**ORIGINAL ARTICLE**

CrossMark

# Your move or mine? Music training and kinematic compatibility modulate synchronization with self- versus other-generated dance movement

Yi-Huang Su[1] · Peter E. Keller[2]

## Abstract

Motor simulation has been implicated in how musicians anticipate the rhythm of another musician's action to achieve interpersonal synchronization. Here, we investigated whether similar mechanisms govern a related form of rhythmic action: dance. We examined (1) whether synchronization with visual dance stimuli was influenced by movement agency, (2) whether music training modulated simulation efficiency, and (3) what cues were relevant for simulating the dance rhythm. Participants were first recorded dancing the basic *Charleston* steps paced by a metronome, and later in a synchronization task they tapped to the rhythm of their own point-light dance stimuli, stimuli of another physically matched participant or one matched in movement kinematics, and a quantitative average across individuals. Results indicated that, while there was no overall "self advantage" and synchronization was generally most stable with the least variable (averaged) stimuli, motor simulation was driven—indicated by high tap-beat variability correlations—by familiar movement kinematics rather than morphological features. Furthermore, music training facilitated simulation, such that musicians outperformed non-musicians when synchronizing with others' movements but not with their own movements. These findings support action simulation as underlying synchronization in dance, linking action observation and rhythm processing in a common motor framework.

## Introduction

Theories of motor simulation (Jeannerod, 2001) have received much attention in the cognitive (neuro)sciences (Eaves, Riach, Holmes, & Wright, 2016; Gallese & Sinigaglia, 2011; Springer, Parkinson, & Prinz, 2013). While having been defined in somewhat different ways depending on the context, simulation generally refers to the mechanism by which our motor system generates an action covertly, or activates the internal representation of an action, in the absence of a real motor act (Eaves et al., 2016; Springer et al., 2013). One main proposal in this framework is that

✉ Yi-Huang Su
yihuang.su@tum.de

[1] Department of Movement Science, Faculty of Sport and Health Sciences, Technical University of Munich, Georg-Brauchle-Ring 62, 80992 Munich, Germany

[2] Western Sydney University, Sydney, Australia

the sensory and motor systems are coupled in terms of the action representation, and that simulation during action observation (Eaves et al., 2016) would engage similar neural mechanisms in the motor system as actually performing that action. Another relevant proposal is that, in a social setting, simulation allows the observers to use their own motor repertoire to cognitively process, predict, or understand the action of others (Decety & Chaminade, 2003; Macerollo, Bose, Ricciardi, Edwards, & Kilner, 2015; Rizzolatti & Sinigaglia, 2010). As such, observing more familiar movements would lead to more optimal simulation, better action prediction (Keller, Knoblich, & Repp, 2007; Stadler, Springer, Parkinson, & Prinz, 2012), and greater activation in the motor system (Calvo-Merino, Glaser, Grèzes, Passingham, & Haggard, 2005; Jola, Abedian-Amiri, Kuppuswamy, Pollick, & Grosbras, 2012). This framework connects well to the common coding theory (Prinz, 1997), which postulates a shared representation of action planning and its perceptual consequences, through which action and perception are not only linked, but could in fact influence each other (Schütz-Bosbach & Prinz, 2007). When observing an action, the perceived event (i.e., the consequence of that action) may activate the corresponding motor code by means of internal

simulation, the ease of which could in turn influence the relevant motor output (e.g., how one should react to the observed action).

Simulation, or more broadly described by action–perception coupling, has in recent years found increasing support in musical contexts (Maes, Leman, Palmer, & Wanderley, 2014; Novembre & Keller, 2014). One aspect of music cognition, the processing of *rhythm*, seems to be suitably accommodated in the framework of motor simulation (Ross, Iversen, & Balasubramaniam, 2016). The action simulation for auditory prediction (ASAP) hypothesis (Patel & Iversen, 2014) postulates that, upon hearing a musical rhythm and without explicit movement, the motor system internally simulates a periodic body motion that corresponds to the most salient periodicity of the rhythm—its beat. Under this hypothesis, the simulated motion is assumed to be abstract and not necessarily associated with a specific body part. However, the simulated motion could in principle also resemble (or at least be related to) each individual's motor repertoire when moving naturally along with music, as these movements often exhibit periodic patterns that match the underlying periodicity of the rhythm (Burger, Thompson, Luck, Saarikallio, & Toiviainen, 2014; Toiviainen, Luck, & Thompson, 2010). Regardless of the precise nature of the simulated movement, one critical postulate is that such covert motor activity may facilitate beat perception by enhancing temporal *prediction* of the beat. Several findings support this hypothesis by showing that the neural motor system activated by the musical beat during passive listening (Grahn, 2012) appears instrumental to beat perception (Grahn & Brett, 2009). Moreover, cortical motor activity is found to oscillate synchronously to different levels of periodicity in the rhythm (Fujioka, Ross, & Trainor, 2015; Iversen, Repp, & Patel, 2009), suggesting a temporal correspondence between the rhythmic structure of music and the internally engaged motion.

Simulating musical rhythm in a predictive manner is beneficial not only for perception but also for coordinating overt movement with music. In situations such as ensemble musicians playing together or a listener dancing to music, sensorimotor synchronization (SMS) enables individuals to time their musical actions to each other or to the musical rhythm (Repp & Su, 2013). Synchronization entails predicting (instead of reacting to) the upcoming events, the success of which likely depends on how effectively one could internally simulate the time course of the synchronization target (Novembre & Keller, 2014; Novembre, Ticini, Schütz-Bosbach, & Keller, 2014). The effectiveness of simulation may in turn be modulated by how similar the external rhythm is to each individual's own action representation. Indeed, it has been shown that pianists synchronize better with their own previously recorded duet part than with that of another pianist (Keller et al., 2007). Complementary to this finding

is that pianists are also better at detecting (artificially introduced) temporal deviations in their own recordings than in those of other pianists, especially at points of greater stylistic timing differences (Repp & Keller, 2010). These results are consistent with the proposed action simulation mechanism, which predicts more successful simulation—and thus better tuned perception and synchronization—for sensory information that more closely matches one's own motor repertoire.

In the present study, we intended to bring the simulation hypothesis for music and rhythm closer to the motor simulation theory originally proposed for action observation (Eaves et al., 2016; Jeannerod, 2001). We combined the two processes in a single SMS task that may be underpinned by the same simulation mechanism: visually observing a music-related rhythmic action that was previously learned, and synchronizing with the visual rhythm of that action. The action here was tightly connected to musical rhythm: dance (Su, 2016a). The rationale for this implementation was threefold: First, while a number of studies have examined the role of action simulation in how musicians synchronize with auditory stimuli (Keller et al., 2007; Novembre et al., 2014; Novembre, Ticini, Schutz-Bosbach, & Keller, 2012), no SMS study thus far has dealt with this issue in a different scenario and modality. Dance appears to be a perfect ecological action for which motor simulation may be required (Cross, Hamilton & Grafton, 2006) to achieve interpersonal coordination, perhaps in ways similar to joint action between musicians (Keller, Novembre, & Hove, 2014). In addition, given the emerging evidence that synchronization with visual rhythms containing realistic motion may be comparable to that found in the auditory counterpart (Hove, Iversen, Zhang & Repp, 2013; Iversen, Patel, Nicodemus, & Emmorey, 2015; Su, 2016b), it seems reasonable to investigate whether SMS with visual dance stimuli could also be modulated by the degree of simulation. Second, dance as a class of whole-body movement has been shown to effectively engage motor simulation in the observers (Calvo-Merino et al., 2005; Cross, Kraemer, Hamilton, Kelley, & Grafton, 2009; Jola et al., 2012). In the present context, dance stimuli may communicate visual rhythms in a comparable manner as auditory musical rhythms (Su, 2016a; Su & Salazar-Lopez, 2016), and their metrical structure has been found to influence SMS accordingly (Su, 2016b). As such, using dance as visual stimuli in a SMS task has the advantage of tying simulation of the *rhythm* to simulation of the *action*, which strengthens the link between theories of rhythm cognition and motor simulation in an action observation scenario. Finally, dance has the additional advantage that, given relatively simple movement sequences, the action can be produced by non-experts without excessive training, and the investigation is thus not limited to a specific expert population (e.g., instrument playing is only possible with musicians).

We manipulated the expected effectiveness of simulation by varying the agency of the observed movement, following the logic of Keller et al. (2007). Similar to their task of having pianists duet with their own or another pianist's recording, here we asked participants to tap synchronously to the visual rhythm of the same point-light dance movement that had been performed by themselves or by another participant (previously recorded in a motion-capture session). The general hypothesis was that, as self-generated actions contain kinematic cues that are most familiar and compatible with one's own repertoire (Knoblich & Prinz, 2001; Sevdalis & Keller, 2010; Wöllner, 2012), they should lead to stronger motor simulation than other-generated actions (Decety & Chaminade, 2003), which should in turn afford better prediction of the movement rhythm and thus better synchronization. Because measuring SMS with realistic dance stimuli is a rather new attempt (see Su, 2016b), several aspects regarding its simulation warrant investigation. In this study, we intended to address a few relevant issues besides the general "self-advantage" (Keller et al., 2007):

First, as whole-body point-light motion contains rich information in many dimensions, what are the critical cues in these stimuli that drive motor simulation in the observers? We hypothesized two candidates, drawing on the literature of biological motion perception (Matheson & McMullen, 2010; Thompson & Baccus, 2012): "form"—the morphological cues that can be derived from the physical proportion of the humanlike figure (Loula, Prasad, Harber, & Shiffrar, 2005; Sevdalis & Keller, 2011), and "motion"—the kinematic cues that can be extracted especially from the velocity profile of the movement (Knoblich & Prinz, 2001; Su, 2014, 2016b). The similarity of both cues to those of one's own could in principle promote action identification and simulation (Loula et al., 2005; Saygin & Stadler, 2012). To reveal which cue played a more influential role, we presented other-generated motion stimuli that were matched either morphologically or in terms of velocity to self-generated ones, and compared SMS behavior amongst these conditions.

Next, most SMS studies have examined simulation in expert musicians who can produce well-controlled, individualistic expressive timing in music as auditory stimuli (Keller et al., 2007; Novembre et al., 2012, 2014; Repp & Keller, 2010). As the musicians in these cases played without a metronome, the rhythm they produced contained *intentional* idiosyncratic temporal deviations from a prescribed underlying periodicity, which were critical cues for simulation. Here, we asked whether a similar advantage in synchronizing with self-generated stimuli would also generalize to actions without such intentional deviations, given that individual motor styles in everyday movements—albeit unintended—can already modulate action prediction (Koul, Cavallo, Ansuini, & Becchio, 2016). Could *unintentional* timing deviations in the movement lead to a "self-advantage" in synchronization?

To answer this question, we recorded dance movements that were paced by a metronome, and we recruited only non-experts (untrained or non-professionally trained participants) to minimize the presence of expressive timing, which might persist in experts even when paced (Repp, 1999). As we presented rhythmic dance stimuli, it was nevertheless of interest how (non-professional) training experience in music mediated SMS and simulation (Karpati, Giacosa, Foster, Penhune, & Hyde, 2016, 2017). We thus grouped the participants based on the years of music training for between-group comparisons.

Finally, while the simulation hypothesis suggests that stimuli more similar to one's own should benefit SMS more, a previous study has reported better synchronization with prototypical motion that is quantitatively averaged ("morphed") across gestures of different individuals (Wöllner, Deconinck, Parkinson, Hove, & Keller, 2012). Although the morphed motion is devoid of individual characteristics, which might hinder simulation, it has smoother trajectories and lower spatiotemporal variability, which likely facilitates SMS. Based on this observation, we also included a condition of morphed motion, i.e., the grand average across all the individuals for the same dance. We aimed to verify whether such a "morph advantage" could outweigh the morphological or kinematic features of individual motion in the present task.

In this study, participants who had no prior experience with swing dance first took part in a training and recording session, in which they learned to perform the basic steps of the *Charleston* dance paced by a metronome (Su, 2016a, b). They learned to dance two versions of the steps: the original version with the whole body moving, and the modified (Riverdance-like) version where the arms remained still (details see "Method" section, and also Su, 2016a, b). The former allowed more unconstrained movement and individualistic expression than the latter, which could be reflected accordingly in the self-related information available in the movement (Sevdalis & Keller, 2010). The motion-captured data were used to generate the stimuli of the dancing point-light figures (PLF, Johansson, 1973). The same participants returned later to complete a SMS task, in which they tapped their finger synchronously to the rhythm of the leg movement (the "visual beat", see Su, 2016b Experiment 2) of four different point-light dancers: self, another participant matched in physical proportion, another participant matched in the mean maximal velocity of the leg movement, and the morph. We asked the following questions: (1) Is synchronization better with self-generated movement, as previously shown in pianists' auditory SMS (Keller et al., 2007)? (2) Which of the two non-self agents engages action simulation more, and is the effect of agency more obvious for less-constrained movements, i.e., whole-body movement compared to the version without arms moving (Sevdalis &

Keller, 2010)? (3) Would music training modulate the ability to simulate? (4) Is synchronization with the morphed motion indeed superior (Wöllner et al., 2012)?

## Method

### Participants

Twenty-two young volunteers without neurological conditions (seven males, mean age 26.3 years, SD 5.0) took part in this experiment. Participants were naïve about the purpose, gave written informed consent prior to the experiment, and received an honorarium of 8 € per hour in return. Fifteen participants had trained in music and thirteen had trained in dance (none in swing dance) as a hobby, amongst whom eight had learned both. The mean training duration was 4.2 years (SD 4.2) for music and 3.9 years (SD 5.0) for dance. Note that data from four participants were later excluded from the analyses due to variable stimulus beat (see the section "Beat variability" in "Stimulus motion analysis" below), and so the sample size reported in the "Results" section was eighteen. The study had been approved by the ethic commission of Technical University of Munich, and was conducted in accordance with the ethical standards of the 1964 Declaration of Helsinki.

### Design

The experiment consisted of two different sessions: dance recording (termed "Dance session"), and the tapping experiment (termed "Tapping SMS session"). The two sessions were separated by 2–3 months, which served to reduce possible effects of episodic memory for the stimuli used in the tapping experiment.

### Dance session

This session contained two parts: the training part and the recording part. Because the dance session was part of a bigger project involving other subsequent experiments, several different dance steps were learned and recorded, only a subset of which were relevant to the present study. Here, we describe only the movements that were used to generate the point-light stimuli for the tapping session[1].

In the training part, each participant was taught individually by a female instructor to perform the basic steps of the *Charleston* dance. Both male and female participants learned the same pattern of movement. One cycle of the basic steps consisted of eight regular bounces of the trunk (stemming from repetitive flexions and extensions of the knees), during which the legs made four kicking movements (left leg to the back, left leg to the front, right leg to the front, right leg to the back) and the arms also swung four times (both arms swinging simultaneously in the opposite direction, forward or backward). When paced by a metronome, the bounce occurred at every beat (beat 1–8), whereas the legs and the arms swung at every second beat (beat 1, 3, 5, and 7, Fig. 1). See Su (2016a) for a detailed description of the movement sequence. Participants learned the sub-components of the movement additively before proceeding to move all the body parts together (i.e., first the bounce, then the legs coordinated with the bounce, and finally the arms swinging on top of them). Participants learned to dance this sequence synchronously to a metronome at the tempo of 150 BPM, corresponding to an inter-beat interval (IBI) of 400 ms. They practiced until they could dance this sequence fluently and cyclically to the metronome. Once a participant succeeded in performing this basic sequence, he/she proceeded to learn one variation of this dance (the "Riverdance-style"), in which the same movement was carried out with the arms placed upon the hips throughout the sequence (see Su, 2016b). Participants also practiced to dance this variation to the metronome. Once they succeeded with the variation, the training part was complete. Depending on the dance experience of each participant, the training part lasted between half an hour and 1 h (involving other dance sequences not relevant to the present study).

The recording part started after the training part. 13 markers in total were attached to the joints and the forehead of each participant as was conventional for generating point-light motion (Johansson, 1973). Participants danced each of the two movement types (with or without arm movement) paced by the same metronome as during the training part. They danced within a marked space of 3 m × 3 m, and their movements were tracked by a 3-D motion capture system (Qualisys Oqus, 7 cameras) at a sampling rate of 100 Hz. Each movement sequence was performed continuously for at least 60 s of the recording. If the movement became out of synchrony to the metronome (according to the instructor) or interrupted during a recording, the recording was repeated until a continuous, successful sequence of at least 60 s was achieved. Participants practiced each sequence briefly before starting the recording, and took a break between recordings. Depending on the quality of the performance, the recording part took between 20 min to 1 h for a single participant (also involving other dance sequences not relevant here).

---

[1] In total 31 participants were recruited for the dance session, who were later invited to take part in several follow-up experiments. 22 of them took part in the present tapping experiment. Participants signed up for the follow-up experiments voluntarily without being pre-selected.
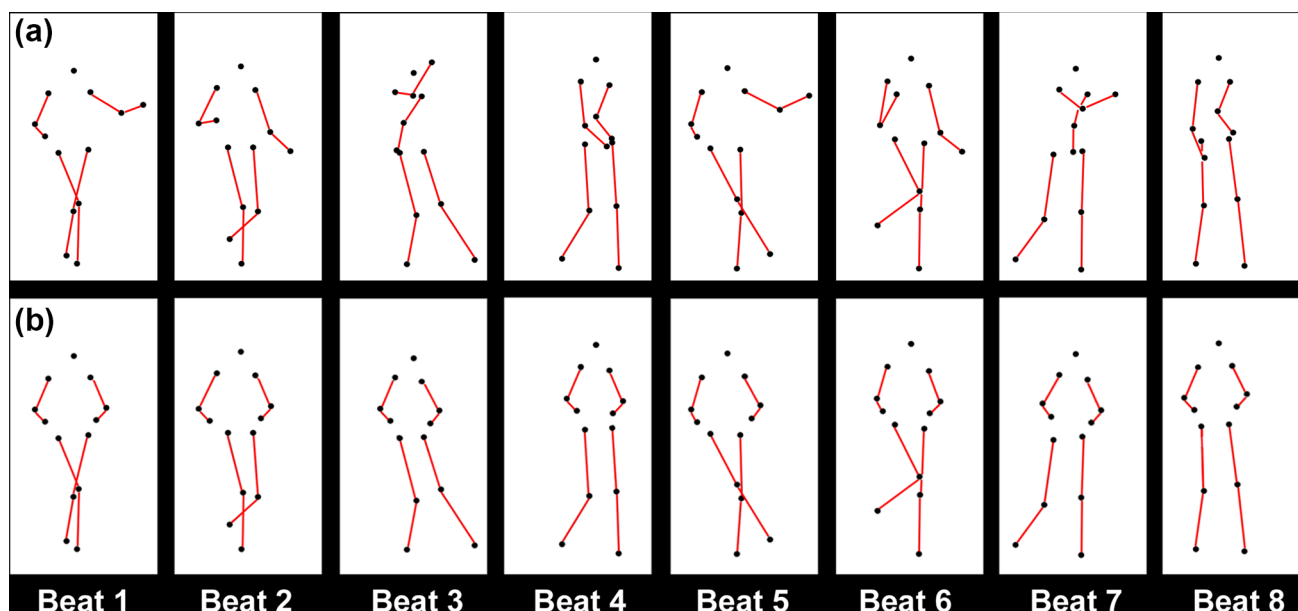
**Fig. 1** Illustration of the point-light stimuli for the SMS task, showing one cycle of the basic steps of the Charleston dance (**a**) with arms and (**b**) without arms. The frames represent the posture at each beat. Participants tapped to the leg movement at beat 1, 3, 5, and 7. The colors are inverted here for the discs and the background, and red lines joining discs are drawn to ease visualization, which were not shown in the experiment

## Tapping SMS session

### Stimuli and materials

The visual stimuli consisted of different PLFs dancing the *Charleston* steps in the original version or in the variation without arm movement, as described above (Fig. 1). The point-light stimuli were generated from the motion-capture data collected in the Dance session. Specifically, for each participant, one best cycle of each movement type (corresponding to eight metronome beats) was selected from the recorded sequence. The procedure and criteria for selecting the best motion cycle followed those described in a previous study (Su, 2016a). In short, the cycle was selected based on the lowest temporal variability of the beat-defining kinematic parameters, namely the end positions and the peak velocities, across the limbs and the trunk. This procedure also further reduced possible idiosyncratic tempo deviations in the movement stimuli.

Next, following the convention of the self-other action recognition paradigm (e.g., Sevdalis & Keller, 2010; Wöllner, 2012), a matching procedure was carried out to select the "other" agent for each participant. Two different matches were created for each participant (termed "*Self*") based on different criteria: The first match, termed "*Other_{PM}*" (*PM* for "physical match"), was selected based on matching gender and physical proportion. The match in physical proportion was defined here as the closest match in the body mass index (BMI), calculated as weight (kg)/height$^2$ (m$^2$). The reason

for using this index, instead of referring only to weight (Sevdalis & Keller, 2010; Wöllner, 2012), was that all the PLFs were going to be scaled to the same height when presented on the screen, and thus the similarity of physical proportion between two agents was better preserved in this manner. The second match, termed "*Other_{KM}*" (*KM* for "kinematic match"), was selected based on matching gender and the closest match in the mean maximal velocity of the leg movements, which have been previously found to define the beat of the dance movement (Su, 2016a, b). The maximal velocity was indicative of the movement amplitude and thus the "beat clarity" embedded in the leg movements (Su, 2014, 2016b). See Table 1 for an overview of the physical and kinematic parameters in the stimuli. Paired t tests showed that, for the physical parameter, the mean absolute difference between *Self* and *Other_{PM}* (~12% of the SD of the *Self* stimuli) was smaller than between *Self* and *Other_{KM}* (~93% of the SD of the *Self* stimuli), $t(17) = -6.06$, $p < 0.001$. Similarly, for the kinematic parameter, the mean absolute difference between *Self* and *Other_{KM}* (~24% of the SD of the *Self* stimuli) was smaller than between *Self* and *Other_{PM}* (> 100% of the SD of the *Self* stimuli), $t(17) = -4.94$, $p < 0.001$.

Finally, a quantitatively averaged version of the point-light motion was generated for each movement type (termed "*Morph*", see Wöllner et al., 2012) across gender. The *Morph* sequence was created by first temporally interpolating the best cycles of all the participants to be the same length, and then computing the mathematical average of the *X, Y,* and *Z* coordinates, respectively, along the time

**Table 1** Means and SDs of the physical and kinematic parameters across stimuli of the 18 participants (*Self*)

| | Physical parameter (BMI) | | | Kinematic parameter (Max. Vel., mm/ms) | | |
|---|---|---|---|---|---|---|
| | Self | Abs. Other$_{PM}$–Self | Abs. Other$_{KM}$–Self | Self | Abs. Other$_{KM}$–Self | Abs. Other$_{PM}$–Self |
| Mean | 22.95 | 0.31 | 2.46 | 3.11 | 0.12 | 0.55 |
| SD | 2.65 | 0.36 | 1.41 | 0.51 | 0.20 | 0.41 |
| | | $t(17) = -6.06, p < 0.001$ | | | $t(17) = -4.94, p < 0.001$ | |

The two columns next to the *Self* under each parameter are the means and SDs of the absolute differences between the *Self* stimuli and their paired physical match (*Other$_{PM}$*), and between the *Self* stimuli and their paired kinematic match (*Other$_{KM}$*). For the physical parameter, the paired t test shows a smaller mean absolute difference between *Self* and *Other$_{PM}$* than between *Self* and *Other$_{kM}$*. Likewise, for the kinematic parameter, the paired t test shows a smaller mean absolute difference between *Self* and *Other$_{kM}$* than between *Self* and *Other$_{PM}$*

vector, for every marker. The averaging procedure yielded prototypical motion for each movement type that eliminated individual characteristics and kinematic variability (Wöllner, 2013; Wöllner et al., 2012).

In sum, for each participant, 2 (movement type: with arms, without arms) × 4 (agent: Self, Other$_{PM}$, Other$_{KM}$, Morph) different point-light stimuli were created, whereby the two *Morph* sequences were the same for all the participants. The 3D point-light motion was displayed on a 2D monitor, using routines from Psychophysics Toolbox version 3 (Brainard, 1997) running on Matlab® R2012b (Mathworks). Every PLF was represented by 13 white discs against a black background, each of which subtended 0.4 degrees of visual angle (°). The whole PLF subtended approximately 5° (width) and 12° (height) when viewed at 80 cm. The PLF was displayed facing the observers, in a configuration as if the observers were watching from 20° to the left of the PLF, which served to optimize depth perception of biological motion in a 2D environment.

## Procedure

The stimuli and experimental program were controlled by a customized Matlab script and Psychtoolbox version 3 routines running on a Linux Ubuntu 14.04 LTS system. The visual stimuli were displayed on a 17-inch CRT monitor (Fujitsu X178 P117A) with a frame frequency of 100 Hz at a spatial resolution of 1024 × 768 pixels. Participants sat with a viewing distance of 80 cm. The finger taps were registered by a customized force transducer that was connected to the Linux computer via a data acquisition device (Measurement Computing®, USB-1608FS). Data were collected with a sampling frequency of 200 Hz, which was controlled and synchronized on a trial basis by the experimental program in Matlab. Participants wore closed studio headphones (AKG K271 MKII) to avoid potential auditory distraction.

The SMS task was similar to that described in Su (2016b, Experiment 2). Participants self-initiated each trial by pressing the space key, after which a PLF was shown dancing the *Charleston* steps cyclically either in the original version or in the version without the arm movement. Each participant observed either their own dance movement (*Self*), the movement of one of the two other participants (*Other$_{PM}$* or *Other$_{KM}$*), or that of the *Morph*. Participants were instructed to observe the PLF movement as a whole and to tap along to the rhythm of the leg movement in a synchronized manner (i.e., for each eight-beat cycle as shown in Fig. 1, they tapped to beat 1, 3, 5, and 7). They tapped with the index finger of their dominant hand on the force transducer. In total, eight complete movement cycles were presented on each trial, equaling 32 leg movements. Participants were not informed of the manipulation of different agents in the experiment.

The experiment followed a 2 (movement type) × 4 (agent) within-participant design, each with ten repetitions. The trials were presented in 5 blocks of 24 trials each, with all the conditions balanced across blocks and the order of conditions randomized within a block. Participants practiced five trials before starting the experiment. The experiment was completed in around 1 h.

## Data analysis

The tapping data were analyzed in the same manner as in Su (2016b), whereby the timing of each tap was extracted by identifying the time point right before the amplitude of the measured force data exceeded a predefined threshold. The tap times were temporally aligned to the onset of the visual stimulus, allowing for the calculation of asynchronies between each tap and the corresponding visual signal. The first two taps in a trial were discarded from analyses. Regarding the timing of the stimulus beat, the previous study (Su, 2016b, Experiment 2) found that, for the *Charleston* dance, peak velocity (as compared to the end position) of each foot trajectory appeared to be the more beat-defining kinematic feature of the leg movements, as it afforded more
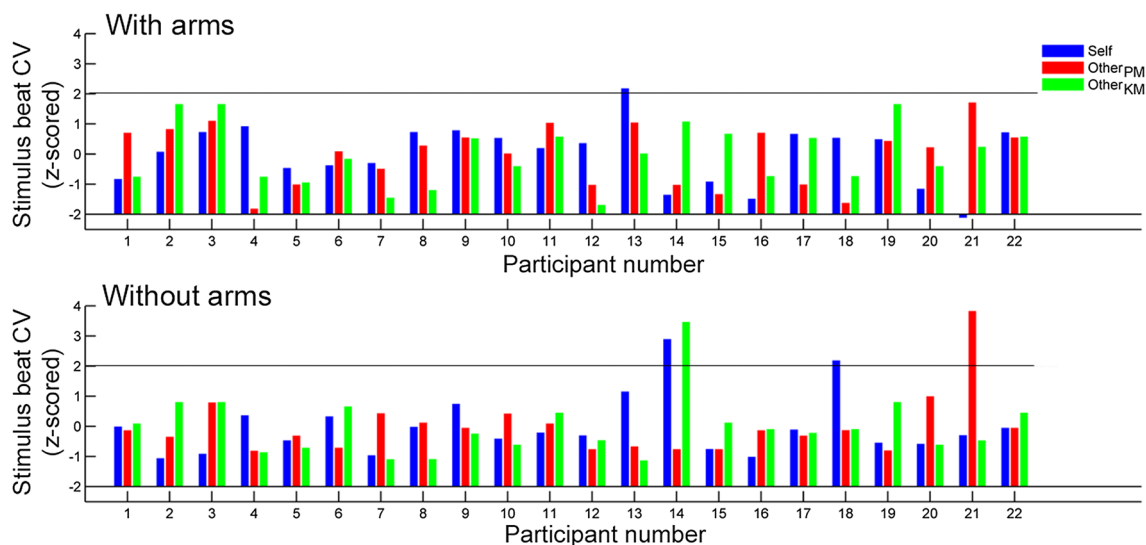
**Fig. 2** The coefficient of variations (CV) of the stimulus beat timing for the three agent conditions for each participant. Bars exceeding the horizontal black lines represent the stimulus sequences whose CV exceeds 2 SD of all the sequences in that condition. This occurred in the stimuli of four participants (nos. 13, 14, 18, and 21), whose tapping data were then excluded from analyses

stable synchronization. Thus, the 3D peak velocity of the foot markers was taken as the synchronization target in the present task (see Su, 2016b for details of its calculation), which was computed for every stimulus sequence.

The asynchronies and the synchronization stability of the taps relative to the beats (Repp & Su, 2013) were analyzed using circular statistics (Berens, 2009; see also Kirschner & Tomasello, 2009 for detailed analysis descriptions). The phase of each tap time relative to its closest beat was calculated on a circular scale (0°–360°), representing the tap-beat asynchrony. Taps with a positive asynchrony would fall into a phase between 0° and 180°, and taps with a negative asynchrony would fall into a phase between 180° and 360° (equivalent to −180° to 0°). For a single trial, the mean direction of the relative phase, $\theta$, was calculated to index the mean magnitude and direction of the tap-beat asynchronies. The main dependent variable of interest—synchronization stability—was indexed by $R$, which was the mean resultant length of the relative phase vector, ranging from 0 (no synchronization) to 1 (perfect synchronization).

For the training factor, participants with music training duration of only up to 1 year were categorized as untrained, while the others were categorized as trained (labeled here as musicians, though none was professionally trained)[2]. This yielded 11 musicians and 7 non-musicians. Note that the total sample size was 18 instead of 22, as data from four

participants were excluded from analysis due to variable stimulus beat (see the section "Beat variability" below). For additional information, we also included non-professional dance training as a between-participant factor in exploratory analyses, whose results can be found in the Supplementary Material.

## Stimulus motion analysis

Several additional analyses of the motion stimuli were conducted to characterize the timing and kinematic differences amongst stimulus conditions, which could potentially affect the synchronization task:

### Beat variability

As we intended to compare synchronization with different movement stimuli, the inevitable differences in the temporal variability of the movement beat needed to be taken into account. For this purpose, the coefficient of variation (CV) of the 32 movement beats was calculated for each of the stimulus sequences, i.e., for each *Self* and its respective *Other_PM* and *Other_KM* (see Fig. 2 for z-transformed CV values). It appeared that the stimulus combinations for four participants (nos. 13, 14, 18, and 21) contained extreme values, as one or more of the stimulus sequences had temporal variability that exceeded 2 SD of all the CV values in that agent condition. These deviations might have led to particularly variable synchronization that would bias the results. As such, tapping data from these four participants were

---

[2] Note that the musician group here were amateurs and were defined according to looser criteria than in studies investigating highly skilled experts (e.g., Karpati et al., 2016).

excluded from subsequent analyses, and the final sample size for the reported results was 18.

### Self–Other subjective movement similarity

Given the various individual kinematics, it was of interest whether movements of $Other_{PM}$ or $Other_{KM}$ were overall perceived to be more similar to the respective *Self*, which might affect synchronization behavior. A separate rating task was thus carried out on 20 observers (some of whom had taken part in the tapping experiment), in which every participant rated the movement similarity of all the Self–Other stimulus combinations presented in the tapping experiment. Specifically, participants watched each of the 18 "*Self*" stimuli paired with its respective $Other_{PM}$ or $Other_{KM}$, for the two movement types separately (yielding in total 72 different combinations for each participant[3]), and rated on each trial the similarity between the two movement sequences on a Likert scale from 1 (not at all similar) to 5 (extremely similar). Participants were instructed to base the rating on their subjective impression of the whole-body movement pattern. A mean rating score across the 20 participants was then computed for each of the 72 stimulus combinations. The mean similarity scores for the 18 $Self–Other_{PM}$ and $Self–Other_{KM}$ combinations were compared in a paired *t* test, which revealed no significant differences in the rating: $t(17) = 1.81$, $p = 0.09$ for stimuli with arm movement [$M = 3.0$ (SD 0.60) for $Self–Other_{PM}$ and $M = 3.3$ (SD 0.48) for $Self–Other_{kM}$], and $t(17) = 1.64$, $p = 0.12$ for stimuli without arm movement [$M = 3.1$ (SD 0.59) for $Self–Other_{PM}$ and $M = 3.5$ (SD 0.69) for $Self–Other_{kM}$]. Thus, neither movements of $Other_{PM}$ nor $Other_{KM}$ were overall judged to be more similar to the movements of *Self*.

### Self–Other motion spectral coherence

To objectively quantify the similarity of the motion signals between different moving agents, cross-spectral analysis (SYNCHRO Toolbox for MATLAB, codes developed by Michael J. Richardson and R. C. Schmidt, http://xkiwilabs.com/software-toolboxes/) was conducted on the time series of the motion data between each *Self* and its $Other_{PM}$ or $Other_{KM}$, for each movement type separately. The analysis yielded an average cross-spectral coherence for each pair

of time series (*Self* paired with $Other_{PM}$ or with $Other_{KM}$, where *Self* was always the referent) over a range of frequencies, which indexed the correlation between the two signals in the frequency domain (see Sofianidis, Hatzitaki, Grouios, Johannsen, & Wing, 2012). The resultant value ranged between 0 (absence of temporal relationship) and 1 (perfect synchrony of signal). Note that the frequency spectrum of the positional data was constrained by the same underlying periodicity, i.e., the metronome tempo to which each agent danced. Thus, the 3D velocity time series were entered for the spectral analyses, which should better reflect the temporal correspondence of the kinematics between different moving agents.

The analysis was carried out on two main components of the movement (Su, 2016b): the foot motion (left and right foot markers) and the trunk motion (averaged across four markers: shoulder and hip on both sides). For the foot motion, the mean spectral coherence values were submitted to a 2 (movement type) × 2 (paired agent) × 2 (limb side: left, right) repeated-measures ANOVA ($N = 18$), which revealed no significant effects except that the right foot motion of *Others* was generally more coherent with *Self* than the left foot motion, $F(1, 17) = 8.66$, $p = 0.009$, $\eta_p^2 = 0.34$. The spectral coherence of $Self–Other_{PM}$ and $Self–Other_{KM}$ did not differ, $F(1, 17) = 0.01$, $p = 0.92$, $\eta_p^2 < 0.001$ (Fig. 3a). For the trunk motion, the 2 (movement type) × 2 (paired agent) repeated-measures ANOVA on the mean spectral coherence revealed an effect of paired agent, $F(1, 17) = 8.20$, $p = 0.01$, $\eta_p^2 = 0.33$, showing that the trunk motion of $Other_{KM}$, compared to $Other_{PM}$, was more coherent with that of *Self* (Fig. 3b).

Across the results of rating and spectral analyses, the movements of $Other_{PM}$ and $Other_{KM}$ differed from each other most consistently in how coherent their trunk motion (velocity time series) was with respect to *Self*, with $Other_{KM}$ being the more coherent amongst the two. This difference was, notably, not reflected in the subjective judgment of movement similarity.

## Results

We first report the ANOVA result for synchronization stability ($R$)—the main measure of performance—in the tapping task, followed by the ANOVA result for stimulus beat variability (CV) as a contrast, to ascertain which of the main effects or interactions in synchronization stability are not attributable to stimulus variability. Next, we report the correlations between stimulus beat variability and SMS variability (circular variance: $1 − R$) to reveal possible links between these two measures across different stimulus conditions. We then report correlations between SMS variability and attributes of Self–Other similarity in the motion stimuli. Finally,
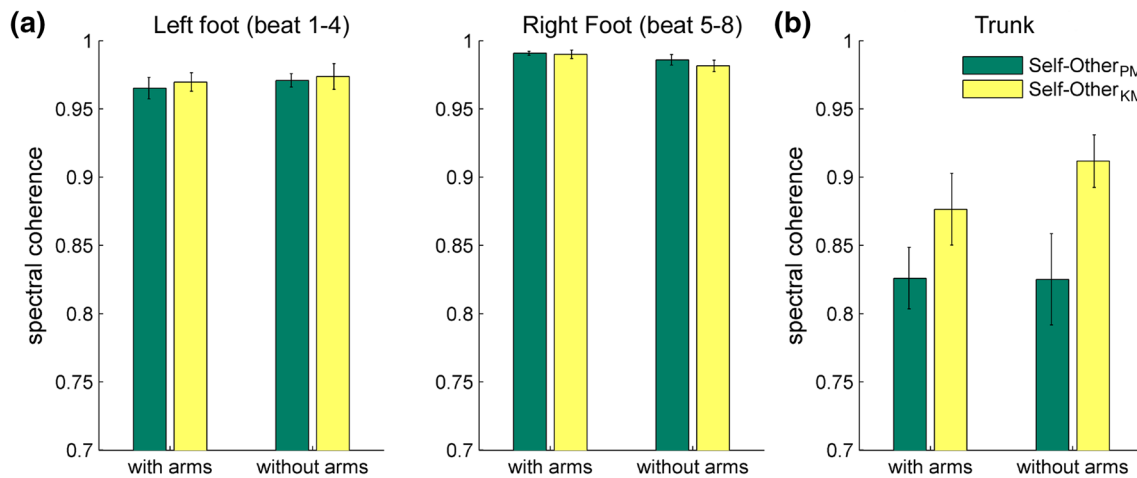
---

[3] For those who had taken part in the tapping experiment, the stimulus pairs involving their own movement accounted for 4 out of 72 combinations they rated. As the mean rating for each combination was calculated from the response across 20 observers, it was unlikely that the mean score for a given stimulus pair could have been swayed by one rater who had seen his/her own movement in that pair. In addition, participants for the rating task were not informed about the agency of the movement stimuli.

**Fig. 3** Mean spectral coherence of the 3D velocity time series between the *Self* and the *Other*$_{PM}$ or *Other*$_{KM}$ stimuli, for **a** the left foot- and right foot motion, and **b** the trunk motion. Error bars are standard error of the means
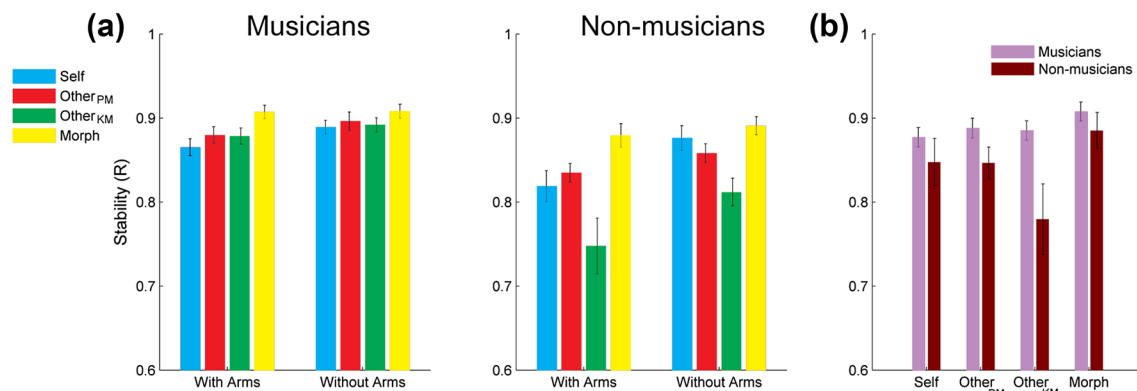


**Fig. 4** Mean synchronization stability as indexed by *R*. **a** Mean *R* for each experimental condition. **b** Mean *R* as a function of movement agent, for musician and non-musician groups separately. Error bars are standard error of the means

we show the result of an additional self-recognition task and its correlation with SMS variability to explore the link between motor simulation and temporal prediction (Keller et al., 2007).

## Synchronization stability

The individual means of *R* were submitted to a full factorial 2 (movement type) × 4 (movement agent) × 2 (music training) mixed ANOVA, with movement type and movement agent as within-participant factors and music training as between-participant factor. For the ANOVAs, Greenhouse–Geisser correction was applied to the *p* values of effects involving movement agent.

There was a main effect of movement type, $F(1, 16) = 20.91$, $p < 0.001$, $\eta_p^2 = 0.56$, showing better synchronization with dance stimuli where the arms remained still. The main effect of movement agent was also significant, $F(3,$

$48) = 9.98$, $p < 0.001$, $\eta_p^2 = 0.38$, for which the post hoc comparisons (paired *t* test with Bonferroni-corrected *p* values) revealed only better synchronization with *Morph* than with all the other agent conditions: *Morph* vs. *Self*, $t(17) = 4.45$, $p = 0.002$, *Morph* vs. *Other*$_{PM}$, $t(17) = 3.55$, $p = 0.01$, and *Morph* vs. *Other*$_{KM}$, $t(17) = 3.23$, $p = 0.03$. There was marginally better synchronization of musicians than non-musicians, $F(1, 16) = 4.69$, $p = 0.05$, $\eta_p^2 = 0.23$ (Fig. 4a).

There was a significant interaction between musical training and movement agent, $F(3, 48) = 5.24$, $p = 0.01$, $\eta_p^2 = 0.25$. To unpack this interaction, the musician and non-musician groups were compared for each of the movement agent conditions separately. Two-sample t tests showed that the two groups did not differ when synchronizing with their own movements, $t(16) = 1.11$, $p = 0.28$, nor when synchronizing with *Morph*, $t(16) = 1.02$, $p = 0.32$. By contrast, musicians were better than non-musicians at synchronizing with *Other*$_{KM}$, $t(16) = 2.93$, $p = 0.009$, and marginally so
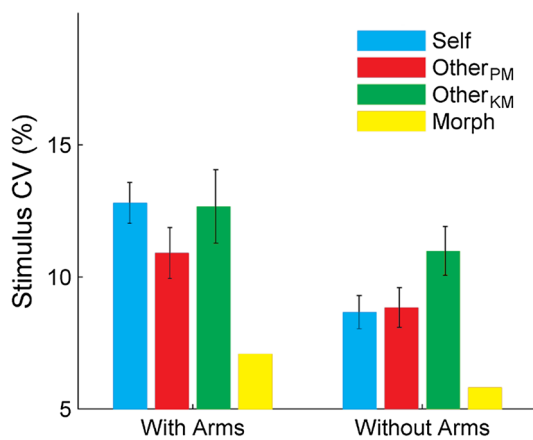
**Fig. 5** Mean CV of the stimulus beat timing for each experimental condition. Error bars are standard error of the means

when synchronizing with $Other_{PM}$, $t(16) = 1.96$, $p = 0.068$ (Fig. 4b).

The interaction between musical training and movement type was also significant, $F(1, 16) = 4.84$, $p = 0.04$, $\eta_p^2 = 0.23$. Two-sampled $t$ tests showed that musicians were better than non-musicians for dance stimuli with arm movement, $t(16) = 2.6$, $p = 0.02$, whereas the two groups did not differ significantly for dance stimuli without arm movement, $t(16) = 1.94$, $p = 0.07$. However, this interaction seemed mainly to reflect a corresponding interaction found in the stimulus beat variability (see the next section), and so its implications will not be considered further.

No other significant main effects or interactions were found.

## Contrasting stimulus beat variability with synchronization stability

To examine whether the effects observed in synchronization stability ($R$) only reflected the differences in beat variability amongst different conditions, a 2 (movement type) × 3 (movement agent, excluding *Morph*) × 2 (music training) mixed ANOVA was conducted on the stimulus beat variability (CV). There was a main effect of music training, $F(1, 16) = 8.26$, $p = 0.01$, $\eta_p^2 = 0.34$, showing that musicians' stimuli ($M = 9.80$, SD 1.57) were less variable than those of non-musicians ($M = 12.37$, SD 2.24). Other than that, there were only the main effect of movement type, $F(1, 16) = 35.46$, $p < 0.001$, $\eta_p^2 = 0.69$ (stimuli with arm movement were overall more variable than those without), and the interaction between movement type and music training, $F(1, 16) = 5.92$, $p = 0.03$, $\eta_p^2 = 0.27$. Post hoc two-sample $t$ tests showed that musicians' stimuli were less variable than those of the non-musicians for stimuli with arm movement, $t(16) = 2.92$, $p = 0.009$, whereas this difference was less obvious for stimuli without arm movement, $t(16) = 2.13$, $p = 0.049$. No other significant effect or interaction was identified (Fig. 5).

To contrast the results of stimulus beat variability (CV) and synchronization stability ($R$), the full factorial mixed-ANOVA on $R$ was repeated without the *Morph* condition, i.e., 2 (movement type) × 3 (movement agent) × 2 (music training). The same pattern of results were obtained as in the previous 2 × 4 × 2 mixed ANOVA on $R$. Specifically, the interaction between movement agent and music training remained, $F(2, 32) = 5.13$, $p = 0.017$, $\eta_p^2 = 0.24$. The corresponding interaction was not found in the results of stimulus CV. Thus, only the effect of movement type and its interaction with music training appeared to be driven by the stimulus variability. Table 2 summarizes the ANOVA results for synchronization stability and stimulus beat variability.

Finally, $t$ tests were applied to compare the beat variability of each stimulus condition to that of the *Morph*, which confirmed that all the 2 (movement type) × 3 (movement agent) conditions had more variable stimulus beat than the respective *Morph*, all $p$ values < 0.001.

**Table 2** Contrasting the main effects and interactions identified in synchronization stability ($R$) and in stimulus beat variability (CV)

| | Synchronization stability ($R$) | Stimulus beat variability (CV) |
|---|---|---|
| **Main effect** | | |
| Movement type | √ | √ |
| Agent | √ | – |
| Music training | – (marginal) | √ |
| **Interaction** | | |
| Music training × agent | √ (both with or without *Morph*) | – |
| Music training × movement type | √ | √ |

Significant effects and interactions are marked by "√" and non-significant ones by "–"

## Correlation between beat variability and SMS variability

To further examine how individual synchronization was related to the beat variability, Spearman's correlational analysis was conducted across participants ($N = 18$) between beat variability (CV) and SMS variability (circular variance, indexed by $1 - R$), for each of the 2 (movement type) $\times$ 3 (movement agent) conditions. Data for both variability measures were $z$-transformed prior to the correlational analysis. A positive correlation would indicate that, for a given condition on an individual participant level, tap-beat asynchronies tended to be more variable (i.e., less stable synchronization) when the stimulus beat was more variable. The analyses yielded significant correlations (corrected for multiple comparisons) in the following conditions: $Other_{PM}$ with arms, $r_s = 0.59$, $p = 0.009$, $Other_{KM}$ with arms, $r_s = 0.81$, $p < 0.001$, and $Other_{KM}$ without arms, $r_s = 0.64$, $p = 0.004$. (Fig. 6). To test whether these correlations were significantly different from those found in the *Self* conditions, Fisher's $r$ to $z$ transformation was applied to compare the strength of the correlations, which revealed that the correlation in $Other_{KM}$ was significantly stronger than in *Self,* $z = 2.27$, $p$ (two-tailed) $= 0.02$ with arms and $z = 2.21$, $p = 0.03$ without arms. By contrast, for both movement types, the correlation in $Other_{PM}$ was not significantly different from that in *Self,* $z = 1.08$, $p = 0.28$ with arms and $z = 0.81$, $p = 0.42$ without arms.

In addition, it was of interest whether the *difference* in beat variability between agents was linked to the *difference* in SMS variability between agents. For this purpose, correlational analyses were conducted between the following two measures: (1) the difference in beat variability between a non-Self ($Other_{PM}$ or $Other_{KM}$) and the *Self* condition, and (2) the difference in SMS variability between the respective conditions. A positive correlation would suggest that an increase/decrease in beat variability in a given Other condition (compared to *Self*) was associated with an increase/decrease in SMS variability in that condition (compared to *Self*). After correcting for multiple comparisons, a significant positive correlation was found only for the $Other_{KM}$–*Self* combination with arms, $r_s = 0.76$, $p < 0.001$ (Fig. 7).

## Relation between stimulus similarity and SMS variability

### Similarity as indexed by rating

An additional research question addressed whether participants synchronized better with a non-Self movement when it was judged to be more similar to his or her own movement. Spearman's correlational analyses were conducted ($N = 18$) between the mean similarity rating scores of the Self–Other

stimuli and the SMS variability, for each of the 2 (non-self agent) $\times$ 2 (movement type) conditions. No significant correlation was found; nor was there any correlation when only the similarity scores were taken of each participant judging his or her own Self–Other stimulus pairs.

### Similarity as indexed by motion spectral coherence

The average spectral coherence between each Self and Other—for the foot motion, specifically—was used as an objective measure of similarity in their motion signals. Spearman's correlation analysis was then conducted between these two measures: (1) spectral coherence between each *Self* and the corresponding $Other_{PM}$ or $Other_{KM}$, and (2) the difference in SMS variability between the respective conditions. A negative correlation would suggest that lower SMS variability (or better synchronization) in a non-Self condition, as compared to in Self, was associated with higher motion coherence between the two agents. A significant negative correlation was found only in the left foot motion (the first two movement beats in every cycle) of the *Self–Other$_{KM}$* pair, for the movement type with arms: $r_s = -0.58$, $p = 0.01$ (Fig. 8).

## Relation to self recognition

As synchronization with movements of different agents might be associated with sensitivity to one's own action (Keller et al., 2007), an additional self-recognition task (Keller et al., 2007; Sevdalis & Keller, 2010; Wöllner, 2012) was conducted in a later session. 16 of the 18 participants returned to take part. Each of these participants observed their own movements or movements of their matched $Other_{PM}$ and $Other_{KM}$, and self recognition for each individual was assessed by the Signal Detection Theory measure of sensitivity, $d'$ (Stanislaw & Todorov, 1999). $d'$ was the $z$-transformed hit rate minus the $z$-transformed false alarm rate, and a greater $d'$ value indicated greater sensitivity to the target (i.e., sensitivity to one's own movement). Here, a hit was a correct answer of "Self" when seeing *Self*, whereas a false alarm was an incorrect answer of "Self" when seeing $Other_{PM}$ or $Other_{KM}$. While self-recognition ($d'$) was overall better for movements with arms than without, $t(15) = 2.67$, $p = 0.017$, further analyses revealed no significant correlation between self-recognition and the variability of synchronizing with *Self,* $r_s = -0.45$, $p = 0.08$ with arms and $r_s = 0.08$, $p = 0.76$ without arms. However, a curious result was found in response bias ($C$): The 2 (other agent) $\times$ 2 (movement type) repeated-measures ANOVA showed that, for both movement types, observers tended more often to misidentify the movements of $Other_{KM}$ ($M = -0.18$) than that of $Other_{PM}$ ($M = 0.26$) as *Self,* $F(1, 15) = 8.80$, $p = 0.01$, $\eta_p^2 = 0.37$.

**Fig. 6** Correlations between the stimulus beat variability (CV) and the SMS variability $(1 - R)$ for each agent condition. Z-scored values were entered for the analyses. Black lines represent the linear regression across data points where correlations were significant
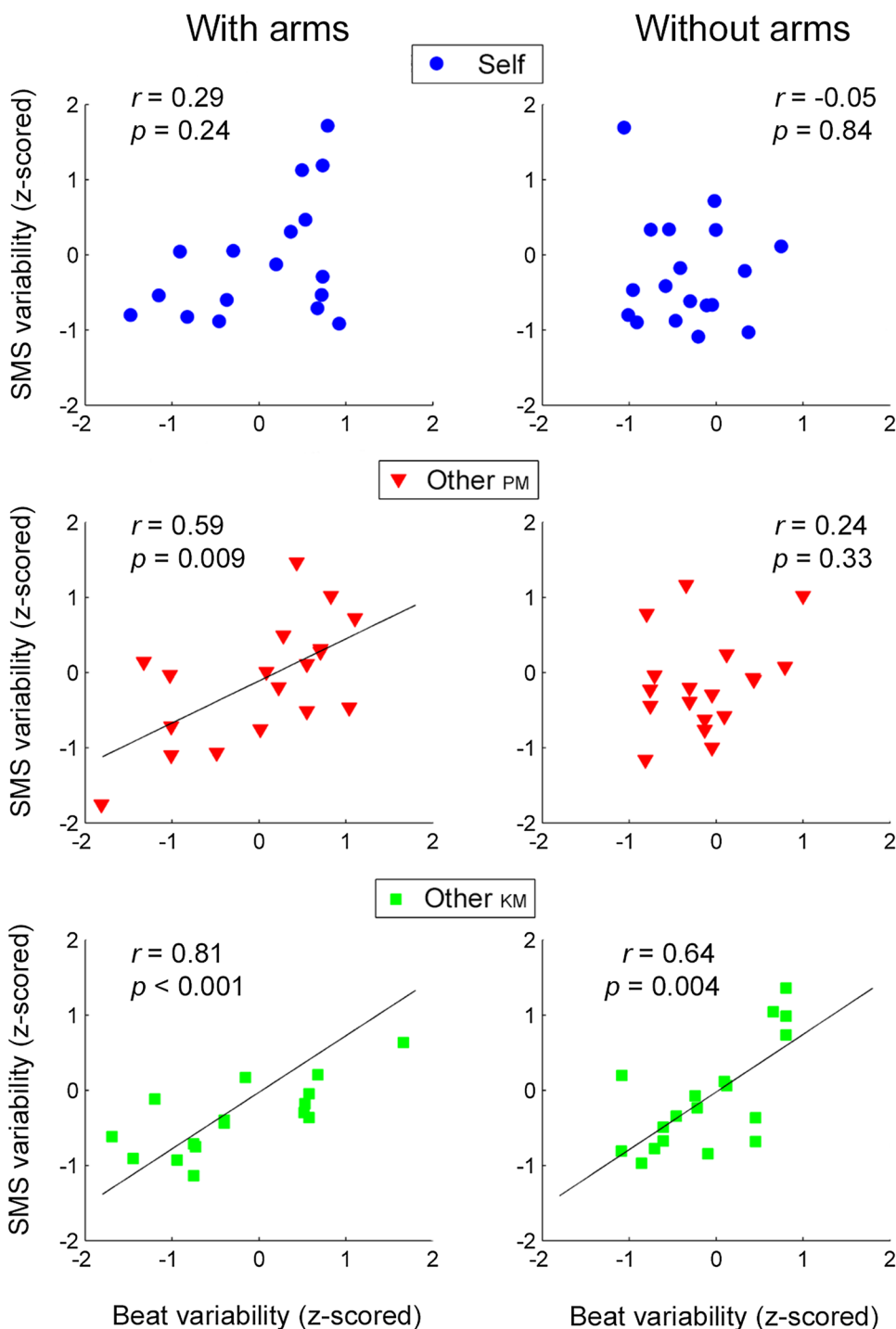


## Result summary

To summarize, on a group level, synchronization with the beat in point-light displays of dancing was most stable with the *Morph* movements. Music training modulated synchronization behavior, such that musicians were better than non-musicians at synchronizing with movements produced by others. This difference was more obvious with *Other$_{KM}$*

(matched in velocity) than with *Other$_{PM}$* (matched in morphology). By contrast, musicians and non-musicians were comparable when synchronizing with their own movements or with those of the *Morph*. Across individuals, higher beat variability in the movement stimuli was correlated with more variable SMS only when synchronizing with the movements of others, and more consistently so for the movements of *Other$_{KM}$*. Furthermore, relative to synchronizing with

**Fig. 7** Correlations between the difference in stimulus beat variability and the difference in SMS variability, the difference being the deviation of $Other_{PM}$ or $Other_{KM}$ from $Self$. The black line represents the linear regression across data points where the correlation was significant
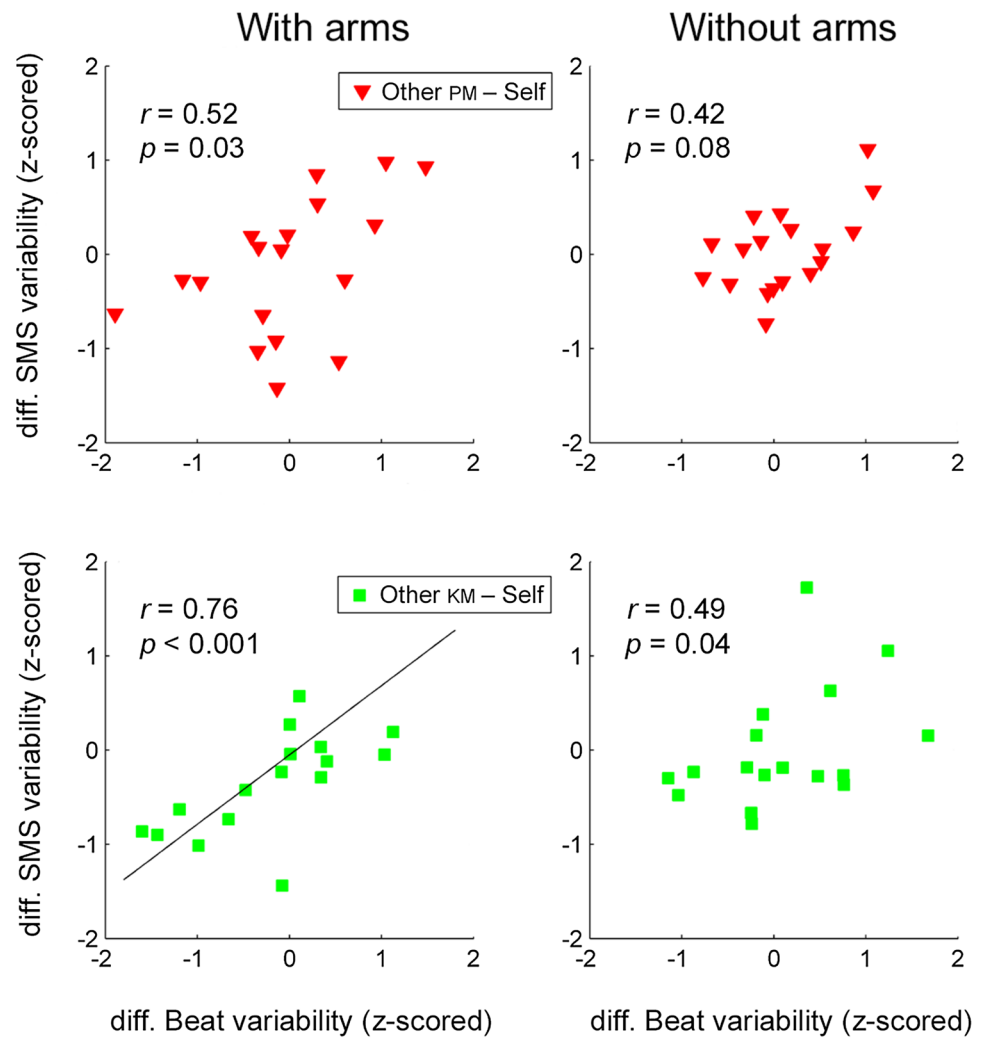


self-generated movements, the increase or decrease of beat variability in $Other_{KM}$ (but not in $Other_{PM}$) was linked to a corresponding increase or decrease of SMS variability, especially for stimuli with arm movement. The variability of synchronizing with $Other_{KM}$ also seemed to be associated with how similar its velocity profile was to that of the $Self$, in terms of the spectral coherence of the left foot motion. Sensitivity to one's own movement was not associated with SMS with the $Self$ stimuli in the present task.

Although most of the effects of movement agent were found in $Other_{KM}$, these movements were not explicitly judged to be more similar to $Self$ than those of $Other_{PM}$, nor was their foot motion more spectrally coherent with $Self$. Instead, besides the maximal velocity of leg movements, the main quantifiable difference between $Other_{PM}$ and $Other_{KM}$ appeared to be in trunk motion, with $Other_{KM}$ being more coherent with that of $Self$. This may be linked to the result that $Other_{KM}$ was more often confused as $Self$ in the recognition task, which has implications in the motor simulation mechanisms adopted in the present SMS (see "Discussion").

## Discussion

In the framework of action simulation for SMS, we investigated how synchronization (finger tapping) with visual rhythms of specific dance movements was modulated by movement agency, i.e., whether the movement had been self-generated or not. According to previous findings (Keller et al., 2007; Repp & Keller, 2010), observing one's own action stimuli should yield more optimal simulation than observing those of others, leading to better temporal prediction and SMS. On top of that, we examined how the effect of agency was influenced by music training, and what cues (morphology or kinematics) in the movement stimuli were more relevant for simulation in a SMS task.

### The absence of a general "self advantage"

First, contrary to previous findings in music performance (Keller et al., 2007; Repp & Keller, 2010), we observed no absolute "self advantage" in our task. Namely,

**Fig. 8** Correlations between the Self–Other spectral coherence of the left foot motion and the difference in SMS variability (between *Self* and the respective *Other*). The black line represents the linear regression across data points where the correlation was significant
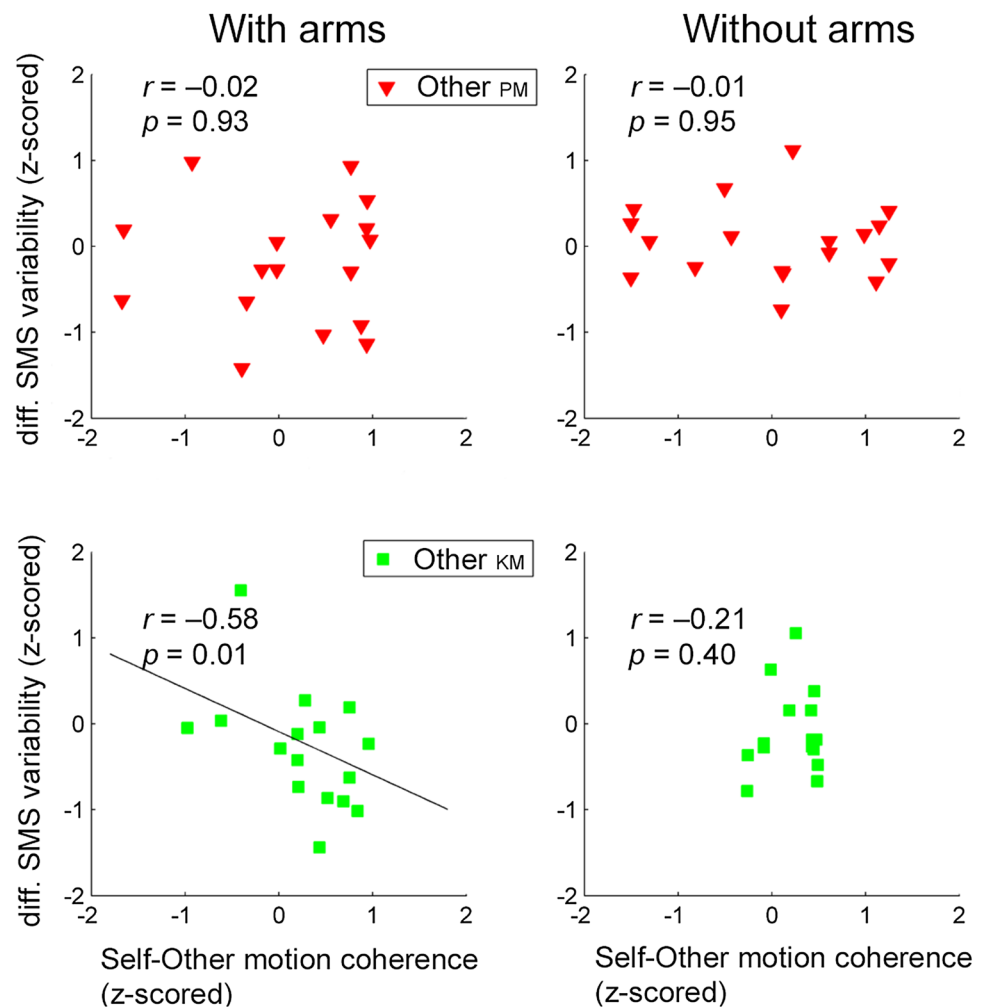


synchronization with one's own point-light dance stimuli was not overall better than with others'. Regarding the effect of movement agent, the most consistent result seems to be that the morphed motion in both styles (with or without arm movement) afforded the best SMS performance, as has been shown before in stimuli of musical conductors' gestures (Wöllner et al., 2012). While at first sight this result seems at odds with the simulation hypothesis, a few reasons need to be considered to explain the lack of a direct self advantage: One critical difference in the present study, as compared to Keller et al. (2007), is that we explicitly minimized the presence of intentional timing deviations in the stimuli through recruitment of non-experts and the performance of paced movements. As opposed to highly skilled experts whose (unpaced) actions are a result of precisely controlled motor sequencing, e.g., the expressive timing in musicians (Repp & Knoblich, 2004) or the well-practiced choreograph in dancers (Calvo-Merino et al., 2005), non-experts who have just learned the dance steps—performed while being paced—likely produced more *unintended* spatiotemporal variability in their movement (Koul et al., 2016). That is,

the individual timing cues were largely unintentional. Our data thus suggest that unintentional individual variations in beat timing could not be effectively used for synchronization purpose. While it has been shown that untrained participants can reliably identify their own visual actions from those of others (Knoblich & Prinz, 2001; Loula et al., 2005; Sevdalis & Keller, 2010), for which internal simulation is arguably required, this mechanism may not account for the unintended variability in the motion signals. Namely, as participants did not intend to produce the spatiotemporal variability when they performed the movement, they could not use this information effectively to predict the stimulus variability during simulation. Such variability in the action stimuli could eventually weaken or hinder prediction for SMS. This might also explain the present lack of correlation between self recognition and synchronization stability for self-generated movement, as opposed to the previously reported correlation in expert musicians (Keller et al., 2007). Judging from the result that the morphed motion with the lowest spatiotemporal variability—despite the movement being least naturalistic—yielded the best SMS, it seems that the predictability of

the visual rhythm derived from consistent motion signals is still more critical for stable synchronization, at least for non-experts. This interpretation, though, does not preclude the possibility that observers nevertheless attempted to simulate the movement rhythm. It may rather suggest that simulation in this case was not optimal due to unintended noise in the signals, as a few other results discussed below still point to simulation as possibly underlying the present SMS.

Another possible explanation for the inefficient simulation when synchronizing with self- compared to other-generated movement is that different neural mechanisms underlying motor simulation may be involved in dance observation and in beat perception. The former is known to engage the action observation network associated with ventral premotor areas (Calvo-Merino et al., 2005; Cross et al., 2009), whereas beat perception in musical contexts may instead rely on the dorsal premotor regions implicated in motor planning, according to the ASAP hypothesis (Patel & Iversen, 2014; Ross et al., 2016). Recent evidence suggests that the auditory dorsal pathway is also involved in beat perception of static visual stimuli (Araneda, Renier, Ebner-Karestinos, Dricot, & De Volder, 2017). Although synchronization with visual action (dance) stimuli appears to be a scenario where both neural networks could be involved, the degree to which rhythm-specific neural substrates are activated during action observation—and how it compares to those activated during auditory rhythm processing—remains to be investigated.

## Music training modulates simulation of self vs. other

Another main finding was that music training modulated the effect of movement agency. Specifically, musicians and non-musicians differed mainly in synchronizing with other-generated movements, with musicians performing better in this regard. It seems that music training facilitates internal simulation of movement patterns (and their timing) that deviate from one's own, thus enabling more accurate prediction of their visual rhythm. While music training is often shown to improve auditory rhythm perception and synchronization (Manning & Schutz, 2016; Repp, 2010; Su & Pöppel, 2012), the present result establishes the training effect in situations of more complex and unpredictable visual rhythms, such as the movements of other agents. On the other hand, despite the lack of a simple "self advantage", the benefit of self-generated movement was indicated when comparing between musicians and non-musicians: Non-musicians synchronized just as well as musicians with their own movements and with the morphed movements, suggesting that these conditions may have been easier and less dependent on additional sensorimotor skills acquired through training (Karpati et al., 2016; Ono, Nakamura, & Maess, 2015; Repp, 2010). Put differently, synchronizing with self-generated movement appeared to be independent of music training as with the

most stable, morphed stimuli, which could be interpreted to indicate that both conditions yielded predictable stimulus timing and thus (relatively) effective simulation—even for untrained participants. As a whole, the interaction between music training and movement agent points to the training-independent benefit of simulating self-generated visual rhythms (i.e., the easier condition), as well as the training-related facilitation of simulating other-generated ones (i.e., the more challenging condition). This pattern is reminiscent of a previous finding (Su & Pöppel, 2012) that musicians and non-musicians performed equally well in an auditory pulse-finding task when they could move along with the rhythm (i.e., rhythm simulation made easier by overt movement, see also Manning & Schutz, 2013), while non-musicians did far worse than musicians when they had to sit still without moving (i.e., a more difficult condition relying only on covert simulation).

When discussing the present training effect of music, a possible limitation lies in the fact that participants categorized as "trained" had heterogeneous degrees of training experience, and none of them had trained professionally (the rationale as outlined in the Introduction). Besides, we did not control whether they were still actively practicing at the time of the investigation. While examining amateurs should make it possible to generalize the effect to a larger population, it stands in contrast to studies investigating the expertise effect specifically in highly skilled, currently active musicians with more than 10 years of training (Karpati et al., 2016). As such, the present training effect was likely associated with the learning experiences of music, rather than its specialization. The latter remains an outstanding question for future investigations, namely, whether the effects would be more pronounced in professional experts.

## The effect of movement agent: kinematic cues matter

With regard to the movement agent, the correlation results show that simulating the dance rhythm seemed to be driven to a greater degree by kinematic similarity (with respect to one's own) than morphological similarity in the stimuli. When comparing between the conditions of $Other_{PM}$ and $Other_{KM}$, which were matched to each $Self$ in terms of physical proportion and movement velocity, respectively, SMS variability correlated more consistently with beat variability when tapping to $Other_{KM}$. The tap-beat variability correlation indicates that the observers attempted to simulate or predict the beat timing in the movement, and that this simulation did not account well for the non-self, unpredictable timing variations (Ragert, Schroeder, & Keller, 2013) and the non-self spatiotemporal noise in the stimuli. Thus, a greater tendency to simulate would lead to a higher correlation between these measures, which was observed most

strongly in SMS with $Other_{KM}$ for both unconstrained (with arms) and constrained (without arms) movements, and to a lesser extent with $Other_{PM}$ for the unconstrained movement (Fig. 6). Moreover, when quantifying the variability of SMS and stimulus beats as their *difference* to the baselines measured in the *Self* conditions, the correlation was most pronounced for the unconstrained movement in $Other_{KM}$ (Fig. 7). This correlation captures how the deviation of movement beat variability from one's own was linked to a corresponding increase or decrease in SMS variability, suggesting that each observer simulated the movement timing of the other agents using a motor template from their own action system (Ragert et al., 2013). Complementary to this result is the observation that, for the same condition of $Other_{KM}$, the difference of SMS variability from the *Self* baseline was negatively associated with the left foot motion coherence between $Other_{KM}$ and *Self* (Fig. 8). That is, the more temporally similar the left foot motion signals between these two agents, the less variable synchronization with $Other_{KM}$ (relative to with *Self*). As to why this correlation was only found in the left foot, one speculation is that the left foot performed beat 1 and beat 2 of the four-beat cycle in the leg movement, whose metrical positions may be more perceptually salient than those of the right foot (beat 3 and beat 4), according to the dynamic attending theory (Large & Jones, 1999).

That the tap-beat variability correlations tended to be stronger for unconstrained movements additionally suggests that observers were more inclined to simulate when the stimuli contained richer and more naturalistic information—in this case whole-body dancing. Although the arm movement in our stimuli was irrelevant to the task and not likely to affect visual beat perception (Su & Salazar-López, 2016) and synchronization (Su, 2016b), it certainly added information of individual characteristics in the dance that could contribute to the action-perception coupling. In agreement with this argument was the result that, indeed, self recognition was more successful for movements with arms than without. Similar results have been reported of better self recognition for more individualistic actions (e.g., dancing as compared to walking or hand clapping, Sevdalis & Keller, 2010) and for more expressive movements (Sevdalis & Keller, 2011), supporting the idea that simulation is facilitated by the individual styles present in the motion stimuli.

The correlations in the $Other_{KM}$ conditions show that movements with similar kinematic (velocity) features to one's own, instead of movements performed by a physically familiar figure, promote simulation in SMS. This is corroborated by the result that participants also more often misidentified $Other_{KM}$ rather than $Other_{PM}$ as themselves. The finding of the velocity cues driving simulation in a SMS task is consistent with previous studies showing that the velocity information in the movement trajectory is critical

for recognizing one's own action (Knoblich & Prinz, 2001) as well as for visual beat perception (Su, 2014) and synchronization (Luck & Sloboda, 2009; Su, 2016b; Wöllner et al., 2012). As opposed to the kinematics, the physical appearance of the motion stimuli has repeatedly been shown to have no influence on performance in rhythm tasks (Hove, Spivey, & Krumhansl, 2010; Ruspantini, D'Ausilio, Mäki, & Ilmoniemi, 2011; Su, 2014). It thus seems that, as far as visual rhythm is concerned, only attributes of the motion profile of an action are relevant for simulation.

It should be noted that the kinematic similarity between $Other_{KM}$ and *Self* lay in the amplitude of peak velocity of the leg motion, and *not* in the temporal profile of this velocity cue. Namely, the beat timing of $Other_{KM}$ was not more consistent with *Self* than $Other_{PM}$, as reflected in the motion coherence analyses (see Keller et al., 2007, who also reported no correlation between timing similarity and self-advantage in simulation). Besides the peak velocity, $Other_{KM}$ was instead more temporally similar to *Self* in terms of their trunk motion, i.e., the periodically vertical bouncing pattern. While the trunk motion constituted another main component of this dance (Su, 2016a), it was essentially irrelevant for the present task. However, it has been shown that when observers synchronize to the leg motion (the beat) of the present dance stimuli, the trunk motion may be perceived in parallel as "subdivisions" between successive beats (Su, 2016b). That is, the trunk motion constitutes part of the visual rhythm communicated by the dance movement. Thus, it is possible that during dance observation, participants picked up all the salient kinematic and timing cues available in the movement, the overall similarity of which to their own would trigger the internal motor resonance, as supported by the present results of synchronization and (false) self-identification.

Overall, these findings hold implications for designing optimal visual or audiovisual rhythmic stimuli involving naturalistic motion, which could capitalize on the kinematic compatibility between the stimulus and the perceiver's intrinsic motor preference. For example, future research could present visual movement rhythms through individualized visual motion stimuli, or even avatars, that incorporate each individual's critical velocity parameters while minimizing temporal variability in the stimuli.

## Conclusions

In conclusion, we found that despite the lack of an absolute advantage for synchronizing with self-generated (paced) dance stimuli, synchronizing with one's own movement appeared less demanding as it was independent of rhythm-related expertise, whereas synchronizing with others' movement benefitted from music training. The similarity of the

kinematic (velocity) cues of an observed movement to one's own, rather than the similarity of the physical appearance of the dancing agent, drove the simulation mechanism for SMS. The degree of simulation also seemed to be associated with how much the other-generated beat timing deviated from one's own. Together these results point to the role of motor simulation in SMS with dance rhythm, which is modulated by each individual's kinematic profile and movement timing. The present findings support the interconnected frameworks of rhythm processing and action simulation, which may operate in a similar manner for both music and dance.

## Compliance with ethical standards

**Ethical approval** All procedures performed in this study involving human participants had been approved by the ethical standards of the Ethics Commission of Technical University of Munich, and were in accordance with the ethical standards of the Ethics Commission of Technical University of Munich and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Written informed consent was obtained from all individual participants included in the study.

**Conflict of interest** The authors declare that they have no conflict of interest.

**Data availability** The datasets generated during and/or analyzed during the current study are available in the Open Science Framework (OSF) repository, https://doi.org/10.17605/OSF.IO/KTDNV.

## References

Araneda, R., Renier, L., Ebner-Karestinos, D., Dricot, L., & De Volder, A. G. (2017). Hearing, feeling or seeing a beat recruits a supramodal network in the auditory dorsal stream. *European Journal of Neuroscience, 45*, 1439–1450. https://doi.org/10.1111/ejn.13349.

Berens, P. (2009). CircStat: A MATLAB toolbox for circular statistics. *Journal of Statistical Software, 31*(10), 1–21.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433–436.

Burger, B., Thompson, M. R., Luck, G., Saarikallio, S. H., & Toiviainen, P. (2014). Hunting for the beat in the body: On period and phase locking in music-induced movement. *Frontiers in Human Neuroscience, 8*, 1–16. https://doi.org/10.3389/fnhum.2014.00903.

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: An FMRI study with expert dancers. *Cerebral Cortex, 15*(8), 1243–1249. https://doi.org/10.1093/cercor/bhi007.

Cross, E. S., Hamilton, A. F. de C., & Grafton, S. T. (2006). Building a motor simulation de novo: Observation of dance by dancers. *NeuroImage, 31*(3), 1257–1267. https://doi.org/10.1016/j.neuroimage.2006.01.033.

Cross, E. S., Kraemer, D. J. M., Hamilton, A. F. de C., Kelley, W. M., & Grafton, S. T. (2009). Sensitivity of the action observation network to physical and observational learning. *Cerebral Cortex, 19*(2), 315–326. https://doi.org/10.1093/cercor/bhn083.

Decety, J., & Chaminade, T. (2003). When the self represents the other: A new cognitive neuroscience view on psychological identification. *Consciousness and Cognition, 12*(4), 577–596. https://doi.org/10.1016/S1053-8100(03)00076-X.

Eaves, D. L., Riach, M., Holmes, P. S., & Wright, D. J. (2016). Motor imagery during action observation: A brief review of evidence, theory and future research opportunities. *Frontiers in Neuroscience, 10*, 1819. https://doi.org/10.1186/1743-0003-5-8.

Fujioka, T., Ross, B., & Trainor, L. J. (2015). Beta-band oscillations represent auditory beat and its metrical hierarchy in perception and imagery. *Journal of Neuroscience, 35*(45), 15187–15198. https://doi.org/10.1523/JNEUROSCI.2397-15.2015.

Gallese, V., & Sinigaglia, C. (2011). What is so special about embodied simulation? *Trends in Cognitive Sciences, 15*(11), 512–519. https://doi.org/10.1016/j.tics.2011.09.003.

Grahn, J. A. (2012). Neural mechanisms of rhythm perception: Current findings and future perspectives. *Topics in Cognitive Science, 4*(4), 585–606. https://doi.org/10.1111/j.1756-8765.2012.01213.x.

Grahn, J. A., & Brett, M. (2009). Impairment of beat-based rhythm discrimination in Parkinson's disease. *Cortex, 45*(1), 54–61. https://doi.org/10.1016/j.cortex.2008.01.005.

Hove, M. J., Iversen, J. R., Zhang, A., & Repp, B. H. (2013). Synchronization with competing visual and auditory rhythms: Bouncing ball meets metronome. *Psychological Research, 77*(4), 388–398. https://doi.org/10.1007/s00426-012-0441-0.

Hove, M. J., Spivey, M. J., & Krumhansl, C. L. (2010). Compatibility of motion facilitates visuomotor synchronization. *Journal of Experimental Psychology: Human Perception and Performance, 36*(6), 1525–1534. https://doi.org/10.1037/a0019059.

Iversen, J. R., Patel, A. D., Nicodemus, B., & Emmorey, K. (2015). Synchronization to auditory and visual rhythms in hearing and deaf individuals. *Cognition, 134*(C), 232–244. https://doi.org/10.1016/j.cognition.2014.10.018.

Iversen, J. R., Repp, B. H., & Patel, A. D. (2009). Top-down control of rhythm perception modulates early auditory responses. *Annals of the New York Academy of Sciences, 1169*, 58–73. https://doi.org/10.1111/j.1749-6632.2009.04579.x.

Jeannerod, M. (2001). Neural simulation of action: A unifying mechanism for motor cognition. *NeuroImage, 14*(1), S103–S109. https://doi.org/10.1006/nimg.2001.0832.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics, 14*(2), 201–211. https://doi.org/10.3758/BF03212378.

Jola, C., Abedian-Amiri, A., Kuppuswamy, A., Pollick, F. E., & Grosbras, M.-H. (2012). Motor simulation without motor expertise: Enhanced corticospinal excitability in visually experienced dance spectators. *PLoS One, 7*(3), e33343. https://doi.org/10.1371/journal.pone.0033343.s003.

Karpati, F. J., Giacosa, C., Foster, N. E. V., Penhune, V. B., & Hyde, K. L. (2016). Sensorimotor integration is enhanced in dancers and musicians. *Experimental Brain Research, 234*(3), 893–903. https://doi.org/10.1007/s00221-015-4524-1.

Karpati, F. J., Giacosa, C., Foster, N. E. V., Penhune, V. B., & Hyde, K. L. (2017). Dance and music share gray matter

structural correlates. *Brain Research, 1657*, 62–73. https://doi.org/10.1016/j.brainres.2016.11.029.

Keller, P. E., Knoblich, G., & Repp, B. H. (2007). Pianists duet better when they play with themselves: On the possible role of action simulation in synchronization. *Consciousness and Cognition, 16*(1), 102–111. https://doi.org/10.1016/j.concog.2005.12.004.

Keller, P. E., Novembre, G., & Hove, M. J. (2014). Rhythm in joint action: Psychological and neurophysiological mechanisms for real-time interpersonal coordination. *Philosophical Transactions of the Royal Society of London Series B Biological Sciences, 369*(1658), 20130394–20130394. https://doi.org/10.1098/rsbl.2009.0670.

Kirschner, S., & Tomasello, M. (2009). Joint drumming: Social context facilitates synchronization in preschool children. *Journal of Experimental Child Psychology, 102*(3), 299–314. https://doi.org/10.1016/j.jecp.2008.07.005.

Knoblich, G., & Prinz, W. (2001). Recognition of self-generated actions from kinematic displays of drawing. *Journal of Experimental Psychology: Human Perception and Performance, 27*(2), 456–465.

Koul, A., Cavallo, A., Ansuini, C., & Becchio, C. (2016). Doing it your way: How individual movement styles affect action prediction. *PLoS One, 11*(10), 1–14. https://doi.org/10.1371/journal.pone.0165297.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review, 106*(1), 119–159.

Loula, F., Prasad, S., Harber, K., & Shiffrar, M. (2005). Recognizing people from their movement. *Journal of Experimental Psychology: Human Perception and Performance, 31*(1), 210–220. https://doi.org/10.1037/0096-1523.31.1.210.

Luck, G., & Sloboda, J. A. (2009). Spatio-temporal cues for visually mediated synchronization. *Music Perception, 26*(5), 465–473. https://doi.org/10.1525/mp.2009.26.5.465.

Macerollo, A., Bose, S., Ricciardi, L., Edwards, M. J., & Kilner, J. M. (2015). Linking differences in action perception with differences in action execution. *Social Cognitive and Affective Neuroscience, 10*(8), 1121–1127. https://doi.org/10.1093/scan/nsu161.

Maes, P.-J., Leman, M., Palmer, C., & Wanderley, M. M. (2014). Action-based effects on music perception. *Frontiers in Psychology, 4*, 1–14. https://doi.org/10.3389/fpsyg.2013.01008.

Manning, F., & Schutz, M. (2013). "Moving to the beat" improves timing perception. *Psychonomic Bulletin and Review, 20*(6), 1133–1139. https://doi.org/10.3758/s13423-013-0439-7.

Manning, F. C., & Schutz, M. (2016). Trained to keep a beat: Movement-related enhancements to timing perception in percussionists and non-percussionists. *Psychological Research, 80*(4), 532–542. https://doi.org/10.1007/s00426-015-0678-5.

Matheson, H. E., & McMullen, P. A. (2010). Neuropsychological dissociations between motion and form perception suggest functional organization in extrastriate cortical regions in the human brain. *Brain and Cognition, 74*(2), 160–168. https://doi.org/10.1016/j.bandc.2010.07.009.

Novembre, G., & Keller, P. E. (2014). A conceptual review on action-perception coupling in the musicians' brain: What is it good for? *Frontiers in Human Neuroscience, 8*, 1–11. https://doi.org/10.3389/fnhum.2014.00603.

Novembre, G., Ticini, L. F., Schutz-Bosbach, S., & Keller, P. E. (2012). Distinguishing self and other in joint action. Evidence from a musical paradigm. *Cerebral Cortex, 22*(12), 2894–2903. https://doi.org/10.1093/cercor/bhr364.

Novembre, G., Ticini, L. F., Schütz-Bosbach, S., & Keller, P. E. (2014). Motor simulation and the coordination of self and other in real-time joint action. *Social Cognitive and Affective Neuroscience, 9*(8), 1062–1068. https://doi.org/10.1093/scan/nst086.

Ono, K., Nakamura, A., & Maess, B. (2015). Keeping an eye on the conductor: Neural correlates of visuo-motor synchronization and musical experience. *Frontiers in Human Neuroscience, 9*, 1–12. https://doi.org/10.3389/fnhum.2015.00154.

Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: The Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in Systems Neuroscience, 8*, 1–14. https://doi.org/10.3389/fnsys.2014.00057.

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology, 9*(2), 129–154.

Ragert, M., Schroeder, T., & Keller, P. E. (2013). Knowing too little or too much: The effects of familiarity with a co-performer's part on interpersonal coordination in musical ensembles. *Frontiers in Psychology, 4*, 368. https://doi.org/10.3389/fpsyg.2013.00368.

Repp, B. H. (1999). Control of expressive and metronomic timing in pianists. *Journal of Motor Behavior, 31*(2), 145–164.

Repp, B. H. (2010). Sensorimotor synchronization and perception of timing: Effects of music training and task experience. *Human Movement Science, 29*(2), 200–213. https://doi.org/10.1016/j.humov.2009.08.002.

Repp, B. H., & Keller, P. E. (2010). Self versus other in piano performance: Detectability of timing perturbations depends on personal playing style. *Experimental Brain Research, 202*(1), 101–110. https://doi.org/10.1007/s00221-009-2115-8.

Repp, B. H., & Knoblich, G. (2004). Perceiving action identity. *Psychological Science, 15*, 604–609.

Repp, B. H., & Su, Y.-H. (2013). Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonomic Bulletin and Review, 20*(3), 403–452. https://doi.org/10.3758/s13423-012-0371-2.

Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. *Nature Reviews Neuroscience, 11*(4), 264–274. https://doi.org/10.1038/nrn2805.

Ross, J. M., Iversen, J. R., & Balasubramaniam, R. (2016). Motor simulation theories of musical beat perception. *Neurocase*. https://doi.org/10.1080/13554794.2016.1242756.

Ruspantini, I., D'Ausilio, A., Mäki, H., & Ilmoniemi, R. J. (2011). Some considerations about the biological appearance of pacing stimuli in visuomotor finger-tapping tasks. *Cognitive Processing, 12*(2), 215–218. https://doi.org/10.1007/s10339-011-0391-2.

Saygin, A. P., & Stadler, W. (2012). The role of appearance and motion in action prediction. *Psychological Research, 76*(4), 388–394. https://doi.org/10.1007/s00426-012-0426-z.

Schütz-Bosbach, S., & Prinz, W. (2007). Perceptual resonance: Action-induced modulation of perception. *Trends in Cognitive Sciences, 11*(8), 349–355. https://doi.org/10.1016/j.tics.2007.06.005.

Sevdalis, V., & Keller, P. E. (2010). Cues for self-recognition in point-light displays of actions performed in synchrony with music. *Consciousness and Cognition, 19*(2), 617–626. https://doi.org/10.1016/j.concog.2010.03.017.

Sevdalis, V., & Keller, P. E. (2011). Perceiving performer identity and intended expression intensity in point-light displays of dance. *Psychological Research, 75*(5), 423–434. https://doi.org/10.1007/s00426-010-0312-5.

Sofianidis, G., Hatzitaki, V., Grouios, G., Johannsen, L., & Wing, A. (2012). Somatosensory driven interpersonal synchrony during rhythmic sway. *Human Movement Science, 31*(3), 553–566. https://doi.org/10.1016/j.humov.2011.07.007.

Springer, A., Parkinson, J., & Prinz, W. (2013). Action simulation: Time course and representational mechanisms. *Frontiers in Psychology, 4*, 1–20. https://doi.org/10.3389/fpsyg.2013.00387.

Stadler, W., Springer, A., Parkinson, J., & Prinz, W. (2012). Movement kinematics affect action prediction: Comparing human to non-human point-light actions. *Psychological Research, 76*(4), 395–406. https://doi.org/10.1007/s00426-012-0431-2.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods Instruments and Computers, 31*(1), 137–149.

Su, Y.-H. (2014). Peak velocity as a cue in audiovisual synchrony perception of rhythmic stimuli. *Cognition, 131*(3), 330–344. https://doi.org/10.1016/j.cognition.2014.02.004.

Su, Y.-H. (2016a). Visual tuning and metrical perception of realistic point-light dance movements. *Scientific Reports, 6*, 1–12. https://doi.org/10.1038/srep22774.

Su, Y.-H. (2016b). Sensorimotor synchronization with different metrical levels of point-light dance movements. *Frontiers in Human Neuroscience, 10*, 1. https://doi.org/10.1525/mp.2011.29.1.65.

Su, Y.-H., & Pöppel, E. (2012). Body movement enhances the extraction of temporal structures in auditory sequences. *Psychological Research, 76*(3), 373–382. https://doi.org/10.1007/s00426-011-0346-3.

Su, Y.-H., & Salazar-Lopez, E. (2016). Visual timing of structured dance movements resembles auditory rhythm perception. *Neural Plasticity*. https://doi.org/10.1007/s00221-011-2835-4.

Thompson, J. C., & Baccus, W. (2012). Form and motion make independent contributions to the response to biological motion in occipitotemporal cortex. *NeuroImage, 59*(1), 625–634. https://doi.org/10.1016/j.neuroimage.2011.07.051.

Toiviainen, P., Luck, G., & Thompson, M. R. (2010). Embodied meter: Hierarchical eigenmodes in music-induced movement. *Music Perception, 28*(1), 59–70. https://doi.org/10.1525/mp.2010.28.1.59.

Wöllner, C. (2012). Self-recognition of highly skilled actions: A study of orchestral conductors. *Consciousness and Cognition, 21*(3), 1311–1321. https://doi.org/10.1016/j.concog.2012.06.006.

Wöllner, C. (2013). How to quantify individuality in music performance? Studying artistic expression with averaging procedures. *Frontiers in Psychology, 4*, 1–3. https://doi.org/10.3389/fpsyg.2013.00361.

Wöllner, C., Deconinck, F. J. A., Parkinson, J., Hove, M. J., & Keller, P. E. (2012). The perception of prototypical motion: Synchronization is enhanced with quantitatively morphed gestures of musical conductors. *Journal of Experimental Psychology: Human Perception and Performance, 38*(6), 1390–1403. https://doi.org/10.1037/a0028130.