

1 **Embodying Melody through a Conducting Baton:**

2 **A Pilot Comparison between Musicians and Non-musicians**

3 Andrea Costantino<sup>1\*</sup>, Nicola Di Stefano<sup>2,3\*</sup>, Fabrizio Taffoni<sup>4</sup>, Giovanni Di Pino<sup>5</sup>, Manuele Casale<sup>1</sup>,  
4 Flavio Keller<sup>3,6</sup>

5  
6 *1 Integrated Sleep Surgery Team UCBM - Unit of Otolaryngology – Integrated Therapies in*  
7 *Otolaryngology, Campus Bio-Medico University, Rome, Italy*

8 *2 Department of Philosophy and Cultural Heritage, Ca' Foscari University of Venice, Venice, Italy*

9 *3 FAST - Institute of Philosophy of Scientific and Technological Practice, Campus Bio-Medico*  
10 *University, Rome, Italy*

11 *4 Advanced Robotics and Human-Centred Technologies - CREO Lab, Campus Bio-Medico University,*  
12 *Rome, Italy*

13 *5 Research Unit of Neurophysiology and Neuroengineering of Human-Technology Interaction*  
14 *(NeXTlab), Campus Bio-Medico University, Rome, Italy*

15 *6 Laboratory of Developmental Neuroscience and Neural Plasticity, Campus Bio-Medico University,*  
16 *Rome, Italy*

17 *\* Authors with equal contribution*

26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48

**Corresponding authors:**

**Flavio Keller, MD**, Lab. of Developmental Neuroscience, Campus Bio-Medico University, Rome, Italy.

**Tel.:** +39 06225419175

**E-mail:** [f.keller@unicampus.it](mailto:f.keller@unicampus.it)

**Andrea Costantino, MD**, Integrated Sleep Surgery Team UCBM - Unit of Otolaryngology – Integrated Therapies in Otolaryngology, Campus Bio-Medico University, Rome, Italy.

**Tel.:** +39 06225411489 **Fax:** +39 06 22541520

**E-mail:** [andrea.costantino94@gmail.com](mailto:andrea.costantino94@gmail.com)

**ORCID ID:** 0000-0001-5551-7785

49

## 50 **Declarations**

51 **Funding:** This work was partially funded by the H2020/ICT European project CONBOTS  
52 (“CONnected through roBOTS: physically coupling humans to boost handwriting and music  
53 learning”, Grant Agreement no. 871803, call topic ICT-09-2019-2020).

54 **Conflicts of interest/Competing interests:** All authors certify that they have no affiliations with or  
55 involvement in any organization or entity with any financial interest (such as honoraria; educational  
56 grants; participation in speakers’ bureaus; membership, employment, consultancies, stock ownership,  
57 or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial  
58 interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the  
59 subject matter or materials discussed in this manuscript.

60 **Ethics approval:** The study was approved by the Ethical Committee of University Campus Bio-  
61 Medico (Prot: 11.19 TS ComEt-CBM 08/2019).

62 **Consent to participate:** Informed consent was obtained from all participants before including them  
63 in the study.

64 **Consent for publication:** Consent for the publication of this study was obtained from both  
65 participants and the Ethical Committee of University Campus Bio-Medico.

66 **Availability of data and material:** Data and material are available on request.

67 **Code availability:** Not applicable.

68 **Authors’ contributions:** AC: Study design, data collection and analysis, manuscript development,  
69 review of final manuscript. NDS: Study design, review of final manuscript. FT: Data analysis,  
70 review of final manuscript. GDP: Manuscript development, review of final manuscript. MC: Study  
71 design, review of final manuscript. FK: Study design, music composition, data analysis, manuscript  
72 development, review of final manuscript.

73 **Abstract**

74 Finger-tapping tasks have been widely adopted to investigate auditory-motor synchronization, i.e.,  
75 the coupling of movement with an external auditory rhythm. However, the discrete nature of these  
76 movements usually limits their application to the study of beat perception in the context of  
77 isochronous rhythms. The purpose of the present pilot study was to test an innovative task that  
78 allows investigating bodily responses to complex, non-isochronous rhythms. A conductor's baton  
79 was provided to sixteen healthy subjects, divided into two different groups depending on the years  
80 of musical training they had received (musicians or non-musicians). Ad hoc-created melodies,  
81 including notes of different durations, were played to the subjects. Each subject was asked to move  
82 the baton up and down according to the changes in pitch contour. Software for video analysis and  
83 modelling (Tracker<sup>®</sup>) was used to track the movement of the baton tip. The main parameters used  
84 for the analysis were the velocity peaks in the vertical axis. In the musician group, the number of  
85 velocity peaks exactly matched the number of notes, while in the non-musician group, the number  
86 of velocity peaks exceeded the number of notes. An exploratory data analysis using Poincaré plots  
87 suggested a greater degree of coupling between hand-arm movements and melody in musicians both  
88 with isochronous and non-isochronous rhythms. The calculated root mean square error (RMSE)  
89 between the note onset times and the velocity peaks, and the analysis of the distribution of velocity  
90 peaks in relationship to note onset times confirmed the effect of musical training. Notwithstanding  
91 the small number of participants, these results suggest that this novel behavioural task could be used  
92 to investigate auditory-motor coupling in the context of music in an ecologically valid setting.  
93 Furthermore, the task may be used for rhythm training and rehabilitation in neurological patients  
94 with movement disorders.

95 **Keywords:** movement; music perception; melody perception; pitch contour; auditory-motor  
96 synchronization

## 97           **Introduction**

98    People often spontaneously accompany music listening with body movements. Such movements are  
99    supposed to mirror rhythmic and acoustical features of musical stimuli. An exhilarating example of  
100   this human ability can be found in the movie [“The Great Dictator”](#), in the scene where the barber  
101   (Charlie Chaplin) shaves a customer while moving to the rhythm of Brahms Hungarian Dance No. 5.  
102   Notably, Chaplin’s movements are not only synchronized with the underlying rhythm: Chaplin uses  
103   the razor as if it were a bow striking the string of a violin, thus conveying changes in harmony and  
104   melody. As the action develops, one has the impression that it is Chaplin himself who is producing  
105   the music.

106   The existence of a strong relationship between music and movement is confirmed by the music  
107   notations that are commonly used to describe movement as well, e.g., *lento*, *glissando*, *ritardando*,  
108   *accelerando*, and *fermata*. As a football fan watching a match would mimic the kick with his own  
109   limb while his favourite player is shooting a penalty, a listener may activate the whole audio-motor  
110   loop when listening to music.

111   This hypothesis has been clarified and investigated in the context of the embodied music cognition  
112   theory (Leman 2008, Leman et al. 2018), according to which music-induced bodily responses are  
113   essential to music perception and cognition, and they are coupled with the melodic and rhythmic  
114   qualities of the auditory stimuli (Godøy 2010; Godøy and Leman 2010). For example, when listening  
115   to music, humans typically use gestures to enact the sounds and to move along with them, as if  
116   gestures co-articulate our inner musical experience (Leman 2016).

117   Accompanying musical rhythm with body movements while playing music is very frequent,  
118   particularly among rock/jazz musicians, and can also help in keeping on the beat. However, how we  
119   move also affects the way we interpret a perceived rhythm during listening. For example, the same  
120   repeating 6-beat rhythm can be perceived as two groups (waltz) or as three groups (march). It has

121 been shown that adults and infants who bounce up and down on every second beat report hearing a  
122 march, while those bouncing on every third beat hear a waltz (Phillips-Silver & Trainor 2005).

123 Such a strict correlation between sound and movement is induced from infancy when babies receive  
124 both sound and movement input as parents swing them while singing. However, such auditory-motor  
125 coupling probably begins in prenatal life (Kisilevsky et al. 2004). Therefore, from an early stage of  
126 development, several experiences might stimulate movement and auditory representations to wire  
127 together in the brain (Trainor 2008). This form of tight auditory-motor coupling is strengthened by  
128 practice and may have a predictive function both during music listening and music production  
129 (Novembre & Keller, 2014). In summary, motor systems may have a top-down influence on the  
130 perception of various aspects of music, such as rhythmic patterns (Ross J.M. et al., 2016).

131 From the wide repertoire of human bodily movements, hand and arm gestures are most often used to  
132 interact with the physical and social external world. Hence, they are also thought to be privileged  
133 ways to respond to music (see Leman et al. 2017 for a review). Indeed, both in sound generation and  
134 sound listening, the hand can be considered “*one with the brain*” in mediating and facilitating action  
135 and perception in relation to musical expression (Leman et al. 2017).

136 The repertoire of hand movements involved in producing music, or in responding to it, is almost  
137 endless. However, the studies that investigated hand movement in response to music have mostly  
138 tested two different kinds of responses: finger tapping or hand-arm movements.

139 In typical finger tapping along with a rhythmic stimulus, the hand-finger unit must hit a hard surface  
140 according to the rhythm. This task has been largely used to study sensorimotor synchronization,  
141 defined as the coupling of movement with an external auditory rhythm (see, e.g., Repp 2010).

142 Other types of responses, e.g., those made by continuous and less constrained movements, typically  
143 involve the hand-arm unit. Such movements are capable of expressing a very wide range of acoustic  
144 features, and they are considered more informative than the hand-finger unit regarding the diverse

145 qualities of music stimuli (Leman et al. 2017). For this reason, conductors' gestures have been  
146 recently investigated to study how musical intentions are encoded in these gestures (e.g., Luck et al.  
147 2010). For example, studies have shown that hand-arm gestures might convey information about  
148 energetic and affective states (e.g., happy vs. sad) attributed to a piece of music, the structures that  
149 are perceived in it (e.g., binary vs. ternary) or what is predicted to come next (e.g., dynamic change)  
150 (see Maes et al. 2014).

151 To date, a large body of research has focused on the hand-finger unit, providing significant insights  
152 into the human ability to synchronize to an external auditory rhythm (see, e.g., Repp 2005 and  
153 2013 for reviews). However, the discrete (on/off) and monodimensional nature of finger-tapping  
154 movement limits its range of application to other, more complex aspects of music. For instance,  
155 while finger-tapping tasks have been used to assess the impact of harmony (i.e.,  
156 consonance/dissonance) on movement form and timing (Komeilipoor et al. 2015), they cannot be  
157 used to investigate how listeners respond to melodic pitch contour, i.e., how they respond to pitch  
158 variation over tones.

159 Therefore, in the last two decades, alternative tasks using continuous movements have been  
160 developed to evaluate more complex aspects of music (Repp and Su 2013), such as circle drawing  
161 (Studenka and Zelaznik 2011), body part oscillations (Lagarde and Kelso 2006), and walking  
162 (Chen et al. 2006, Styns et al. 2007, Moumdjian et al. 2018).

163 In the present study, we wanted to investigate musicians' and non-musicians' unconstrained bodily  
164 responses to unfamiliar melodies, searching for differences in rhythm synchronization, and especially  
165 for differences in the movement qualities between the two groups. To this aim, we developed and  
166 introduced a novel task for studying the response of the hand-arm unit to melody perception.

167 Participants were asked to move a conductor's baton up and down while listening to an unfamiliar  
168 sequence of notes of variable duration. The movement of the tip of the baton was tracked with a video

169 camera and analysed. As an important difference with the finger tapping task, to highlight how  
170 participants physically enacted the melody, we focused on the peak velocity of the movement, which  
171 may show a wide range of values and be linked with personal aspects of auditory-motor coupling.

172 We predicted that musicians, because of their extensive training in auditory-motor coupling, would  
173 perform better than non-musicians and, in particular, show a higher degree of coupling with rhythm  
174 and melodic contour. In addition, we were interested in determining whether, in our task, the hand-  
175 arm gesture was driven by anticipatory mechanisms, as occurs in finger-tapping tasks.

176

## 177 **Methods**

### 178 *Participants*

179 This study was carried out in accordance with the principles of the Helsinki Declaration and further  
180 amendments. The behavioural test was performed by 16 healthy adult volunteers. Each subject  
181 underwent liminal pure-tone audiometry before the behavioural test, and only normally hearing  
182 subjects were included in the study. Subjects were considered to have normal hearing function if  
183 the pantonal audiometric threshold was  $\leq 25$  dB. The exclusion criteria were hearing loss, tinnitus,  
184 other ear disorders, and motor function impairment. No participant had to be excluded because of  
185 the abovementioned criteria. The study was approved by the ethical committee of University  
186 Campus Bio-Medico (Prot: 11.19 TS ComEt-CBM 08/2019).

187 Based on their musical skills, the participants were assigned to two different groups, i.e., musicians  
188 (group M; 6 males, 2 females; age:  $23.6 \pm 1.6$ ) and non-musicians (group N; 3 males, 5 females;  
189 age:  $24.2 \pm 0.5$ ).

190 A participant was included in group M if s/he had received professional musical training for at  
191 least five years or if s/he was a professional musician at the time of the test. For group N, musical  
192 training longer than six months in the previous five years was considered an exclusion criterion.



193

194

195 *Musical Stimuli*

196 All tunes were originally composed by one of the authors (F.K.), using Sibelius (Avid  
197 Technology®) run on a MacBook Pro (2.9-GHz Intel® [Santa Clara, CA] Core™ 2 Duo, 4-GB  
198 1067-MHz DDR3, Apple®, Cupertino, CA). Melodies were generated with a synthetic piano from  
199 the Sibelius instrument library (32-bit, 44.1 kHz) and saved as mp3 files. Melodies were composed  
200 using notes of three different durations (half, quarter, eighth) played at 100 bpm. The durations of  
201 the notes therefore corresponded to 1.2 s, 0.6 s and 0.3 s, respectively. Melodies were written in  
202 4/4 time. Three different melodies (labelled “A”, “B”, and “C”) were composed (Fig. 1). All  
203 melodies consisted of 14 measures and contained all three note durations. Melody A consisted of  
204 67 notes, Melody B consisted of 51 notes, and Melody C consisted of 46 notes. The total duration  
205 of each melody was 31.2 s. Each melody consisted of two different phases: the first three measures  
206 had an isochronous rhythm (Fig. 1, underlined); the remaining measures had a non-isochronous  
207 rhythm. The first isochronous phase of each melody was presented for the same length of 7.2 s  
208 with different note durations as follows: Melody A, note duration of 0.3 s; Melody B, 0.6 s; Melody  
209 C, 1.2 s. Each melody was played in three different octaves (1, 2, 3) to cover an overall frequency  
210 range of approximately 400-4000 Hz. In summary, nine different melodies (according to sequence  
211 of notes and octaves) were played to each subject, corresponding to 483 notes (not considering the  
212 final note of each melody), divided as follows according to note duration: 234 of 0.3 s (47.56%),  
213 147 of 0.6 s (29.87%), and 102 of 1.2 s (20.73%). The melodies were pseudo-randomized in the  
214 following sequence to obtain the final stimulus: A1-C2-B3-A2-C3-B1-A3-C1-B2.

215 The melodies were played through the headphones of a portable audiometer (Amplaid® 171),  
216 suitably calibrated according to the isosensation curves (isophonic curves) defined by the ISO  
217 226:2003 standard. During the subsequent data analysis, the investigator did not listen to the

218 melody to avoid its influence on the analysis. The only feedback provided to the investigator was  
219 the switching on of the audiometer screen at the onset of the first note of each melody (Fig. 2).

220

### 221 *Task and experimental setting*

222 A conductor's baton (length 37.5 cm; weight 40 g) was provided to the subject, who was asked to  
223 accompany the musical stimuli by moving the baton. A red polystyrene ball approximately 1 cm  
224 in diameter (weight 1 g) was placed on the tip of the baton. The subject sat comfortably in a chair  
225 in front of a black screen at a distance of approximately 100 cm (Fig. 2A). The screen had the dual  
226 purpose of reducing external confounding factors and providing spatial reference parameters for  
227 calibration (two thin red lines placed at a known distance of 10 cm drawn on the screen). The  
228 subject's upper limb was not restrained by any means so that all limb muscles were involved in the  
229 execution of the movement. The subject was free to hold the baton with her/his right or left hand,  
230 and all subjects chose to hold the baton with the right hand. After performing pantonal audiometry,  
231 the subject was instructed to accompany each note of a melody moving the baton to obtain a 1:1  
232 coupling between movements and notes. The subject was asked to move the baton up and down  
233 in the vertical direction as much as possible. Two test melodies were used as pre-test  
234 familiarization to ensure that the subject understood the investigator's instructions. The subject's  
235 actions were video recorded, and the position data of the baton's tip were extracted using software  
236 for video analysis (see the following section).

237

### 238 *Data acquisition and analysis*

239 The task was video recorded with the frontal camera of an iPhone X (Apple®, Cupertino, CA) placed  
240 behind and above the head of the subject (Fig. 2A). The camera was configured in full high-definition  
241 and slow-motion mode (1080p/120 fps). Software for video analysis and modelling ([Tracker®](#))  
242 ([Brown and Cox 2009](#)) running on a MacBook Air (1.4-GHz Intel® [Santa Clara, CA] Core™ i5

243 dual-core a, 4-GB 1600-MHz DDR3, Apple<sup>®</sup>, Cupertino, CA) allowed us to track the position (x, y)  
244 of the red polystyrene ball in each time frame. The measures, expressed in pixels, were transformed  
245 to mm thanks to the reference lines on the black screen in front of the subject. Video traces were  
246 aligned with audio stimuli by marking as time zero the time of the first photogram showing the  
247 audiometer screen switching on. Position data were time-derived to estimate velocity (v). According  
248 to the instruction to move the baton up and down in the vertical direction as much as possible (i.e.,  
249 along the y axis), the main parameter used for the analysis was velocity on the vertical axis ( $v_y$ ).  
250 Positive velocity peaks were recorded during upward movements (upbeats). An example of the  
251 velocity data provided by the Tracker<sup>®</sup> software is shown in figure 2B. Velocity traces were visually  
252 inspected, and only  $v_y$  peaks with an amplitude greater than 30% of the maximum value measured in  
253 the trial and separated by a minimum interval of 150 ms (i.e., 50% of the duration of the shortest  
254 note) were considered for the analysis. The times at which each peak occurred were stored in a  
255 Microsoft Excel spreadsheet. Statistical analyses were performed using the statistical package  
256 STATA version 13 (StataCorp LP, College Station, TX, USA) and MATLAB<sup>™</sup> version R2017b  
257 (Mathworks, Inc., Natick, MA). Graphs were prepared using GraphPad Prism 6.0 for MacBook Air  
258 (GraphPad, La Jolla, CA) and MATLAB<sup>™</sup> version R2017b.

259 Starting from the first note in each melody and the first velocity peak, a progressive number was  
260 assigned to each note and each velocity peak. The time interval between the onset of note  $i$  and the  
261 following note  $i+1$  was defined as  $\Delta t_n(i)$ . In analogy, the time interval between the velocity peak  $i$   
262 and the peak  $i+1$  was named  $\Delta t_p(i)$  (Fig. 3).

263 To graphically assess the distributions of velocity peaks with respect to the target notes, Poincaré  
264 plots were employed (Fig. 4). In the plot, red circles represent the relation between the duration of  
265 two consecutive notes [i.e.,  $\Delta t_n(i)$  on the x axis vs  $\Delta t_n(i+1)$  on the y axis]. In analogy, blue crosses  
266 represent the relation between two consecutive intervals between velocity peaks [i.e.,  $\Delta t_p(i)$  vs.  
267  $\Delta t_p(i+1)$ ].

268 The area of each red circle is proportional to the total number of note pairs of a given duration, and  
269 the circles on the identity line represent isochronous rhythms.

270 A greater clustering of crosses around the centre of each circle suggests a better match of note duration  
271 and peak interval duration.

272 As a broad measure of the ability to match melody with movement, we employed the difference  
273 between the number of velocity peaks and the number of notes in each trial.

274 To test whether musical expertise, the duration of notes and pitch impacted this ability, we normalized  
275 the data with a normal quantile transformation and ran ANOVA with three factors (Conover, 1999),  
276 i.e., *Groups* (two levels: Musicians vs Non-musicians), *Melody* (three levels: A, B, C) and *Octave*  
277 (three levels: 1, 2, 3) with the difference between number of peaks and number of notes being the  
278 dependent variable. The trials in which the total number of velocity peaks was higher than the  
279 median+3\*standard deviation of the group were removed as outliers (3 out of 144 trials).

280

281 To measure the capacity to move accordingly to the rhythm in the isochronous phase of each melody,  
282 we calculated the root mean square error (RMSE) between the duration of notes  $\Delta t_n$  and the interval  
283 between two consecutive velocity peaks  $\Delta t_p$  as follows:

$$284 \quad RMSE = \sqrt{\frac{\sum_{i=1}^K [\Delta t_n(i) - \Delta t_p(i)]^2}{K}} \quad (1)$$

285 with K being the total number of notes for each isochronous phase.

286 RMSE, normalized with a normal quantile transformation, was employed as the dependent  
287 variable for ANOVA with three factors, *Groups* (two levels: Musicians [M]- vs non-musicians  
288 [N]), *Melody* (three levels: A, B, C) and *Octave* (three levels: 1, 2, 3), which were run to test the  
289 impact of music expertise, duration of notes and pitch on the capacity to move according to the  
290 rhythm.

291

292 Finally, we performed a time window analysis to assess the distribution frequency of the  
293 distribution of  $v$  peaks in the interval between two consecutive notes, both in the isochronous and  
294 non-isochronous phases. The duration of each note was divided into 50-ms time windows, the first  
295 time window starting with the onset of the note, and the last ending with the onset of the following  
296 note. We counted the number of  $v$  peaks falling in each time window.

297 The results of the velocity peak distribution suggested that the peaks falling in the last 150 ms,  
298 instead of being related to the previous note, were related to the following one (see Results and  
299 Discussion).

300 Anticipation was tested for 600-ms and 1200-ms notes by comparing the number of peaks in the  
301 150-ms interval following the onset of the note with the number of peaks in the 150 ms preceding  
302 note onset.

303 To assess any effect of musical training on anticipation, we compared the number of peaks falling  
304 in a given time window and the number of peaks falling out of that time window. Any difference  
305 between the number of  $v$  peaks measured in the two groups was tested with a chi-square test  
306 performed separately for the three different note durations.

307 **Results**

308 Figure 4 shows the Poincaré plots for one subject in group M (Fig. 4A) and group N (Fig. 4B), as  
309 well as the overall results from both groups (Fig. 4C, group M; Fig. 4D, group N). Inspection of  
310 these figures suggested a greater clustering of v peaks around the centre of the red circles in group  
311 M, particularly in the case of isochronous rhythm (red circles along the identity line).

312

313 *Matching the number of notes and peaks*

314 On average, in group M, the number of peaks exactly matched the number of notes in each melody,  
315 while this was not the case in group N. ANOVA run on the difference between the number of  
316 peaks and notes showed a significant main effect of the factor *Groups* ( $p=0.0004$ ). No further  
317 effects or interactions were found to be significant. Table 1 reports the number of notes and the  
318 mean number of v peaks per melody for each group.

319

320 *TABLE 1*

Melody	Number of notes	Number of v peaks	
		<i>Median ± SD</i>	
		GROUP M	GROUP N
A	67	$67 \pm 4.21$	$69 \pm 5.47$
B	51	$51 \pm 3.89$	$53 \pm 7.87$
C	46	$46 \pm 2.60$	$49 \pm 5.14$

321

322

323

324

325

326 *Ability to move according to the rhythm*

327 ANOVA run on the RMSE showed a significant main effect of the factor *Groups* ( $p=0.0274$ ) and  
328 *Melody* ( $p<0.0001$ ). No further effects or interactions were found to be significant. Table 2 reports  
329 the RMSE (mean $\pm$ SD) measured for the groups and melodies (isochronous rhythms).

330

331 TABLE 2

		GROUPS		
		M	N	M+N
MELODY	A (300 ms)	0.079 $\pm$ 0.01	0.089 $\pm$ 0.021	0.084 $\pm$ 0.017
	B (600 ms)	0.103 $\pm$ 0.034	0.176 $\pm$ 0.05	0.140 $\pm$ 0.056
	C (1200 ms)	0.227 $\pm$ 0.124	0.258 $\pm$ 0.166	0.242 $\pm$ 0.142
	A+B+C	0.123 $\pm$ 0.039	0.161 $\pm$ 0.064	

332

333 Considering melodies regardless of the group, RMSEs progressively increased from melody A  
334 (0.084) to melody C (0.242). Comparisons of melodies after Bonferroni correction revealed that  
335 the RMSEs for melody C were significantly higher than those for melody B ( $p=0.0008$ ) and  
336 melody A ( $p<0.0001$ ), and the RMSE of melody B was significantly higher than the RMSE of  
337 melody A ( $p=0.0452$ ). These results may reflect a lower degree of regularity/predictability in  
338 melodies B and C.

339

340

341 *Time window analysis*

342 Figure 5 shows the distribution of v peaks for each note duration.

343 The distribution of peaks in group M for the quarter note (600 ms, Fig. 5B) was bimodal, while  
344 the distribution in both groups M and N for the half note (1200 ms, Fig. 5C) was trimodal. In the  
345 case of quarter notes, v peaks in group M were more frequent in the interval at 150-200 ms and in  
346 the interval at 500-550 ms, while v peaks were less frequent in the interval at 300-350 ms (which  
347 corresponded to half of the note duration). In the case of half notes, both groups presented three  
348 local maxima: at 100-150 ms, 500-550 ms, and 1000-1150 ms for group M and at 150-200 ms,  
349 550-600 ms, and 1100-1150 ms for group N.

350 Both groups (and particularly group M) presented a high frequency of peaks in the last three time  
351 windows (150 ms) before the following note and a very low frequency in the time windows  
352 preceding that 150-ms interval (Fig. 5A and 5B); this suggested that the peaks falling the last 150  
353 ms, instead of marking the previous note, were anticipating the following one.

354 Concerning the eighth note (Fig. 5A), although the v peaks were concentrated in the time windows  
355 towards the end of the interval, especially in group M, its duration was too short to highlight a  
356 clear bimodal distribution, and we preferred not to include it in the following analysis on  
357 anticipation.

358 Anticipation was tested for the 600-ms and 1200-ms notes by analysing the number of v peaks  
359 falling in the time windows immediately preceding and following the note onset time (i.e., +/- 150  
360 ms) using a chi-square test (Fig. 5D and 5E).

361 The results revealed significant differences between groups for both note durations. Group M  
362 showed a larger number of v peaks before note onset time (i.e., a higher tendency to anticipate the  
363 beat) in comparison with group N, both for the 600-ms ( $p < 0.0001$ ) and 1200-ms notes ( $p = 0.01$ ).  
364 No significant differences were observed between groups in the number of velocity peaks falling  
365 after the note onset time.

366

367



368 **Discussion**

369

370 In the present work, we introduced a novel task to test auditory-motor coupling using a conducting  
371 baton and ad hoc-composed melodies with variations in pitch and rhythm that are typical of  
372 classical Western music. We found that the musicians were better at matching movement velocity  
373 peaks with the number and duration of notes and that they were more prone to anticipate the note  
374 with movement than the non-musicians. Below, the main novelties and findings of the work are  
375 discussed in detail.

376

377 The Experimental Paradigm

378 The synchronization task proposed here introduced several novelties compared to the more  
379 common finger-tapping task. Here, we consider common aspects and differences between the two  
380 tasks, focusing on five aspects: movement kinematics, motor control mechanisms, characteristics  
381 of musical stimuli, anticipation, and differences between musicians and non-musicians.

382

383 *Movement Kinematics*

384 The most obvious difference between finger tapping and the task described here is that, while  
385 finger tapping measures the instant when the finger hits the surface, our task is based on an  
386 oscillatory arm movement and uses the velocity peaks to assess auditory-motor synchronization.  
387 However, this difference may not be so crucial, as finger tapping has been conceptualized by  
388 several authors as an oscillation that is interrupted by contact with the surface. According to this  
389 view, the virtual target of synchronization in finger tapping is the velocity minimum at the end of  
390 the oscillation (see Repp 2005 for review). Another difference with finger tapping is the fact that  
391 the present paradigm considers both positive (upbeat) and negative (downbeat) velocity peaks.

392

393 *Motor Control Mechanisms*

394 Whereas in music practice, the hand is often conceived as a hand-finger unit (for example, in piano  
395 playing), in music listening, the hand is often moved as a hand-arm unit (Leman et al 2017,  
396 Amelynck et al., 2014). Therefore, the present task, in which participants listened to music while  
397 performing vertical movements with their forearm, represents a more ecological motor response  
398 to music than finger tapping.

399 Apart from the ability to predict the onset time of the next note, it may be speculated that the  
400 simple rhythmic finger-tapping task has a lot in common with a simple reaction time task where  
401 the subject has to tap in response to an auditory stimulus. Both tasks are simple, conscious and  
402 voluntary, and more importantly, both involve a motor action that is always the same and is pre-  
403 planned in advance. More complex motor control aspects, such as selection, modulation or other  
404 higher-level functions, are likely not needed.

405 In this light, it is reasonable that they share, at least in part, similar neural substrates. Marking the  
406 rhythm by tapping requires that the previously prepared motor programme held in a temporary  
407 buffer is transferred to M1 and that the motoneurons controlling the relevant muscles are activated.  
408 Despite widespread cortical activation in simple reaction time tasks involving the frontal, temporal  
409 and cerebellar cortices (Kansaku et al., 2004), bilateral premotor cortices seem to be the areas  
410 where the motor programme is buffered (Zangrandi et al. 2019).

411 Responding to music with the hand-arm unit is a more open, multidimensional task, where the  
412 conductor moves the baton and physical features of the movement (e.g., speed, direction,  
413 smoothness) encode emotional features of the music. It may be speculated that when music  
414 becomes the immersive environment, dancing and conducting are ways the body becomes the  
415 interpreter of the physical interaction with music. Compared to rhythmic finger tapping, moving a  
416 conducting baton more strictly requires an internal model of the required interaction kinematics  
417 (direction, speed) and dynamics (forces, torque) that have to be translated in the motor command

418 to achieve the desired outcome (Kawato, 1999; Shadmehr and Krakauer, 2008). This hypothesis  
419 is consistent with a deeper involvement of the frontoparietal network and of the cerebellum in our  
420 task compared to simpler rhythmic finger tapping.

421

#### 422 Musical Stimuli

423 Finger-tapping tasks are typically based on an external isochronous rhythm without pitch variation,  
424 i.e., metronome clicks. Moreover, in those tasks, rhythm variation is usually limited to 10-15% of  
425 the baseline stimulus duration (corresponding, in musical terms, to *tempo rubato*). In the present  
426 task, we investigated sensorimotor responses to auditory stimuli with more ecologically valid  
427 musical stimuli consisting of sequences of notes with a definite pitch contour and showing a  
428 greater, more realistic variation (up to 400%) in the duration of contiguous notes. In addition to  
429 more realistically representing the features of music we typically listen to, ecological stimuli are  
430 better candidates to elicit, and therefore investigate, spontaneous responses to music in a laboratory  
431 environment.

432

#### 433 Anticipation

434 Considering melodies regardless of the group, the RMSE measured in the isochronous phase showed  
435 an increase with the duration of the note. This observation is in line with finger tapping studies, in  
436 which an increased inter-onset interval (IOI) leads to a greater negative mean asynchrony (NMA),  
437 i.e., to anticipation that becomes larger with increasing IOI (Zendel et al. 2011; Repp 2008).

438 In agreement with this latter study, the time window analysis revealed a tendency of the v peaks to  
439 fall in the time windows *preceding* the note onset time, particularly in group M (Fig. 6D and 6E).

440 These results suggest that finger tapping and our paradigm, according to the classic NMA, assess the  
441 same phenomenon as anticipation.

442 It should be noted that the interval of 0.3 s is too short to draw conclusions about the actual  
443 "voluntary" anticipation strategy; indeed, it could be a delayed gesture referring to the previous  
444 note (Fischinger 2013), thus reflecting a mere stimulus-response mechanism. However,  
445 anticipation and hence voluntariness was confirmed with the 600-ms and especially the 1200-ms  
446 notes (Fig. 5B-E).

447 It has been hypothesized that this anticipatory strategy represents a form of stimulus-driven  
448 attention (Jones et al. 2002). In classic finger tapping, regular IOIs allow attention to move ahead  
449 of the beginning of the next tone, resulting in NMA. In our test, the tendency to anticipate the next  
450 note seems to be present both in the isochronous and non-isochronous phases of rhythm. This may  
451 indicate that the presence of pitch variations facilitates the prediction of musical rhythms and  
452 stimulates the sense of agency in the listener (see below for a detailed discussion). Further analyses  
453 are needed to better quantify this anticipatory trend in the non-isochronous phase and to relate it  
454 to the stimulus duration. In fact, it seems that movements related to longer IOIs are regulated by  
455 the cognitively controlled timing system (Lewis and Miall 2003; Buhusi and Meck 2005), which  
456 is prone to an important attentional modulation. The possibility of isolating the (anticipatory)  
457 synchronization component with longer stimuli durations could hypothetically represent a way to  
458 isolate the attentional component.

459 Interestingly, anticipatory mechanisms have also been conceived in music perception as a means  
460 to allow for a reversal of the perceptuo-motor sequence, thus making the relationship between  
461 movement and sound even more strict (Leman 2008). In the present task, the coupling between  
462 action and sound may have induced the sense of agency in participants (see, e.g., Synofzik et al.  
463 2008 and Tsakiris et al 2007 for theoretical reviews).

464 Agency, i.e., the feeling of initiating and controlling a physical action of a body part and, more  
465 generally, an event (Braun et al. 2018, Longo et al. 2008) is one of the main determinants of  
466 embodiment. However, embodiment is a multifaceted concept that requires complex estimations  
467 of body part positions and locations so that agency may be independently processed by the human

468 brain with respect to other features of embodiment, such as the sense of ownership (Kalckert and  
469 Ehrsson, 2012).

470 Recently, we weighted the relative impacts of single sensory modalities and of the motor command  
471 in building the representation of the hand along its sensorimotor control loop. Both with and  
472 without the motor command, the sense of ownership was determined by spatiotemporal  
473 congruency of sensory feedback. However, the sense of agency was perceived only when the  
474 active movement of the hand was present (Pinaridi et al. 2020), suggesting motor efference as an  
475 agency prerequisite (Kalckert and Ehrsson, 2012).

476 If we translate this evidence to the present work, we may hypothesize that embodying the music  
477 goes through the behaviour, adopted especially by musicians, of moving the baton as if participants  
478 were producing the sounds through their own movement. In this light, responding to music through  
479 arm movements implies that the motor act not only precedes or anticipates the sound event in  
480 music but also announces the sound through the gesture in advance, thus making perceived sound  
481 the confirmation or the realization of the movement (Leman et al. 2017).

482

483

#### 484 *Differences between Musicians and Non-Musicians*

485 The simplest parameter used to compare the behavioural performances between the two groups was  
486 the difference between the total number of v peaks and the number of notes. On average, in group M,  
487 the total number of v peaks was equal to the number of notes, while in group N, it was significantly  
488 higher. This “excess” motor activity in non-musicians has also been observed in other paradigms,  
489 e.g., during covert singing: less accurate singers tend to produce more motor activity in laryngeal  
490 muscles when compared with experienced singers (Pruitt et al. 2018). The difference observed in our  
491 protocol can be explained by melodic expectations (Margulis 2005), i.e., the fact that musicians are  
492 able to build a heuristic model for making predictions about the occurrence of the next note grounded

493 on features of the perceived melody (e.g., direction) (see, e.g., von Hippel 2002). This interpretation  
494 is also in line with the Poincaré plots: the greater clustering of time intervals between consecutive v  
495 peaks ( $\Delta t_p$ ) around the melodic attractors is an indirect measure of better prediction and tempo  
496 flexibility.

497 Analysis of RMSE (Table 2) showed that musicians were better at matching the time delay between  
498 notes with the delay between peaks. This observation is in line with previous studies with rhythmic  
499 finger tapping. However, our task extended this finding because it involved using the arm-hand unit  
500 and analysing the velocity of movement: both aspects increase the freedom to match not only the  
501 rhythm but also the melody contour and its emotional content.

502         The time window analysis highlighted further differences between the two groups that were  
503 dependent on note duration. Musicians appeared to better anticipate notes of longer duration (0.6  
504 and 1.2 s, Fig. 6 B-E). This observation is in agreement with data by Doelling and Poeppel (2015),  
505 who recorded brain activity using magnetoencephalography (MEG) during passive listening to  
506 different melodies and compared two groups of listeners, those with or without musical training.  
507 Data from non-musicians demonstrated entrainment of cortical activity by musical stimuli (in  
508 particular in the beta frequency range) when subjects were passively listening to musical stimuli  
509 but only with faster tempi. In contrast, musicians also showed evidence of tracking of musical  
510 rhythm at the slowest tempi (below 1 note per second). Since beta oscillations appear to be  
511 involved in sensorimotor synchronization and in predicting new rhythms (Krause et al. 2010), it  
512 may be hypothesized that embodiment of musical rhythm requires beta-oscillatory activity in the  
513 cortical sensorimotor systems. Indeed, anatomical substrates of this ability were found in several  
514 studies of functional brain anatomy in musicians. In particular, the study of Bengtsson et al. (2009),  
515 which investigated functional brain activity in musicians listening to rhythm sequences of varying  
516 temporal predictability, showed that the activity pattern in supplementary motor area (SMA) and  
517 pre-SMA depended on sequence predictability. From a functional point of view, a greater degree

518 of coupling was found between the auditory cortex and the cortical or subcortical motor areas,  
519 while musicians either listened to or played a piano (Bangert et al. 2006; Jäncke 2012).

520 Finally, the M group might have also been advantaged by the experimental task, which is based  
521 on moving a conductor's baton. In fact, musicians who are used to playing in an ensemble are  
522 trained to pay attention to the baton of a conductor, synchronizing their movements with the rest  
523 of the ensemble (Keller 2008). Synchronization between musicians of an ensemble is mediated by  
524 the hand-arm gestures of the conductor (Hallam 2006). It has been suggested that this favoured  
525 functional connection between musical ability and the processing of temporal information could  
526 depend on the strict coupling of auditory and sensory-motor representations in the central nervous  
527 system.

528

#### 529 Future perspectives

530 The present study is consistent with previous paradigms (e.g., finger tapping) evidencing the  
531 phenomenon of anticipation and showing robust differences between musicians and non-  
532 musicians. With a similar protocol, interesting results may be obtained from the comparison  
533 between groups with different kinds of musical expertise that require different hand/arm  
534 techniques, such as conductors vs. musicians or singers vs. musicians.

535 Finally, more ecologically valid and complex music stimuli associated with continuous bodily  
536 movements could also be useful to assess hand-arm responses to music in disease and  
537 rehabilitation. For example, it would be interesting to test a group of patients who are affected by  
538 Parkinson's disease with the present protocol. Indeed, Parkinson's disease impairs sensorimotor  
539 integration (Abbruzzese and Berardelli 2003, Conte, Belvisi et al. 2017), and there is currently a  
540 strong interest in employing sensorimotor synchronization in the rehabilitation of Parkinsonian  
541 patients (see Moumdjian et al. 2018 for a review). Recent work has shown that rhythmic priming  
542 with finger tapping improves gait parameters in Parkinson's patients. Interestingly, this effect was

543 not seen when the patients performed swinging, alternating movements with both arms (Janzen et  
544 al., 2019). One possible explanation for this discrepancy is that while the finger-tapping protocol  
545 stimulates connections between the auditory cortex and cortical hand motor areas, swinging arm  
546 movements may be generated by central pattern generators at the spinal cord level and thus may  
547 be less efficient in reinforcing cortical auditory-motor coupling. Our prediction would be that arm  
548 movements such as those stimulated by moving a conductor's baton should be at least equally  
549 efficient as finger-tapping movements in improving gait parameters.

550

## 551 **References**

552 Bangert M, Peschel T, Schlaug G, et al. Shared networks for auditory and motor processing in professional  
553 pianists: evidence from fMRI conjunction. *Neuroimage* 2006;30(3):917–26.

554 Bengtsson SL, Ullén F, Ehrsson HH, et al. (2009) Listening to rhythms activates motor and premotor  
555 cortices. *Cortex* 45(1):62–71.

556 Braun N., Debener S., Sychala N., Bongartz E., Sörös P., Müller H.H.O., Philippsen A. (2018) The senses  
557 of agency and ownership: A review. *Front. Psychol.* 2018 Apr 16;9:535.

558 Brown D, Cox AJ. Innovative Uses of Video Analysis. *Physics Teacher* 2009;47(3):145–50. Buhusi CV,  
559 Meck WH. What makes us tick? Functional and neural mechanisms of interval timing. *Nat Rev Neurosci*  
560 2005;6(10):755–65.

561 Chen H-Y, Wing AM, Pratt D. The synchronisation of lower limb responses with a variable metronome:  
562 the effect of biomechanical constraints on timing. *Gait Posture* 2006;23(3):307–14.

563 Chen JL, Penhune VB, Zatorre RJ. The role of auditory and premotor cortex in sensorimotor  
564 transformations. *Ann N Y Acad Sci* 2009;1169:15–34.

565 Clayton M, Sager R, Will U. In time with the music: the concept of entrainment and its significance for  
566 ethnomusicology. *European Meetings in Ethnomusicology* 2005;11,:3–142.

567 Conover W.J. (1999) *Practical Nonparametric Statistics*. Third Edition, John Wiley & Sons, New York.



568 Fischinger T. Timing and synchronisation of professional musicians: A comparison between orchestral  
569 brass and string players. Conference Abstract: 14th Rhythm Production and Perception Workshop  
570 Birmingham 11th - 13th September 2013.

571 Fujii S, Schlaug G. The Harvard Beat Assessment Test (H-BAT): a battery for assessing beat perception  
572 and production and their dissociation. *Front Hum Neurosci* 2013;7:771.

573 Godøy RI. *Gestural Affordances of Musical Sound, Musical Gestures*, Taylor & Francis Group. (2010).

574 Godøy, R. I. & Leman, M. *Musical Gestures: Sound, Movement, and Meaning (Paperback)* - Routledge.  
575 (2010).

576 Hallam S. Musicality. In: *The child as musician: A handbook of musical development*. New York, NY,  
577 US: Oxford University Press; 2006. p. 93–110.

578 Jäncke L. The dynamic audio-motor system in pianists. *Ann N Y Acad Sci* 2012;1252:246–52.

579 Jones MR, Moynihan H, MacKenzie N, Puente J. Temporal aspects of stimulus-driven attending in  
580 dynamic arrays. *Psychol Sci* 2002;13(4):313–9.

581 Kalckert A., Ehrsson H.H. (2012) Moving a rubber hand that feels like your own: a dissociation of  
582 ownership and agency. *Front. Hum. Neurosci.* Mar 14;6:40.

583 Kansaku K., Hanakawa T., Hallett M. (2004) A shared neural network for simple reaction time.  
584 *Neuroimage* 22(2):904-911.

585 Kawato, M. (1999) Internal Models for Motor Controls and Trajectory Planning. *Current Opinion in*  
586 *Neurobiology* 9(6): 718-727.

587 Keller PE. Joint action in music performance. In: *Enacting intersubjectivity: A cognitive and social*  
588 *perspective on the study of interactions*. Amsterdam, Netherlands: IOS Press; 2008. p. 205–21.

589 Kisilevsky S., Hains S.M.J., Jacquet A.Y., Granier-Deferre C., Lecanuet J.P. (2004) Maturation of fetal  
590 responses to music (2004) *Dev. Sci.* 7(5):550-559.

591 Kisilevsky, S., Hains, S. M., Jacquet, A. Y., Granier-Deferre, C., Lecanuet, J. P., (2004). Maturation of  
592 fetal responses to music. *Developmental Science.* 7, 5, 550-559.

593 Komeilipoor N, Rodger MWM, Craig CM, Cesari P. (Dis-)Harmony in movement: effects of musical  
594 dissonance on movement timing and form. *Exp Brain Res* 2015;233(5):1585–95.

595 Krause V, Pollok B, Schnitzler A. Perception in action: the impact of sensory information on sensorimotor  
596 synchronization in musicians and non-musicians. *Acta Psychol (Amst)* 2010;133(1):28–37.

597 Lagarde J, Kelso J a. S. Binding of movement, sound and touch: multimodal coordination dynamics. *Exp*  
598 *Brain Res* 2006;173(4):673–88.

599 Leman M, Maes P-J. The Role of Embodiment in the Perception of Music. *Empirical Musicology Review*  
600 2015;9(3–4):236–46.

601 Leman M. *Embodied Music Cognition and Mediation Technology*. (MIT Press, 2008).

602 Leman M., Maes PJ., Nijs L., Van Dyck E. (2018) What Is Embodied Music Cognition?. In: Bader R. (eds)  
603 *Springer Handbook of Systematic Musicology*. Springer Handbooks. Springer, Berlin, Heidelberg

604 Leman, M., Nijs, L., & Di Stefano, N. (2017). On the Role of the Hand in the Expression of Music. In  
605 Bertolaso, M., Di Stefano, N. (eds) *The Hand* (pp. 175-192). Springer, Cham.

606 Lewis PA, Miall RC. Distinct systems for automatic and cognitively controlled time measurement:  
607 evidence from neuroimaging. *Curr Opin Neurobiol* 2003;13(2):250–5.

608 Longo M.R., Schüür F., Kammers M.P., Tsakiris M., Haggard P. (2008) What is embodiment? A  
609 psychometric approach. *Cognition* 107(3):978-998.

610 Luck, G., Toiviainen, P., & Thompson, M. R. (2010). Perception of expression in conductors' gestures: A  
611 continuous response study. *Music Perception*, 28(1), 47-57.

612 Maes, P. J., Leman, M., Palmer, C., & Wanderley, M. (2014). Action-based effects on music  
613 perception. *Frontiers in psychology*, 4, 1008.

614 Margulis, E. H. (2005). A model of melodic expectation. *Music Perception: An Interdisciplinary*  
615 *Journal*, 22(4), 663-714.

616 Moore GP. Piano Trills. *Music Perception: An Interdisciplinary Journal* 1992;9(3):351–9.

617 Moundjian L., Buhmann J., Willems I., Feys P., Leman M. (2018) Entrainment and synchronization to  
618 auditory stimuli during walking in healthy and neurological populations: a methodological systematic  
619 review. *Front. Hum. Neurosci.* Jun 26;12:263.

620 Münte TF, Altenmüller E, Jäncke L. The musician's brain as a model of neuroplasticity. *Nat Rev Neurosci*  
621 2002;3(6):473–8.

622 Novembre G., Keller P.E. (2014) A conceptual view on action-perception coupling in the musicians' brain:  
623 what is it good for? *Frontiers in Human Neuroscience* 8, Aug 21;8:603.

624 Phillips-Silver J, Trainor LJ. Feeling the beat: movement influences infant rhythm perception. *Science*  
625 2005;308(5727):1430.

626 Phillips-Silver J, Trainor LJ. Hearing what the body feels: auditory encoding of rhythmic movement.  
627 *Cognition* 2007;105(3):533–46.

628 Phillips-Silver, J., Trainor, L. J., (2005). Feeling the beat: movement influences infant rhythm perception.  
629 *Science*. 308, 1430-1430.

630 Pinardi, M., Ferrari, F., D'Alonzo, M., Clemente, F., Raiano, L., Cipriani, C., Di Pino, G. (2020).  
631 Doublecheck: A sensory confirmation is required to own a robotic hand, sending a command to feel in  
632 charge of it. *Cognitive Neuroscience*, In press.

633 Pruitt TA, Halpern AR, Pfordresher PQ (2019) Covert singing in anticipatory auditory imagery.  
634 *Psychophysiology* 2019;56:e13297.

635 Rammsayer TH, Buttkus F, Altenmüller E. Musicians Do Better than Nonmusicians in Both Auditory and  
636 Visual Timing Tasks. *Music Percept* 2012;30(1):85–96.

637 Repp BH, Doggett R. Tapping to a Very Slow Beat: A Comparison of Musicians and Nonmusicians. *Music*  
638 *Percept* 2007;24(4):367–76.

639 Repp BH, Su Y-H. Sensorimotor synchronization: a review of recent research (2006-2012). *Psychon Bull*  
640 *Rev* 2013;20(3):403–52.

641 Repp BH. Metrical Subdivision Results in Subjective Slowing of the Beat. *Music Percept* 2008;26(1):19–  
642 39.

643 Repp BH. Musical synchronization. In: *Music, motor control and the brain*. New York, NY, US: Oxford  
644 University Press; 2006. p. 55–76.

645 Repp BH. Sensorimotor synchronization and perception of timing: effects of music training and task  
646 experience. *Hum Mov Sci* 2010;29(2):200–13.

647 Repp BH. Sensorimotor synchronization: a review of the tapping literature. *Psychon Bull Rev*  
648 2005;12(6):969–92.

649 Ross JM, Iversen JR, Balasubramaniam R (2016) Motor simulation theories of musical beat perception.  
650 *Neurocase* 22:6, 558-565.

651 Shadmehr R., Krakauer J. W. (2008) A computational neuroanatomy for Motor Control. *Experimental*  
652 *Brain Research* 185(3): 359-381.

653 Studenka BE, Zelaznik HN. Circle drawing does not exhibit auditory-motor synchronization. *J Mot Behav*  
654 2011;43(3):185–91.

655 Styns, F., van Noorden, L., Moelants, D., & Leman, M. (2007). Walking on music. *Human movement*  
656 *science*, 26(5), 769-785.

657 Synofzik, M., Vosgerau, G., & Newen, A. (2008). I move, therefore I am: A new theoretical framework to  
658 investigate agency and ownership. *Consciousness and cognition*, 17(2), 411-424.

659 Toiviainen P, Luck G, Thompson MR. Embodied Meter: Hierarchical Eigenmodes in Music-Induced  
660 Movement. *Music Perception: An Interdisciplinary Journal* 2010;28(1):59–70.

661 Trainor, L. J., (2008). The neural roots of music. *Nature*. 453, 598-599.

662 Trumpp NM, Kliese D, Hoenig K, Haarmeier T, Kiefer M. Losing the sound of concepts: damage to  
663 auditory association cortex impairs the processing of sound-related concepts. *Cortex* 2013;49(2):474–86.

664 Tryfon A, Foster NE, Ouimet T, et al. Auditory-motor rhythm synchronization in children with autism  
665 spectrum disorder. *Research in Autism Spectrum Disorders* 2017;35:51–61.

666 Tsakiris, M., Schütz-Bosbach, S., & Gallagher, S. (2007). On agency and body-ownership:  
667 Phenomenological and neurocognitive reflections. *Consciousness and cognition*, 16(3), 645-660.

668 von Hippel, P. (2002, May). Melodic-expectation rules as learned heuristics. In *Proceedings of the 7th*  
669 *International Conference on Music Perception and Cognition, Sydney*.

670 Zangrandi, A., A. Mioli, M. D'Alonzo, D. Formica, G. Pellegrino and G. Di Pino (2019). Conditioning  
671 transcranial magnetic stimulation of ventral premotor cortex shortens simple reaction time. *Cortex* 121:  
672 322-331.

673 Zarco W, Merchant H, Prado L, Mendez JC. Subsecond timing in primates: comparison of interval  
674 production between human subjects and rhesus monkeys. *J Neurophysiol* 2009;102(6):3191–202.

675 Zatorre RJ, Chen JL, Penhune VB. When the brain plays music: auditory-motor interactions in music  
676 perception and production. Nat Rev Neurosci 2007;8(7):547–58.

677 Zendel BR, Ross B, Fujioka T. The Effects of Stimulus Rate and Tapping Rate on Tapping Performance.  
678 Music Percept 2011;29(1):65-78

679

## 680 **Table legends**

681 **Table 1:** Data are expressed as the mean  $\pm$  SD. \*  $p$  value statistically significant.

682

## 683 **Figure legends**

684 **Figure 1:** Explanatory music sheet with the three melodies (A-B-C). The underlined part of each  
685 melody represents the isochronous phase.

686 **Figure 2:** Experimental setup with audiometer for generation of music stimuli and movement  
687 recording camera. B. Example of the velocity peaks ( $v$ ) on the vertical axis ( $v_y$ ) provided by the  
688 software Tracker<sup>®</sup>. The time interval between two consecutive peaks ( $\Delta t_p$ ) was defined as shown.

689 **Figure 3:** Definition of the intervals between notes and between velocity peaks (see also Fig. 4)

690 **Figure 4:** Poincaré plots expressing the performance of one subject in group M (**A**) and one in group  
691 N (**B**) and the overall data related to group M (**C**) and group N (**D**). The Poincaré plot visualizes the  
692 correlation between two consecutive data points in a time series. The red circles reported in these  
693 figures are the target intervals given by effective note durations of the nine melodies. The blue  
694 crosses represent the time intervals from velocity peaks.

695 **Figure 5:** Distribution of velocity peaks in 50-ms time windows between two consecutive notes.  
696 The time windows related to the notes of 0.3 s (**A**), 0.6 s (**B**) and 1.2 s (**C**) are represented on the x  
697 axis. The y axis represents the number of  $v$  peaks in the corresponding time window.

698

699