Motor and Audiovisual Learning Consolidate Auditory Memory of Tonally Ambiguous Melodies

Andrea Schiavio & Renee Timmers
University of Sheffield, Sheffield, United Kingdom

THE PRESENT STUDY INVESTIGATED THE ROLE OF motor and audiovisual learning in the memorization of four tonally ambiguous melodies for piano. A total of one hundred and twenty participants divided into three groups — pianists, other musicians (i.e., not pianists), and nonmusicians — learned the melodies through either playing them on a keyboard (*playing condition*), through performing the melodies on a piano without auditory feedback (*silent playing condition*), through watching a video with a performer playing the melodies (*seeing condition*), or through listening to them (*control condition*). Participants were exposed to each melody four times during the learning phase (in additional to hearing it once during a familiarization phase). This exposure consisted of an alternation between hearing the melody *and engaging with the melody* in the way determined by the learning condition. Participants in the control group only received the auditory aspect of the learning phase and listened to each melody twice. Memory of the melodies was tested after a 10-minute break. Our results indicate a benefit of motor learning for all groups of participants, suggesting that active sensorimotor experience plays a key role in musical skill acquisition.

ONE OF THE MAIN RECENTLY EMERGING insights of the *embodied paradigm* is that brains, bodies, and environment participate dynamically in driving cognitive processes (Chemero, 2009; Colombetti, 2014; Hurley, 1998; Varela, Thompson, & Rosch, 1991). By this view, cognition is best understood as an *activity* involving the entire living system in its coupling with the world, and is thus not reducible to skull-bound computational processes (Gallagher, 2005; Gallagher & Zahavi, 2008; Lakoff & Johnson, 1999; Merleau-Ponty, 2002; Rizzolatti & Sinigaglia, 2008; Shapiro, 2010; Sheets-Johnstone, 2010, 2011). The centrality of action for cognition is well exemplified by the notion of *sensorimotor contingencies*, such that sensorimotor contingencies are "the structure of the rules governing the sensory changes produced by various motor actions" (O'Regan & Noë, 2001, p. 941). Living systems thus acquire concepts that lead to the mastery of a particular (i.e., musical) perceptual experience through an active engagement with the world, whereby relevant sensorimotor knowledge is developed (Noë, 2004). In other words, embodied frameworks suggest that we are not passive receivers of external stimuli that are perceived, processed, and stored "in the head." Rather, our bodies - in their situated dialogue with the environment - do much of the work needed to make sense of the world without necessarily involving complex intellectual processes (Wilson & Golonka, 2013). As music is a human activity that involves a vast range of intellectual, ecological, and bodily, resources, the embodied framework just mentioned might help us gain a richer understanding of its complex experiential and cognitive dynamics, and to explore to what extent motor activity shapes our musical mind more generally.

The study of audio-motor integration in musical contexts, for example, has increased over the last few decades. Much of this work has focused on the field of musical learning, which represents an excellent domain to shed further light on the *motor basis* of musical experience (Palmer & Meyer, 2000). Recent studies show a direct correlation between motor knowledge and auditory experience; in particular after long-term motor training, as with adult professional musicians (e.g., Koeneke, Lutz, Wustenberg, & Jancke, 2004; Phillips-Silver & Keller, 2012). Gaining adequate motor knowledge is indeed an important prerequisite for developing relevant musical skills (Bowman, 2004; Schiavio, 2014a). Along these lines, several studies suggest that the mastery of specific, goal-directed actions facilitates the automatic recognition of (the goals underlying) witnessed or heard actions (Haslinger et al., 2005; Kohler et al., 2002; Umiltà et al., 2001, 2008). Put simply, it appears that knowing how to perform a specific (musical) action impacts the way in which an agent understands

the same action, even if he or she perceives only the sounds associated with that goal-directed movement (Overy & Molnar-Szakacs, 2009). For example, the sound produced by guitar strings will elicit different responses in a skilled guitarist compared to an amateur because of their different *history of structural coupling* with guitar-related actions. A guitarist will have first-hand experience of the goal-directed movements necessary to produce a particular sound on the guitar, and will therefore recruit the same neural population in both performing these movements and in perceiving the sounds provoked by that motor behavior, showing that motor musical expertise impacts auditory motor integration (Bangert & Altenmüller, 2003; Drost et al., 2005; Haueisen & Knösche, 2001; Schiavio, Menin, & Matyja, 2014).

This suggests that the listeners' or performers' ability to understand, perceive, and learn music is primarily a form of sensorimotor activity that depends on the level of musically related motor knowledge of the agents (Novembre & Keller, 2011) in their continuous and dynamical coupling with the environment (Clarke, 2005; Krueger, 2014; Matyja & Schiavio, 2013). Accordingly, we maintain that differences between expert musicians and amateurs are not so much a matter of differences in intellectual processes and theoretical knowledge of music, but instead are based on their concrete ability to generate and recognize meaningful musical actions. As such, a musical event can be intended "in terms of the intentional, hierarchically organized sequences of expressive motor acts behind the signal" (Overy & Molnar-Szakacs, 2009, p. 492). The concrete implications for musical learning, however, are still not fully clear. Is a musician's memory instrument-specific? And how is a nonmusician's memory affected by the visual and auditory perception of music-directed actions? Findings within the visual domain suggest that the "mere observation of movements leads to the formation of a lasting specific memory trace in movement representations that resembled that elicited by physical training" (Stefan et al., 2005, p. 9344). It is also well known that auditory-motor associations acquired while learning to play a piece can be determinant for later auditory recognition (Brown & Palmer, 2012). However, most of the relevant literature focuses on a population of well-trained musicians engaged in expertise-related tasks with their own instrument. For example, in their study on musical learning, Brown and Palmer (2012, p. 567) argue that "motor learning can aid performers' auditory recognition of music beyond auditory learning alone, and that motor learning is influenced by individual abilities in mental imagery and by variation in acoustic features." However,

the authors consider only a population of expert pianists, thus referring to subjects whose music-directed motor vocabulary (that is, the repertoire of actions that a musician develops through practice with his or her instrument) is already partially constituted.

The present study, therefore, aims to investigate the *constitution* of a musical-motor vocabulary of acts (Lahav, Saltzman, & Schlaug, 2007). Here we compare pianists, other musicians (i.e., not pianists), and non-musicians in the same piano-based task to see not only whether the motor system's engagement can be similar in musicians and nonmusicians when the task requirements are musically relevant, but also to understand by what terms a musician's ability to make sense of the musical patterns can be understood as cross-modal (see Eitan & Timmers, 2010). That is, we are interested in investigating the concrete *modalities* by which *musical subjects* (listeners, learners, and/or performers) with different musical expertise engage with *musical objects* (the musical stimuli). We therefore asked participants to familiarize themselves with four *tonally ambiguous* melodies in different conditions and with diverse degrees of motor engagement (playing, playing without any auditory feedback, only seeing the hand of a performer playing, and a control condition with no clear motor engagement). In doing so, we tested their ability to recognize the melodies acoustically, alternating the target ("correct") melodies with similar ones. It is important to note that we used music that lacks a strong tonal centre in order to avoid any facilitation in remembering melodic fragments with the help of common tonal rules. A study by Novembre and Keller (2011), for example, highlighted the "syntactic knowledge" shared between the rules of tonal music and the relevant chain of actions of trained musicians. Their paper shows how the tonal norms of the context facilitate the imitation of a given chord, despite the absence of any auditory feedback. This is not surprising, considering the literature reviewed above: trained musicians store meaningful motor knowledge in their music vocabulary of acts — the kind of knowledge that impacts training and performing. Thus, given that many musicians in Western contexts are familiar with tonal structures, we expected that having a strong tonal context in our study would facilitate the performance of the pianists. Since we did not want subjects whose motor vocabulary could match with the actions underlying a tonal context, we employed tonally ambiguous stimuli. Generally speaking, we expected the subjects' newly acquired motor expertise (in terms of first-hand experience with the learning condition involving playing the piano mute or with sound) to show its advantages beyond other

learning modalities. In particular, we hypothesized that a coupling between auditory and motor information would allow the participants who performed the melodies on the piano (with both auditory and motor feedback) to recall them with greater accuracy than subjects exposed to auditory and visual information only. For this reason, we also expected an advantage for pianists in familiarizing with the stimuli.

Method

PARTICIPANTS

One hundred and twenty healthy adult volunteers – split into three groups – participated in this study. Group A included 40 musicians who did not play the piano (median age = 25). Group B included 40 pianists (median age = 25). Group C included 40 nonmusicians (median age = 24.5), none of whom received formal musical training or studied music in a systematic manner. The participants in Groups A and B had a minimum of five years of music training and had diverse backgrounds, including study of classical music, rock, metal, folk, and jazz. Group A participants studied the following instruments: guitar ($n = 16$), voice ($n = 9$), violin ($n = 3$), clarinet ($n = 3$), drums ($n = 2$), saxophone ($n = 2$), trumpet ($n = 1$), cello ($n = 1$), oboe ($n = 1$), accordion ($n = 1$), and French horn ($n = 1$). Each group was split into four groups of ten participants to learn the melodies under a particular condition. Tables 1 and 2 show age and musical experience for each subgroup. It should be noted that the pianists in Group B received

more years of music training than the musicians in Group A (median for pianists = 15; median for other musicians = 9). We report on age and training effects across groups in the Results section.

MATERIALS

The stimuli for the experiment included four tonally ambiguous piano melodies and three variations of each melody, providing a total of sixteen melodies (four correct melodies + twelve variations). The stimuli were performed by a professional pianist and recorded at the Department of Music of the University of Sheffield using Cubase®. The four original piano melodies were taken (entirely or with slight modifications) from C. Ives' *114 Songs for Voice and Piano* (1922) and lasted between 10 and 15 s each. The melodies were all made of five notes, allowing the participants to use a single hand position on the keyboard, pressing one key with each finger. The four stimuli that participants learned are shown in Figure 1.

The other stimuli were melodic or rhythmic variations of the original melodies. Figure 2 shows an example of how a melody (the first one in Figure 1) was varied.

The stimuli were presented to participants over headphones at a comfortable volume. Participants performed on a Yamaha keyboard (model SY22) without weighted keys and watched a video that indicated the key to be performed while playing with or without sound (red). This video was preceded by instructions that showed the hand position for the melody for 15 seconds, displaying the right or left hand depending on the handedness of the participants. Participants in the seeing condition saw the hand of the professional pianist performing the melodies instead of the highlighted keys. Recordings were made with a video camera placed above the piano, giving an aerial view of the performer's hand (Figure 4).

PROCEDURE

The experiment involved three different phases. In the *familiarization phase* (1) all participants listened to each stimuli once at a comfortable volume. Afterwards, during the *learning phase* (2) participants went through their specific learning condition. Subgroup A learned the melodies through listening and playing: participants were comfortably seated in front of a keyboard, which was placed by the screen of a computer. They listened to a melody and then repeated it on the keyboard while watching the screen for instructions about what key to press. Subgroup B learned the melodies through listening and silent playing. The procedure was the same as the playing condition except that participants did not

TABLE 1. *Median and range of age per participant group*

	Non-musicians		Musicians		Pianists	
	Median	Range	Median	Range	Median	Range
Control	23	19 – 33	26.50	19 – 38	25.50	21 – 48
Playing	24.50	18 – 67	25	18 – 58	25	19 – 31
Silent	23	18 – 41	19	18 – 25	24	20 – 60
Seeing	25.50	22 – 38	26	21 – 36	27.50	18 – 43

TABLE 2. *Median and range of years of musical training per participant group*

	Non-musicians		Musicians		Pianists	
	Median	Range	Median	Range	Median	Range
Control	0	0	10	6 – 15	16.50	5 – 35
Playing	0	0	10	5 – 20	14	8 – 30
Silent	0	0	6	4 – 11	16.50	5 – 45
Seeing	0	0	9.50	5 – 21	12.50	3 – 40

FIGURE 1. The four stimuli used in the experiment.

receive auditory feedback while playing on the keyboard. A similar procedure was followed for subgroup C, who learned the melodies by listening to the melodies and watching a video of the pianist's hand playing the melodies. Finally, subgroup D (control) alternated hearing the melodies with 15-20 s of silence. This process was repeated two times for each melody. In other words, each participant from subgroups A, B, and C was exposed to a melody four times during the learning phase (in additional to hearing it once during the familiarization phase). This exposure consisted of an alternation between hearing the melody *and engaging with the melody* in the way determined by the learning condition (Figure 5). Participants in the control group only received the auditory aspect of the learning phase and listened to each melody twice. The order of learning the melodies was randomized across participants. After a 10-min break, participants' memory for the melodies

was tested in the *recognition phase* (3). They listened to all 16 melodies (variants and original melodies) in random order and were asked to judge after each melody whether it was a melody they had learned previously or not. A computer-based questionnaire was given to the volunteer to respond Y (yes) or N (no) after each stimulus.

Results

PRELIMINARY ANALYSIS

We first statistically compared the level of music training of the three participant groups across the four learning conditions. Level of music training was first log-transformed to make it normally distributed. The value 1 was added to the raw music training values because the log of 0 cannot be taken. A univariate analysis of variance was run with learning condition and group as

FIGURE 2. The three variations used for the first melody.

FIGURE 3. Playing and silent playing condition; highlighted keys to assist the playing task.

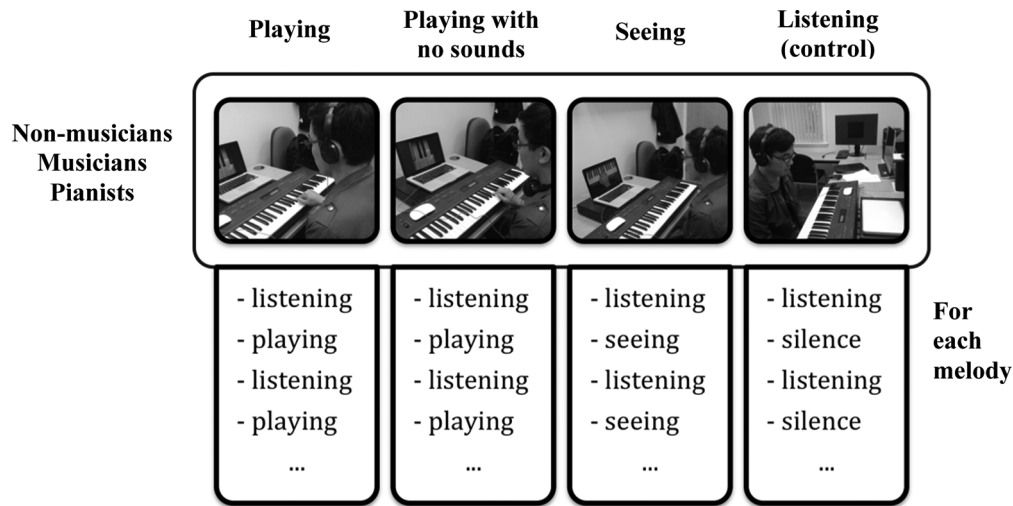**FIGURE 4.** Seeing condition; aerial view of the performer's hand.

	Playing	Playing with no sounds	Seeing	Listening (control)	
Non-musicians Musicians Pianists					**For each melody**
	- listening - playing - listening - playing ...	- listening - playing - listening - playing ...	- listening - seeing - listening - seeing ...	- listening - silence - listening - silence ...	

FIGURE 5. Learning phase illustrated per condition.

between-subjects variables and the log of years of music training as the dependent variable. The effect of group was highly significant, $F(2, 108) = 705.42$, $p < .001$, $r = .96$, with a significant difference between nonmusicians and the other two groups ($p < .001$), and a significant difference between pianists and musicians ($p < .001$, using Bonferroni corrections for multiple testing). Pianists had more years of music training (median = 15) than the other musicians (median = 9). There was no main effect of learning condition, $F(3, 108) = 0.84$, $p = .48$, and no significant interaction between the effects of group and learning condition, $F(6, 108) = 1.92$, $p = .09$. Although this interaction was not significant, we nevertheless conducted follow-up tests to examine whether music training was comparable across conditions within each participant group. For nonmusicians and pianists, this was the case, $F(3, 108) = 1.00$, $p = .40$, for nonmusicians, $F(3, 108) = 0.63$, $p = .60$, for pianists. For musicians, however, the effect remained significant after the log transformation to music training, $F(3, 108) = 3.29$, $p = .03$, $r = .46$).

A post hoc test using Bonferroni corrections showed that music training in the silent playing condition was relatively low compared to the other conditions. The difference of the silent playing condition with the playing and seeing conditions approached significance ($p = .06$ and $p = .09$, respectively). These analyses show that overall the groups and conditions were comparable in music training, with the potential exception of the silent playing condition for the musicians, which was marginally significantly different from the playing and seeing conditions for this group in music training. A similar analysis was run to compare the age across groups and conditions. Age was first log-transformed to correct for skewness of the data and make it normally distributed. A univariate analysis of variance was run with group and condition as independent variables and log-transformed age as the dependent variable. The test showed no significant effect of group, $F(2, 108) = 0.58$, $p = .56$, or of condition, $F(3, 108) = 1.64$, $p = .16$, and no significant interaction between the effects of group and condition, $F(6, 108) = 1.37$, $p = .23$.

EXPERIMENT RESULTS

Mean proportion of correct responses across all 16 melodies was calculated per participant. This procedure led to a normally distributed dependent variable (Skewness = 0.30, $SE = 0.22$; Kurtosis = 0.03, $SE = 0.44$). A univariate analysis of variance was conducted with learning condition and group as between-subjects variables and the mean proportion correct as the dependent variable. The log-transform of years of music training was included as covariate. This analysis showed a significant main effect of condition, $F(3, 107) = 5.38$, $p = .002$, $r = .36$. There was no significant effect of group, $F(2, 107) = 0.27$, $p = .76$, nor a significant interaction between group and condition, $F(6, 107) = 0.08$, $p = .99$. There was no significant association between proportion correct and years of music training, $F(1, 107) = 0.30$, $p = .59$. The same results are obtained if music training is not included as covariate, i.e., the only significant effect is that of learning condition ($p = .001$). The effect of group and the interaction between group and condition were nonsignificant ($p > .77$). Given the absence of an effect of music training, post hoc tests were run for the three groups combined (musicians, pianists,

TABLE 3. *Pairwise comparisons between mean proportion correct of each learning condition*

	Playing		Silent		Seeing	
	Difference	*p*	Difference	*p*	Difference	*p*
Silent playing	.034	1.000				
Seeing	−.059	.564	−.093	.054		
Control	−.095	.046	−.130	.002	−.036	1.000

TABLE 4. *Mean and standard error of the mean proportion correct per condition for each musicianship group*

	Non-musicians		Musicians		Pianists	
	Mean	SE	Mean	SE	Mean	SE
Control	.538	.048	.548	.035	.569	.034
Playing	.644	.047	.656	.040	.638	.046
Silent playing	.669	.048	.688	.060	.688	.040
Seeing	.569	.041	.581	.040	.613	.029
Overall	.605	.024	.617	.023	.627	.019
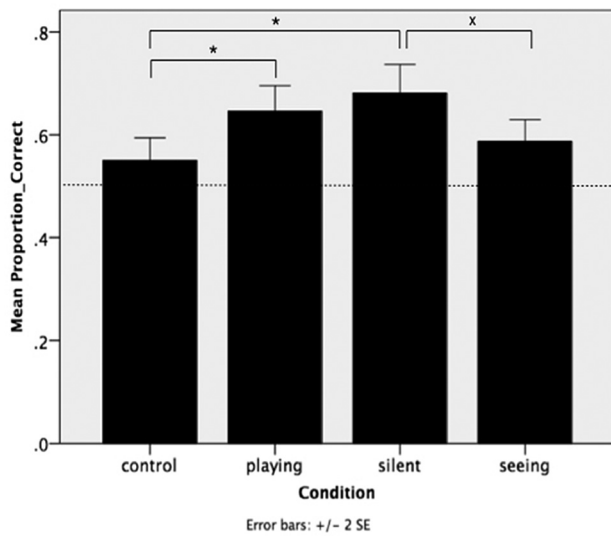
FIGURE 6. Mean proportion correct per condition. Conditions that are significantly different from each other are marked (*). X refers to an effect that is marginally significant (< .06). Chance level is indicated with a dashed line.

and nonmusicians). A pairwise post hoc test was conducted using Bonferroni corrections for multiple comparisons to investigate the differences between the learning conditions (see Table 3). This analysis showed that the mean proportion correct was lower for the control condition than the playing ($p = .05$) and silent playing ($p = .002$) conditions. It also revealed that the scores in the silent playing condition were marginally better than in the seeing condition ($p = .05$). The means and standard error per condition are shown in Figure 6. Each of these means is significantly different from chance level of .50, as tested with a one-sample *t*-test, $t(29) = 2.26$, $p = .031$ for the control condition, $t(29) = 5.89$, $p < .001$ for the playing condition, $t(29) = 4.19$, $p < .001$ for the seeing condition, and $t(29) = 6.51$, $p < .001$, for the silent condition.

These results show a clear rank ordering of the learning conditions; conditions involving a motor activity (silent playing and playing) led to better recognition

performance than the more passive conditions. Importantly, this pattern is present irrespective of musicianship group, as is apparent from the nonsignificant interaction between group and condition. Table 4 provides the mean values per condition for each group, showing the same rank order for each participant group. It is interesting to note that musicians' and nonmusicians' scores are equal in providing correct answers (recognizing the correct melodies and naming the wrong ones as incorrect). Nor did the pianists show any advantage in learning the melodies. Our data suggest that playing music or silent playing gives the participants a greater chance of learning the melodies regardless of musical expertise, at least when it concerns the learning of tonally ambiguous melodies.

These results were obtained by taking the average proportion of correct responses across melodies and ignoring whether participants correctly rejected variations of the original or correctly accepted the original. To examine whether there was a difference between correctly rejecting or correctly accepting, we compared the mean proportion correct for original melodies and for variations of the melodies using a paired-samples *t*-test. The greater number of variations than originals may have led to a response bias to reject the melodies as original (which would lead to a higher correct response in general), or rather, to falsely accept variations of the melodies (which would have led to a lower correct response, in particular for the variations). The paired-samples *t*-test showed no significant difference in mean proportion correct for original melodies and variations of melodies, $t(119) = −0.53$, $p = .60$. In both cases, the mean proportion correct was just above .60 (.61 for originals and .62 for variations).

Additionally, we examined whether some melodies were better recognized than others. A repeated-measures analysis of variance was run with melody as the independent variable and mean proportion correct as the dependent variable. The effect of melody on proportion correct was marginally significant, $F(3, 357) = 2.63$,

TABLE 5. *Mean proportion correct for each version of each melody*

		Mel.1		Mel.2		Mel.3		Mel.4	
Version	N	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Original	120	.583	.045	.600	.045	.6333	.044	.608	.045
Variation 1	120	.642	.044	.583	.045	.492	.046	.675	.043
Variation 2	120	.742	.040	.675	.043	.592	.045	.625	.044
Variation 3	120	.550	.046	.642	.044	.550	.046	.667	.043

$p = .05$ (using Greenhouse Geisser's correction for violations of sphericity). Pairwise comparisons using Bonferroni correction for multiple comparisons showed that only the difference between the mean proportion correct of melody 3 and 4 approached significance ($p = .08$). The mean proportion correct for melody 3 was marginally worse ($M = .57$, $SE = .02$) than for the other melodies (which had values above .63), in particular melody 4 ($M = .64$, $SE = .02$).

Finally, Table 5 shows the mean proportion correct for each variation of each melody. It shows that for melody 2 and 4 the mean proportion correct were highly comparable across variations and the original. Melody 1 showed a relatively high correct response to variation 2, while melody 4 showed a relatively low correct response to variation 1. Although undesirable, it seems unlikely that these particular variations in correct response have influenced the main results, which are based on the mean proportion correct across melodies and versions.

Discussion

The results of the study confirm the hypothesis that motor engagement facilitates the learning of unfamiliar melodies. The measurements of nonmusicians, pianists, and other musicians did not show systematic differences, and the benefits of learning through motor-based conditions emerged. Somewhat unexpected is the result that this is true even without auditory feedback, as in the silent playing condition. Interestingly, our hypothesis that pianists may show an advantage above the other participants was also not borne out, showing equal advantage of motor engagement for all participants. One could interpret these results in terms of memory being facilitated by multimodal engagement with musical stimuli (Aiello & Williamon, 2002). Memory performance is improved if more modalities are engaged in learning (Chen & Fu, 2003; Shams & Seitz, 2008), which was the case for the silent playing and playing conditions. These were arguably also the more challenging conditions, suggesting that the need to

engage with the task may have played a role as well. Moreover, the data also seem to point to a more "embodied" interpretation, in which the nature of the memorization processes might be reconsidered. Instead of perceiving, coding, storing, and representing, we could think of these results in more dynamic terms (van Gelder, 1998), where what truly impacts a participant's ability to remember the stimuli is active sensorimotor engagement with music (Engel et al., 2012). In order to make sense of the melodies, and then to recognize them among others, a participant needs to be motorically linked to the musical stimuli (Molnar-Szacaks & Overy, 2006). The newly acquired motor knowledge gained through playing and silent playing conditions provides participants with a stronger modality to engage with the musical environment, above conceptual, conscious reflection (Maes, Leman, Palmer, & Wanderley, 2014; Schiavio, 2014b).

With this in mind, note that these data might not fit into more traditional paradigms of music cognition, often aligned with cognitivist and representational approaches to mind. As these perspectives (broadly speaking) tend to emphasize computational brain-bound specific structures, which physically implement higher-level cognitive abilities, they may not fully appreciate the direct impact motor experience has on our musical understanding. Scholars operating within such cognitivist paradigm, indeed, may predict that in a similar task to our study musicians — thanks to their knowledge of musical structures — would perform significantly better than nonmusicians, using strategies such as anticipation and grouping to facilitate their understanding of the excerpt (see Snyder, 2001). With regard to this point, one could argue that musicians would not actually have any advantage, given the tonally ambiguous nature of the stimuli (Krumhansl, Sandell, & Sergeant, 1987); that is, it would be unclear how a musician would use the theoretical knowledge associated with traditional Western music to remember the stimuli. However, the musicians' familiarity with musical structures in general, with the piano, and with learning or listening strategies in particular, would have certainly played a key role in the task. On the contrary, what

seems to matter for music learning, memory, and auditory recognition is the motor engagement that links learners and music, and not complex representational abilities (Schiavio, 2012). This experiment, therefore, suggests that musicians and nonmusicians are engaged with new musical stimuli without recruiting their theoretical knowledge. Otherwise, musicians (pianists and other musicians) would have performed better compared to nonmusicians. In other words, it seems that the stimuli are best understood in motor rather than conceptual terms. Many nonmusicians, while engaging in the playing or silent playing condition, were unaware of the position of their hand on the keyboard (because of their lack of knowledge of the keyboard itself, and because their attention was more focused on correctly performing non-familiar movements). However, their results were not significantly different from musicians, and they seemed to benefit equally from the motor involvement.

Our work also resonates with Truitt and colleagues's (1997) study based on piano sight-reading. The researchers measured the eye-hand span (temporal distance between seeing the note on the score and playing it on the piano) and the perceptual span (the spatial region from which the agent would extract information, close to the note on the score) in expert and less-expert pianists. The results concerning the latter were surprisingly similar: as in our study, there was not so much difference between the measures of expert and less expert musicians. (The only parameter of eye movement significantly different between groups was the average fixation duration, with the group of more skilled pianists having a shorter duration. The position of the fixation and the probability of fixating on a note, were not significantly different.) According to Jänke (2006), this can be explained by considering the sensorimotor optimization underlying musical performance as belonging to the *motor* side of the process, rather than on the *visual* one. Both experts and novices, indeed, seem to employ the same basic sensorimotor dynamics to actively engage in new musical activities, regardless of their differing expertise. Differences in skilled and less skilled pianists, we might argue, cannot be ascribed to their ability to elaborate and represent complex patterns of visual information more quickly, or — generalizing — to their capacity to form high-level representations of music. In both cases we would expect a significant difference in the *strategies* adopted to read musical scores (i.e., concerning the position of the fixation), or to remember tonally ambiguous melodies.

This being said, we should add that by emphasizing the role of body and actions we certainly do not want to imply that mere muscular, kinematic activity is more important than musical concepts or theoretical knowledge. Instead, we argue against the static metaphoric representation that sees the body as a puppet controlled by the brain (see Di Paolo, Rohde, & De Jaegher, 2007), and that separates the "objective" world from a "subjective," inner mind. Perception and behavior, action and cognition, are instead mutually dependent and deeply intermixed in our wordly experience (Thompson, 2007; Varela et al., 1991). With this interpretation, therefore, we do not suggest that in our study — or in the aforementioned paper by Truitt and colleagues — motor facilitation is independent from experience. Rather, we maintain that the lack of a significant expertise effect is due to the fact that in both experiments trained musicians did not make use of their motor vocabulary of acts. The tasks (involving piano melodies in our paper vs. sight reading in the other paper) were unpredictable and both experts and amateurs had to rely on a newly acquired repertoire of actions, instead of already acquired chains of goal-directed movements. Indeed, by emphasizing the importance of motor and audiovisual learning over theoretical knowledge we do not argue that the latter does not play a role for the learners' development of musical skills. Instead, our data might help to reconsider the *explanatory power* of theoretical knowledge when it comes to music learning. This interpretation also resonates with the work of Palmer and Meyer (2000), who argue that mental plans for actions and actual movements become independent from each other only at later stages of learning. This shows that the dynamics of skill acquisition are governed by structures that are primarily distributed across the body of the subject, and thus not reducible to processes only "in the head" from the beginning.

This said it is important to keep in mind that the study has a number of methodological limitations including a small number of participants in a single condition ($n = 10$). This may have increased the chance of a null result for the interaction between music training and learning condition. However, each population group showed the same rank ordering of learning conditions (control, seeing, playing, silent playing), confirming the robustness of the effect of learning condition across groups. Another methodological concern relates to a possible response bias caused by using three false melodies for every original melody. However, the correct response means per version did not show a response bias towards false acceptances or towards greater rejections of the original, as we report above. Concerning the main effect of learning condition on the learning of the melodies, it may be argued that the conditions including a motor action and a visual component may have been more engaging or might have

attracted more attention than the control condition. This might partly explain the advantage of these learning conditions over the control condition. On the other hand, these conditions may have led to a greater cognitive load, which would have lowered the performance on the recognition tests. We do not have a way of controlling for these differences across conditions. Measuring arousal and engagement as a covariate would be useful for future verification. We observe that learning was enhanced by having these additional modalities.

Conclusion

The key findings of our study are: 1) the absence of a systematic difference in recognition ability between nonmusicians, pianists, and other musicians, and 2) the benefit of learning melodies with a motor component. These results are compatible with the "embodied" perspective we briefly discussed, and may inspire a deeper understanding of what music learning entails. We maintain that experiencing and learning music can be seen as processes fundamentally immersed in the dynamics of action (Gerson, Schiavio, Timmers, & Hunnius, 2015; Krueger, 2011; Leman, 2007; Menin & Schiavio, 2012; Schiavio, 2014b; Schiavio & Høffding, 2015),, and not reducible to acquisition, elaboration, and development of complex mental representations fully realizable in the head (see Gruhn, 2006). We argue that this perspective downplays the role of action for musical experience, and creates a separation between inner processing and outer behavioral events. In contrast, our view suggests that musical learning is best understood as an embodied coupling that dynamically engages subjects and the musical environment. A learner (expert or amateur) needs to participate in, explore, and modify the sonic world, gaining a first-hand experience of it. As such, learning music is not a modification of the behavioral output only (in light of more conceptual clarity of a given music-related representation), but rather a change in the whole brain-body-world system (see Kelso, 1995). We are confident that this may represent the starting point for a paradigm shift in the psychology of music, as has begun already in cognitive science (Thompson, 2005). Particularly, this study contributes to the existing body of knowledge highlighting the benefit that musically naïve and experienced subjects can obtain through active engagement with music, relative to more passive forms of music learning (Bowman, 2004; Keebler, Wiltshire, Smith, Fiore, & Bedwell, 2014; Schiavio & Cummins, 2015). The memorization of a melody is facilitated by the cross-modal patterns of active behavior. Although more research is needed to understand the cognitive structures underlying these processes, an embodied approach to music cognition represents a promising turning point from traditional representationalist models to describe human musical perception, learning, and experience.

Author Note

Correspondence concerning this article should be addressed to Andrea Schiavio, Department of Music, The University of Sheffield, 34 Leavygreave Road, Sheffield S7 2RD, United Kingdom. E-mail: a.schiavio@sheffield.ac.uk

References

Aiello, R., & Williamon, A. (2002). Memory. In R. Parncutt & G. McPherson (Eds.), *The science and psychology of music performance: Creative strategies for teaching and learning* (pp. 167-181). Oxford, UK: Oxford University Press.

Bangert, M., & Altenmüller, E. (2003). Mapping perception to action in piano practice: A longitudinal DC-EEG study. *BMC Neuroscience, 4*, 26-36.

Bowman, W. (2004). Cognition and the body: Perspectives from music education. In L. Bresler (Ed.), *Knowing bodies, moving minds: Toward embodied teaching and learning* (pp. 29-50). Netherlands: Kluwer Academic Press.

Brown, R. M., & Palmer, C. (2012). Auditory-motor learning influences auditory memory for music. *Memory and Cognition, 40*, 567-578.

Chemero, A. (2009). *Radical embodied cognitive science.* Cambridge MA: MIT Press.

Chen, G., & Fu, X. (2003). Effects of multimodal information on learning performance and judgment of learning. *Journal of Educational Computing Research, 29*(3), 349-362.

Clarke, E. (2005). *Ways of listening: An ecological approach to the perception of musical meaning.* Oxford, UK: Oxford University Press.

Colombetti, G. (2014). *The feeling body: Affective science meets the enactive mind*. Cambridge MA: MIT Press.

Di Paolo, E., De Jaegher, H., & Rohde, M. (2010), Horizons for the enactive mind: Values, social interaction, and play. In J. Stewart, O. Gapenne, & E. Di Paolo, Ezequiel (Eds.), *Enaction: Towards a new paradigm for cognitive science* (pp. 33-87). Cambridge MA: MIT Press.

Drost, U. C., Rieger, M., Brass, M., Gunter, T. C., & Prinz, W. (2005). Action-effect coupling in pianists. *Psychological Research, 69*(4), 233-241.

Eitan, Z., & Timmers, R. (2010). Beethoven's last piano sonata and those who follow crocodiles: Cross-domain mappings of auditory pitch in a musical context. *Cognition, 114*, 405-422.

Engel, A., Bangert, M., Horbank, D., Hijmans, B. S., Wilkens, K., Keller, P. E., & Keysers, C. (2012). Learning piano melodies in visuo-motor or audio-motor training conditions and the neural correlates of their cross-modal transfer. *NeuroImage, 63*, 966-978.

Gallagher, S. (2005). Dynamic models of body schematic processes. In H. De Preester & V. Knockaert (Eds.), *Body image and body schema* (pp. 233-250). Amsterdam: J. Benjamins.

Gallagher, S., & Zahavi D. (2008). *The phenomenological mind*. New York: Routledge.

Gerson, S., Schiavio, A., Timmers, R., & Hunnius, S. (2015). Active drumming experience increases infants' sensitivity to audiovisual synchrony during observed drumming actions. *PLoS ONE, 10*(6), e0130960.

Gruhn, W. (2006). Music learning in schools: Perspectives of a new foundation for music teaching and learning. *Action, Criticism, and Theory for Music Education, 5*(2), 2-27.

Haslinger, B., Erhard, P., Altenmüller, E., Schroeder, U., Boecker, H., & Ceballos-Baumann, A. O. (2005). Transmodal sensorimotor networks during action observation in professional pianists. *Journal of Cognitive Neuroscience, 17*(2), 282-293.

Haueisen, J., & Knösche, T. R. (2001). Involuntary motor activity in pianists evoked by music perception. *Journal of Cognitive Neuroscience, 13*(6), 786-792.

Hurley, S. (1998). *Consciousness in action*. Cambridge, MA: Harvard University Press.

Jänke, L. (2006). From cognition to action. In E. Altenmüller, J. Kesselring, & M. Wiesendanger (Eds.), *Music, motor control and the brain* (pp. 25-36). New York: Oxford University Press.

Keebler, J. R., Wiltshire, T. J., Smith, D. C., Fiore, S. M., & Bedwell, J. S. (2014). Shifting the paradigm of music instruction: implications of embodiment stemming from an augmented reality guitar learning system. *Frontiers in Psychology, 5, 471*. DOI: 10.3389/fpsyg.2014.00471

Kelso, S. (1995). *Dynamic patterns*. Cambridge, MA: MIT Press.

Koeneke, S., Lutz, K., Wustenberg, T., & Jancke L. (2004). Long-term training affects cerebellar processing in skilled keyboard players. *Neuroreport, 15*, 1279-1282.

Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science, 297*, 846-848.

Krueger, J. (2011). Enacting musical content. In R. Manzotti (Ed.), *Situated aesthetics: Art beyond the skin*. Exeter, UK: Imprint Academic.

Krueger, J. (2014). Affordances and the musically extended mind. *Frontiers in Psychology, 4*, 1-13.

Krumhansl, C. L., Sandell, E. J., & Sergeant, D. C. (1987). The perception of tone hierarchies and mirror forms in twelve tone serial music. *Music Perception, 5*, 31-78.

Lahav, A., Salztman, E., & Schlaug, G. (2007). Action representation of sound: Audiomotor recognition network while listening to newly acquired actions. *Journal of Neuroscience, 27*, 308-314.

Lakoff, G., & Johnson M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to Western thought*. New York: Basic Books.

Leman, M. (2007). *Embodied music cognition and mediation technology*, Cambridge, MA: MIT Press.

Maes, P. J., Leman, M., Palmer, C., & Wanderley, M. M. (2014). Action-based effects on music perception. *Frontiers in Psychology, 4*, 1008. DOI:10.3389/fpsyg.2013.01008

Matyja, J., & Schiavio, A. (2013). Enactive music cognition. Background and research themes. *Constructivist Foundations, 8*(3), 351-357.

Menin, D., & Schiavio, A. (2012). Rethinking musical affordances. *AVANT. Trends in Interdisciplinary Studies, 3*(2), 202-215.

Merleau-Ponty, M. (2002). *Phenomenology of perception* (2nd ed.). London & New York: Routledge.

Molnar-Szakacs, I., & Overy, K. (2006). Music and mirror neurons. From motion to e'motion. *Social Cognitive and Affective Neuroscience, 1*(3), 235-241.

Noë, A. (2004). *Action in perception*. Cambridge, MA: MIT Press.

Novembre, G., & Keller, P. E. (2011). A grammar of action generates predictions in skilled musicians. *Consciousness and Cognition, 20*, 1232-1243.

O'Regan, J. K., & Noë, A. (2001). A sensorimotor approach to vision and visual consciousness. *Behavioral and Brain Sciences, 24*(5), 939-973.

Overy, K., & Molnar-Szakacs, I. (2009). Being together in time: Musical experience and the mirror neuron system. *Music Perception, 26*, 489-504.

Palmer, C., & Meyer, R. K. (2000). Conceptual and motor learning in music performance. *Psychological Science, 11*, 63-68.

Phillips-Silver, J., & Keller, P. E. (2012). Searching for roots of entrainment and joint action in early human interactions. *Frontiers in Human Neuroscience*, Special Research Topic: Brain and Art. DOI: 10.3389/fnhum.2012.00026

Rizzolatti, G., & Sinigaglia, C. (2008). *Mirrors in the brain. How our minds share actions and emotions*. Oxford, UK: Oxford University Press.

Schiavio, A. (2012). Constituting the musical object. A neuro-phenomenological perspective on musical research. *Teorema, 31*(3), 63-80.

Schiavio, A. (2014a). *Music in (en)action. Sense-making and neurophenomenology of musical experience* (Unpublished doctoral dissertation). The University of Sheffield, Sheffield, United Kingdom.

Schiavio, A. (2014b). Action, enaction & inter(en)action. *Empirical Musicology Review, 9*(3-4), 254-262.

Schiavio, A., & Cummins, F. (2015). An inter(en)active approach to musical agency and learning. In R. Timmers, N. Dibben, Z. Eitan, R. Granot, T. Metcalfe, A. Schiavio, V. & Williamson (Eds.), *Proceedings of the International Conference on the Multimodal Experience of Music*. Sheffield, UK: HRI Online Publications. Available online at: http://www.hrionline. ac.uk/openbook/chapter/ICMEM2015-Schiavio

Schiavio, A., & Høffding S. (2015). Playing together without communicating? A pre-reflective and enactive account of joint musical performance. *Musicae Scientiae, 19*(4), 366-388

Schiavio, A., Menin, D., & Matyja, J. (2014). Music in the flesh. Embodied simulation in musical understanding. *Psychomusicology. Music, Mind and Brain, 24*(4), 340-343.

Shams, A., & Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in Cognitive Science*, *12*(11), 411-417.

Shapiro, L. (2010). *Embodied cognition*. London, UK: Routledge.

Sheets-Johnstone, M. (2010). Movement: The generative source of spatial perception and cognition. In F. Dolins & R. W. Mitchell (Eds.), *Spatial cognition, spatial perception: Mapping the self and space* (pp. 323-340). New York: Cambridge University Press.

Sheets-Johnstone, M. (2011). *The primacy of movement*. Philadelphia, PA: John Benjamins Publishing.

Snyder, R. (2001). *Music and memory. An introduction*. Cambridge, MA: MIT Press.

Stefan, K., Cohen, L. G., Duque, L., Mazzocchio, R., Celnik, P., Sawaki, L., Ungerleider, L., & Classen, J. (2005). Formation of a motor memory by action observation. *Journal of Neuroscience, 25*(4), 9339-9346.

Thompson, E. (2005). Sensorimotor subjectivity and the enactive approach to experience. *Phenomenology and the Cognitive Sciences, 4*(4), 407-427.

Thompson, E. (2007). *Mind in life: Biology, phenomenology, and the sciences of mind*. Cambridge, MA: Harvard University Press.

Truitt, F. E., Clifton, C., Pollatsek, A., & Rayner, K. (1997). The perceptual-span and eye-span in sight-reading music. *Visual Cognition, 4*, 143-161.

Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., & Rizzolatti, G. (2001). 'I know what you are doing': A neurophysiological study. *Neuron, 32*, 91-101.

Umiltà, M. A., Escola, L., Intskirveli, I., Grammont, F., Rochat, M., Caruana, F., et al. (2008). How pliers become fingers in the monkey motor system. *Proceedings of The National Academy of Sciences, 105*(6), 2209-2213.

van Gelder, T. (1998). The dynamic hypothesis in cognitive science. *Behavioral and Brain Science, 21*(5), 615-628.

Varela, F., Thompson, E., & Rosch, E. (1991). *The embodied mind*. Cambridge, MA: MIT Press

Wilson, A. D., & Golonka, S. (2013). Embodied cognition is not what you think it is. *Frontiers in Psychology, 4*, 58. DOI: 10.3389/fpsyg.2013.00058