusician group). Given that baseline trials clearly did not constitute an adequate "neutral" condition and the fact that the observed effect of sequence type was clearly driven by the difference between congruent and incongruent trials types, our subsequent analyses concentrated only on data from congruent and incongruent trials.

We next transformed the response time data into ratios (mean response time for incongruent sequence type divided by mean response time for congruent sequence type). This allowed us to focus solely on differences in response time between congruent and incongruent trials, as a function of version, regardless of the absolute response times, which constituted a large source of variance across participants. Ratios greater than 1 indicated a longer mean response time for incongruent than for congruent sequence type; ratios less than 1 indicated the reverse. Mean response time ratios according to version can be seen, for both groups, in Figure 3.

We used a repeated measures ANOVA with one within-subjects variable, version (2 levels: standard, stretched) and one between-subjects variable, group (2 levels: pianist, nonmusician). This revealed a main effect of group, $F(1, 28) = 22.38$, $MSE = 0.004$, $p < .001$, and an interaction between version and group, $F(1, 28) = 5.69$, $MSE = 0.002$, $p = .02$. Pairwise comparisons showed that the ratio of response times differed significantly for the two versions for the pianists only, reflecting a greater response time difference between incongruent and congruent trials for the standard than for the stretched version: pianists, $t(14) = 2.28$, $p = .04$; nonmusicians, $t(14) = -1.11$, $p = .29$. While the ratio of incongruent to congruent response times was significantly greater than 1 for both versions in the pianist group [standard, $t(14) = 4.28$, $p = .001$; stretched, $t(14) = 5.43$, $p < .001$], this was not the case for either version for the nonmusicians, who were unaffected by the sequence type in both cases: standard, $t(14) = -1.63$, $p = .12$; stretched, $t(14) = 0.221$, $p = .83$.[4]

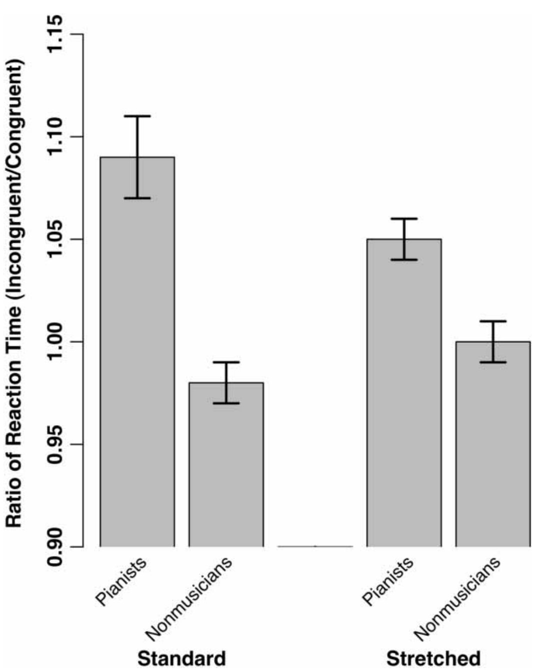

Figure 3. *Mean ratios of response times (incongruent/congruent) for Experiment 2, standard and stretched.*

GENERAL DISCUSSION

The principle of common coding implies that a joint representation is formed when actions are repeatedly paired with a specific perceptual event (Prinz, 1990). Musicians are an ideal group in which to examine the presence and characteristics of such integrated representations, since they are occupationally specialized with regard to a particular type of action–effect binding. They must be intimately attuned to the coupling between actions and their auditory consequences, in order to monitor and refine their performance, both technically and aesthetically. Two previous studies have used interference paradigms to demonstrate the existence of action–effect associations in pianists (Drost et al., 2005a, 2005b). The results of these experiments are consistent with the existence of such associations but findings concerning spatial compatibility effects that are greater in musicians preclude an interpretation solely in terms of learned couplings between actions and perceptual events. The aim of the present study was to develop and use a paradigm that could test for the existence of action–effect associations in pianists, while controlling for the effects of spatial compatibility effects (Lidji et al., 2007; Rusconi et al., 2006). In Experiment 1, response time differences between congruent and incongruent sequences in pianists, but not nonmusicians, were consistent with the presence of action–effect associations but did not preclude the possibility that such effects were due to enhanced spatial compatibility effects in musicians. Experiment 2 compared the effects of congruency manipulation in the original version of the task with those in a "stretched" version, where imperative stimuli were presented at pitches spanning a range that clearly did not correspond to the pitches that would normally be produced by the fingers of a single hand. Consistent with the demonstration of an exaggerated horizontal SMARC/SPARC effect in musicians (Lidji et al., 2007; Rusconi et al., 2006), pianists, but not nonmusicians, showed an effect of sequence type in the stretched version. However, the significantly reduced effect of sequence type in the stretched compared with the standard version suggests that the effects seen in Experiment 1 are, at least partly, attributable to learned action–effect associations.

This, to our knowledge, is the first demonstration of the existence of action–effect associations in musicians that takes into account the possible contribution of a SMARC/SPARC effect. In addition, the results challenge previous claims that action–effect associations are instrument specific (Drost et al., 2007) since, in the present paradigm, the imperative and to-be-ignored dimensions of the stimulus were both carried by the voice. While there may indeed be instrument-specific contributions to the action–effect associations seen in musicians, the present paradigm has demonstrated that pitch can be a sufficiently strong cue by itself to influence response time.

It is important to note that the interference effects shown here were seen even though the response was made on a computer keyboard as opposed to a piano keyboard (see also Repp & Knoblich, 2007, 2009). As with the notational Stroop effect observed in pianists (Stewart et al., 2004), the interference effects appear to depend on the correspondence of pitch and space between consecutive tones/actions, as opposed to the congruence of the pitch–space relations of individual elements of the sequence. Studies in which auditory feedback has been manipulated during performance also demonstrate that similarity between the planned action sequences and the perceived sequences arises from the relationship between movement transitions between perceived events since participants were affected to a similar degree when auditory feedback was delayed and transposed as opposed to only delayed (Pfordresher, 2008).

Several behavioural studies, including the present one, have focused on the extent to which motor responses can be primed by the perception of an irrelevant sensory attribute that is similar to the sensory consequences of the planned action. Interestingly, there is also evidence that merely imagining the effects of actions can prime motor responses (for a review, see Koch, Keller, & Prinz, 2004), a process that seems particularly well developed in musicians. Keller and Koch (2008) asked participants to produce simple motor sequences using three keys, which were vertically aligned. Different colours, visually presented, cued participants to produce one of four different sequences, and key presses resulted in auditory feedback that could be compatible or incompatible with respect to the keys pressed. "Compatible" feedback involved the pairing of low-, medium-, and high-pitched tones for responses made at low, medium, and high positions, in vertical space, in line with the documented association between vertical position and pitch height (Lidji et al., 2007; Rusconi et al., 2006). "Incompatible" feedback involved either the reverse mapping or an arbitrary pairing between keys and tones. Response initiation times were faster with compatible feedback, and this effect increased with the level of musical training, suggesting that trained musicians may make use of such imagined action effects when planning movements in a musical context.

The present findings motivate several further questions for future investigation. Comparisons with other instrumental groups can be informative here. First, while the use of numbers as the imperative stimulus exploits an overlearned mapping that is present in both musicians and nonmusicians alike, future comparisons between pianists and other instrumentalists besides pianists would be able to investigate whether musicians (in general) additionally possess additional mappings between numbers and scale degrees, which might contribute to the effect we have observed.

[4] Ratios were also computed for congruent relative to baseline trials, for both versions. No significant differences were found between pianists and nonmusicians: standard, $t(28) = -0.11$, $p = .91$; stretched, $t(28) = -0.28$, $p = .78$.

Secondly, it would be informative to disentangle a possible effect of tonal ambiguity (lack of a tonal centre), which was inherent to the pitch manipulation used in the stretched condition. Thirdly, it will be important to determine the extent to which the learned pitch–space mappings demonstrated in the present study are dependent on hand configuration and/or the orientation of the instrument. This question could be explored by adapting the paradigm for violinists, for whom the hand can be rotated independently of the orientation of the instrument. In this way, pitch–space mappings could be made to be congruent with the precise finger movements typically required to produce such sounds, or with the spatial locations on the instrument that would normally produce such sounds, or both of these. The prediction can be made that the effect of congruence would depend upon both the hand position and the instrument orientation. Finally, it will be valuable to explore the extent to which audio–motor coupling is facilitated by certain pedagogical approaches; that is, do approaches, such as the Suzuki method, which emphasize sound to action coupling before the development of musical literacy, foster more robust or earlier developing common representations between the auditory and motor domains? In sum, the investigation of learned perception–action coupling in different instrumental groups provides a rich model with which to address questions about the relation of action and perception, the development of such mappings at different ages, levels of skill, and the impact of different pedagogical approaches.

Original manuscript received 29 March 2012
Accepted revision received 10 April 2012
First published online 20 June 2012

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*(3), 917–926.

Baumann, S., Koeneke, S., Schmidt, C. F., Meyer, M., Lutz, K., & Jancke, L. (2007). A network for audio-motor coordination in skilled pianists and non-musicians. *Brain Research*, *1161,* 65–78.

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

D'Ausilio, A., Altenmüller, E., Olivetti Belardinelli, M., & Lotze, M. (2006). Cross-modal plasticity of the motor cortex when listening to a rehearsed musical piece. *European Journal of Neuroscience*, *24,* 955–958.

Drost, U. C., Rieger, M., Brass, M., Gunter, T. C., & Prinz, W. (2005a). Action–effect coupling in pianists. *Psychological Research*, *69,* 233–241.

Drost, U. C., Rieger, M., Brass, M., Gunter, T. C., & Prinz, W. (2005b). When hearing turns into playing: Movement induction by auditory stimuli in pianists. *The Quarterly Journal of Experimental Psychology*, *58*(8), 1376–1389.

Drost, U., Rieger, M., & Prinz, W. (2007). Instrument specificity in experienced musicians. *Quarterly Journal of Experimental Psychology*, *60*(4), 527–533.

Elsner, B., & Hommel, B. (2001). Effect anticipation and action control. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(1), 229–240.

Glaser, M. O., & Glaser, W. R. (1982). Time course analysis of the Stroop phenomenon. *Journal of Experimental Psychology: Human Perception and Performance*, *8,* 875–894.

Haueisen, J., & Knösche, T. R. (2001). Involuntary motor activity in pianists evoked by music perception. *Journal of Cognitive Neuroscience*, *13,* 786–792.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, *24,* 849–878.

Hughes, C. M., & Franz, E. A. (2007). Experience-dependent effects in unimanual and bimanual reaction time tasks in musicians. *Journal of Motor Behaviour*, *39*, 3–8.

Kawahara, H., Masuda-Katsuse, I., & de Cheveigné, A. (1999). Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous frequency based F0 extraction. *Speech Communication*, *27,* 187–207.

Keller, P. E., & Koch, I. (2008). Action planning in sequential skills: Relations to music performance. *Quarterly Journal of Experimental Psychology*, *61,* 275–291.

Koch, I., Keller, P., & Prinz, W. (2004). The ideomotor approach to action control: Implications for skilled performance. *International Journal of Sport and Exercise Psychology*, *2*(4), 362–375.

Lidji, P., Kolinsky, R., Lochy, A., & Morais, J. (2007). Spatial associations for musical stimuli: A piano in the head? *Journal of Experimental Psychology: Human Perception and Performance*, *33*(5), 1189–1207.

Maidhof, C., Rieger, M., Prinz, W., & Koelsch, S. (2009). Nobody is perfect: ERP effects prior to performance in musicians indicate fast monitoring processes. *PLoS ONE*, *4*(4), e5032. doi:10.1371/journal.pone.0005032

Patston, L. L., Hogg, S. L., & Tippett, L. J. (2007). Attention in musicians is more bilateral than in non-musicians. *Laterality*, *12*, 262–72.

Pfordrescher, P. Q. (2008). Auditory feedback in music performance: The role of transition-based similarity. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(3), 708–725.

Prinz, W. (1990). A common coding approach to perception and action. In O. Neumann & W. Prinz (Eds.), *Relationships between perception and action*. Berlin, Germany: Springer-Verlag.

Repp, B. H., & Knoblich, G. (2007). Action can affect auditory perception. *Psychological Science*, *18,* 6–7.

Repp, B. H., & Knoblich, G. (2009). Performed or observed keyboard actions affect pianists' judgments of relative pitch. *Quarterly Journal of Experimental Psychology*, *62,* 2156–2170.

Rusconi, E., Kwan, B., Giordano, B. L., Umiltà, C., & Butterworth, B. (2006). Spatial representation of pitch height: The SMARC effect. *Cognition*, *99,* 113–129.

Stewart, L., Walsh, V., & Frith, U. (2004). Reading music modifies spatial mapping in pianists. *Perception and Psychophysics*, *66,* 183–195.