

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Toiviainen, Petri; Carlson, Emily

Title: Embodied Meter Revisited : Entrainment, Musical Content, and Genre in Music-Induced Movement

Year: 2022

Version: Published version

Copyright: © 2022 the Regents of the University of California

Rights: In Copyright

Rights url: <http://rightsstatements.org/page/InC/1.0/?language=en>

Please cite the original version:

Toiviainen, P., & Carlson, E. (2022). Embodied Meter Revisited : Entrainment, Musical Content, and Genre in Music-Induced Movement. *Music Perception*, 39(3), 249-267.

<https://doi.org/10.1525/mp.2022.39.3.249>

EMBODIED METER REVISITED: ENTRAINMENT, MUSICAL CONTENT, AND GENRE IN MUSIC-INDUCED MOVEMENT

PETRI TOIVIAINEN, & EMILY CARLSON
University of Jyväskylä, Jyväskylä, Finland

PREVIOUS RESEARCH HAS SHOWN THAT HUMANS tend to embody musical meter at multiple beat levels during spontaneous dance. This work that been based on identifying typical periodic movement patterns, or eigenmovements, and has relied on time-domain analyses. The current study: 1) presents a novel method of using time-frequency analysis in conjunction with group-level tensor decomposition; 2) compares its results to time-domain analysis, and 3) investigates how the amplitude of eigenmovements depends on musical content and genre. Data comprised three-dimensional motion capture of 72 participants' spontaneous dance movements to 16 stimuli including eight different genres. Each trial was subjected to a discrete wavelet transform, concatenated into a trial-space-frequency tensor and decomposed using tensor decomposition. Twelve movement primitives, or eigenmovements, were identified, eleven of which were frequency locked with one of four metrical levels. The results suggest that time-frequency decomposition can more efficiently group movement directions together. Furthermore, the employed group-level decomposition allows for a straightforward analysis of interstimulus and interparticipant differences in music-induced movement. Amplitude of eigenmovements was found to depend on the amount of fluctuation in the music in particularly at one- and two-beat levels.

Received: November 17, 2020, accepted August 18, 2021.

Key words: music and dance, entrainment, music information retrieval, genre, movement

THE TERM *RHYTHMIC MOVEMENT* CAN BE USED to describe motor behaviors ranging from the beating flagella by which microorganisms move through water to the drumming of a woodpecker against a hollow tree. In contrast with discrete movements such as reaching and grasping, rhythmic movement is more efficient and accurate (Smits-Engelsman et al., 2002) and requires less cortical control (Schaal

et al., 2004). Rhythmic movement pervades everyday human life in forms we share in common with our phylogenetic ancestors, such as walking, chewing our food, or knocking on a closed door.

While rhythmic movement is found across the animal kingdom, humans possess a unique ability to adapt such movement to be in time with an external signal (Bispham, 2018). This process of entrainment is manifested in the precisely simultaneous steps of a marching band, the clapping and chanting games of children at a playground, and, more subtly, in the ebb and flow of speech between conversation partners (Hawkins, 2014; Ogden & Hawkins, 2015). Research suggests a significant relationship between this ability and our similarly unparalleled abilities for social cognition including empathy, shared intentionality, and prosocial behavior (Feldman, 2006; Herrmann et al., 2007; Kirschner & Tomasello, 2010; Tomasello, 2020). While examples of rhythmic entrainment have been identified in a small number of nonhuman species (Merchant & Honing, 2014; Patel et al., 2009), these are comparatively crude when compared the full breadth of human rhythmic entrainment, which can be as simple as tapping a finger to a metronome or as complex as dance movements which engage the whole body and multiple beat levels.

The complexity of human entrainment is manifest in and, most likely, profoundly related to the rhythmic complexity and ordered structures found in human music. A significant and pervasive aspect of this is the concept of meter. In the Western music theoretical tradition, meter is defined, for example, by the Grove Dictionary of Music as a “temporal hierarchy of subdivisions, beats and bars that is maintained by performers and inferred by listeners which functions as a dynamic temporal framework for the production and comprehension of musical durations” (London, 2001). London goes on to suggest that meter can therefore be understood an aspect of human behavior, rather than of music per se. It is not necessary, however, to parse such a dense explanation in order to understand what is meant by meter; a crowd gathered in a sports stadium experiences meter when stomping and clapping along to Queen’s “We Will Rock You,” as does an elderly couple swept along in dance by the strains of *The Beautiful Blue*

Danube. Meter provides structure to a steady beat. For example, a common pattern in Western music is that of four-beat structures. In such a sequence, the first beat receives the most stress, the third beat receives a smaller amount of stress, and the second and fourth beats are relatively unstressed. The waltz, a three-beat meter in which the first beat is stressed provides another familiar example. Musical rhythm is derived from the hierarchical segmentation of such beat and bar structures.

Fitch (2016) has noted that “meter cannot be properly understood without reference to movement and dance” (p. 2). Music is, in many ways, inextricable from human movement; dance of some form is among very few cultural universals in music identified by ethnomusicological studies (Nettl, 2001), and movement is among the most commonly reported responses to heard music (Lesaffre et al., 2008). Not only do we naturally move in response to music, even from a very young age (Eerola et al., 2006), such movements have been shown to vary in relation to rhythmic and timbral music features of whatever music is heard (Burger et al., 2013). Movement appears even to play a role in our perception of music; Phillips-Silver and Trainor (2007), for example, showed that when adults were trained to move to a metrically ambiguous (unaccented) rhythmic pattern in a way that reflected either a march or a waltz, they subsequently identified unambiguous waltz or march patterns as similar to the previously heard music.

Such findings are frequently framed in reference to embodied cognition, a philosophical and research perspective that emphasizes a two-way relationship between cognition and bodily states, movements, and postures (Mahon, 2015). Leman (2008) has defined a theory of embodied music cognition in which music-induced movements are seen as embodied resonances with musical fluctuations, which ultimately allow listeners to comprehend and experience empathic reactions to heard music. That is, listeners become entrained to heard music through covert (mental or “internal”) or overt (bodily) imitation of, and synchronization with, musical sounds; these movements in imitation of music bring to mind emotional states associated with similar movements, allowing the listener to empathically experience emotion expressed by the music (Leman, 2008, p. 122). This description borrows heavily from theoretical models of empathy between humans, in which the visual stimulus of another person’s movement triggers covert imitation in the observer, and thus the understanding and even experience of the observed person’s cognitive or emotional state (Zahavi, 2001).

Given that spatial and temporal hierarchy is pervasive within biology (probably necessarily so, see Mobus & Kalton, 2015) it is not surprising that the temporal hierarchies of musical rhythm are reflected in, and perhaps influenced by, the hierarchical patterns of movement in the human body. Leman and Naveda (2010) used topical gesture analysis of the movements of two professional and two student dancers performing the samba and Charleston, captured using temporal-spatial motion capture, and found that their movements encoded multiple beat levels. Using principal component analysis (PCA) decomposition of motion capture data taken from participants’ free, improvised dance movements, Toiviainen et al. (2010) extracted hierarchical eigenmovements from dancers’ movements that corresponded to multiple metrical levels; a mediolateral body sway at the four-beat level (whole note), the swaying of limbs at the two-beat level (half note), and vertical bouncing at the one-beat level (quarter note). Support for this finding was found by Burger et al. (2014), although rather than performing eigenmovement decomposition they analyzed only synchronization along different movement directions. It must be noted that Toiviainen et al. (2010) used only one musical stimulus, and therefore the degree to which eigenmovements within dance movement manifest across musical styles, or vary between musical styles, is thus not yet fully understood.

In previous work (Amelynck et al., 2014; Toiviainen et al., 2010), eigenmovements were obtained using a time-domain analysis, in which the obtained time series data were subjected to PCA. This approach groups variables based on their mutual covariance structure, and therefore, when applied to time series data, they group together variables that are phase-locked, either in-phase or anti-phase. By comparison, time series that display a mutual phase shift of, say, 90 degrees do not covary and thus are not grouped together. In spontaneous dance, different body parts are often frequency-locked but have mutual phase shifts. For instance, in hip hop dance the periodic motion of the head often shows a phase difference with that of the other body parts (Sato et al., 2015). Furthermore, different movement directions of a body part may exhibit mutual phase shifts. Samba dancers, for example, often use periodic circular hand movements (Leman & Naveda, 2010; Naveda & Leman, 2010). To group together movement components that comprise such movement patterns it may be more efficient to group the movement variables based on their degree of frequency-locking instead, thus ignoring their mutual phase relationships. Time-frequency analysis provides

a method to obtain such decompositions. This kind of analysis can be carried out using, for instance, short-term fourier transform (STFT), or discrete wavelet transform (DWT), both of which provide a representation of the signal's instantaneous amplitude and phase at a given time point and frequency. The main difference between the two is that the former uses a linear frequency division while the latter uses a logarithmic one. In the present study we chose to use DWT for two reasons. First, a logarithmic frequency division corresponds better to our perception in various modalities, in line with the Weber-Fechner law, according to which the just noticeable difference in a stimulus feature is proportional to the initial stimulus. Second, DWT has been used to model various musical activities, including perception of rhythm (Smith & Honing, 2008) and melody (Velarde et al., 2016), as well as movement interaction (Eerola et al., 2018).

To extract eigenmovements, Toiviainen et al. (2010) relied on participant-level decomposition of movement data followed by clustering. That is, each participants' movements were analyzed using principal components analysis (PCA), and only afterwards compared to attempt to identify commonalities via clustering. This method, however, does not lend itself well to easy comparison between participants and musical stimuli, as these individual patterns identified by PCA are bound to vary between dancers. Comparing dancers' movement components is then likely, at least in some cases, to be an instance of the proverbial problem of comparing apples and oranges, or at least comparing the twist to the dougie. The significance of this is highlighted by the finding of Carlson et al. (2020) that individual movement patterns in free dance are indeed so unique that the application of machine learning to a set of dance data was able to identify individual dancers from a group of 72 with an accuracy rate of 94.1%, startlingly higher than the chance rate of 1.37%. Thus, if the aim of an analysis is to identify commonalities in movement patterns across dancers, it is more useful to identify commonalities across the group first, before determining how they are manifested in individual dancers.

Fortunately, such group-level decomposition analysis techniques already exist, having been usefully applied to, for example, EEG data (Huster et al., 2015; Wang et al., 2020; Wang et al., 2018) and fMRI data (Calhoun et al., 2009). With EEG data, these approaches start with a frequency decomposition based on STFT or DWT. In group-level decomposition, data is divided into its participant-specific components only after being subjected to analysis, meaning that each component is extracted from the group data as a whole. This allows

for patterns of movement to be detected across the entire group, which may be manifested to a greater or lesser degree within different individuals within the group. More importantly, since this approach estimates a set of components that is common to all subjects along with their respective strengths (analogously to PCA or factor scores), it facilitates to population-level inferences and can easily be applied to comparisons between participants, participant groups, or stimuli.

Various variants of tensor decomposition provide suitable methods to perform group-level decomposition to data obtained by applying time-frequency analysis to whole-body dance movement with several musical stimuli and multiple subjects. These methods can be regarded as generalizations of more commonly used dimensionality reduction methods such as PCA and factor analysis in the sense that they are able to decompose data arrays that have a dimensionality higher than two (the latter also referred to as matrices). Tensor decomposition of time-frequency representations of multidimensional time series have been successfully applied, for instance, in the domain of EEG and MEG analysis (Cong et al., 2015; Kolda & Bader, 2009). In the present study we will use non-negative tensor decomposition.

The current paper revisits the question of eigenmovements within complex dance movement using group-level decomposition and wavelet analysis to examine the movements of participants dancing to a variety of musical genres. The first goal is to compare decompositions performed in time- and time-frequency-domains. More specifically, the first research question is:

- 1) How do time-domain and frequency-domain decomposition of dance movement data differ in terms of their dimensionality?

Additionally, we will address the following research questions related to the nature of eigenmovements and their dependence on musical content and genre:

- 2) Which eigenmovements are most associated with entrainment (that is, are frequency-locked with the music) and which are more gestural?
- 3) How do eigenmovements in dance movement resonate with musical structure?
- 4) What are the most universal eigenmovements across musical genres, and what, if any, differences are there in salience of eigenmovements between genres?

In light of theoretical considerations presented above, we assume that frequency-domain decomposition provides a more compact decomposition (i.e., with lower

dimensionality) than time-domain decomposition. In light of previous research, we hypothesize that the existence of eigenmovements within dance movement at a range of metrical levels will be supported, and include eigenmovements related to bodily sway at the four-beat level, limb sway at the two-beat level, and vertical bounce at the beat level. Eigenmovements are expected to correspond to the hierarchical organization of the human body; that is, slower eigenmovements are expressed via larger bodily components and faster eigenmovements such as the torso and faster eigenmovements by smaller components such as the arms and hands. We also hypothesize that the strength of the eigenmovements is mostly affected by the fluctuation (i.e., strength of periodic patterns) at one-beat level in the musical stimulus, in the vicinity of the most dominant frequency of human locomotion, 2 Hz (McDougall & Moore, 2005) and preferred pulse period of 500–600 ms (Fraisse, 1982; London et al., 2019; McAuley, 2010). We did not attempt to formulate specific hypotheses about the relationship of specific eigenmovements to particular genres, but rather considered this part of the analysis as exploratory, as we did not have a priori knowledge about which eigenmovements would be identified, and previous analysis of eigenmovements in dance movement has not yet examined different genres.

Method

DATA COLLECTION

A motion capture study was designed to collect free dance movement data from participants using naturalistic (commercially available) musical stimuli representing different genres (see Procedure). Full details of the experiment can be found in Carlson et al. (2018).¹

Participants

A total of 73 participants (54 females) completed the motion capture experiment. Participants ranged in age from 19 to 40 years ($M = 25.74$, $SD = 4.72$). Thirty held Bachelor's degrees while 16 held Master's degrees. Thirty-three reported having received some formal music training; six reported one to three years, eleven reported seven to ten years, while 16 reported ten or more years of training. Seventeen participants reported having received some formal dance training; ten reported one to three years, five reported four to six

¹The motion capture data and scripts for the calculation and decomposition of wavelet tensors are available at <https://jyx.jyu.fi/handle/123456789/74858>

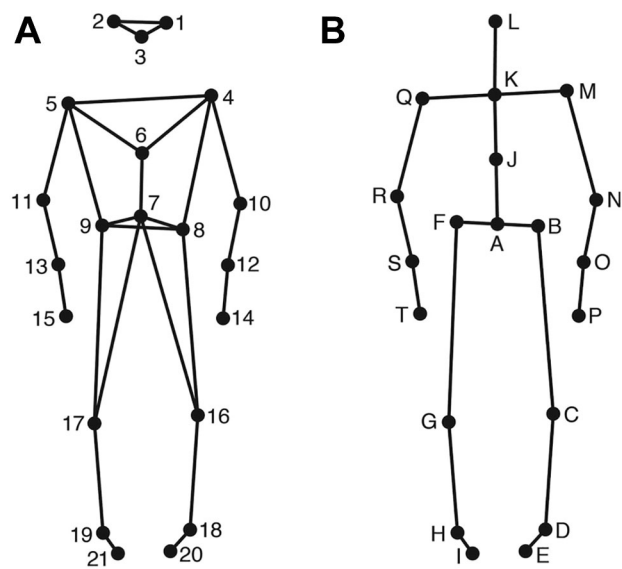


FIGURE 1. Marker and joint locations. (A) Anterior view of the marker locations as a stick figure illustration; (B) anterior view of the locations of the secondary markers/joints used in animation and analysis of the data.

years, while two reported seven to ten. Participants were of 24 different nationalities, with Finland, the United States, and Vietnam being the most represented. For attending the experiment, participants received two movie ticket vouchers each. All participants spoke and received instructions in English.

Apparatus

Participants' movements were recorded using a twelve-camera optical motion capture system (Qualisys Oqus 5+) tracking, at a frame rate of 120 Hz, the three-dimensional positions of 21 reflective markers attached to each participant. Markers were located as follows (L = left, R = right, F = front, B = back) 1: LF head; 2: RF head; 3: B head; 4: L shoulder; 5: R shoulder; 6: sternum; 7: stomach; 8: LB hip; 9: RB hip; 10: L elbow; 11: R elbow; 12: L wrist; 13: R wrist; 14: L middle finger; 15: R middle finger; 16: L knee; 17: R knee; 18: L ankle; 19: R ankle; 20: L toe; 21: R toe, visible in Figure 1A. The musical stimuli were played in a random order in each condition via four Genelec 8030A loudspeakers and a subwoofer. The direct (line-in) audio signal of the playback and the synchronization pulse transmitted by the Qualisys cameras when recording were recorded using ProTools software so as to synchronize the motion capture data with the musical stimulus afterwards.

Stimuli

The stimuli comprised 35-second excerpts from 16 musical pieces from eight genres: Blues, Country,

Dance, Jazz, Metal, Pop, Rap, and Reggae. The stimuli were selected using a computational process based on social-tagging and acoustic data. The selection pipeline was designed to select naturalistic stimuli that were uncontroversially representative of their respective genres, which would also be appropriate to use in a dance setting. To this end, a total of 2,407 tracks were collected from online music service Last.fm from those tagged by users as “danceable,” “dancing,” “head banging,” or “headbanging,” and which had been tagged with only one genre label (e.g., “Country” or “Jazz”). Tracks were retained only if they had a non-zero danceability score according to Echo Nest (the.echonest.com, an online music and data intelligence service where music categorization is determined by computational analysis of a given track’s acoustic features, including beat strength, tempo, and loudness). Two randomly selected excerpts from each of the eight genres were checked for tempo and stylistic consistency by the researchers. For a detailed description of the stimulus selection process, see Carlson et al. (2017). The musical stimuli used in this study are listed in Appendix A.

Procedure

Groups of three or four dancers at a time attended the experiment and were instructed to move freely to the randomized musical stimuli, as they might in a dance club or party setting. They moved both individually (without seeing any other dancers) and in dyads, although only individual data is considered in the current analysis. They were asked to listen to the music and move freely as they desired, staying within the marked capture space. The aim of these instructions was to create a naturalistic paradigm, such that participants would feel free to behave as they might in the real world. To limit the effects of fatigue, participants were informed that they were free to ask for a break or stop the experiment at any time, and were additionally offered water, juice, and biscuits as light refreshments.

DATA ANALYSIS

Data Preprocessing

Using the Motion Capture (MoCap) Toolbox (Burger & Toiviainen, 2013) in MATLAB, movement data of the 21 markers were first trimmed to match the duration of the musical excerpts. Gaps in the data were linearly filled. Following this, the data were transformed into a set of 20 secondary markers—subsequently referred to as joints. The locations of these 20 joints are depicted in Figure 1b. The locations of joints B, C, D, E, F, G, H, I, M, N, O, P, Q, R, S, and T are identical to the locations of one of the original markers, while the locations of the

remaining joints were obtained by averaging the locations of two or more markers; Joint A: midpoint of the two back hip markers; J: midpoint the shoulder and hip markers; K: midpoint of shoulder markers; and L: midpoint of the three head markers.

For each trial, the motion capture data were trimmed to contain the interval between 10 and 20 s from the beginning of the recording. Subsequently, the data were transformed to a local coordinate system in which, for each frame, the origin was located at the vertical projection on the floor level of the midpoint between the ankle markers (H and D) and the medio-lateral axis was perpendicular to the line joining the hip markers F and B. Finally, the velocities of each joint and direction were estimated using numerical differentiation with a Savitzky-Golay smoothing FIR filter with a window length of seven samples and a polynomial order of two.

Wavelet Transform

For each trial (i.e., subject and stimulus), the velocity data of each spatial component (i.e., for each joint in each of the three directions) were subjected to DWT using Morse wavelets (the most commonly used wavelet type) with sixteen voices per octave, and ranging over eight octaves. We used Matlab (version R2020b) and wavelet Toolbox (version 5.5) for the calculations. The obtained wavelet transforms were subsequently transformed from absolute frequencies to frequencies relative to the tactus beat frequency of each respective stimulus, with the range of relative frequencies spanning four octaves, from $(1/8)\sqrt{2}$ times beat frequency to $2\sqrt{2}$ times beat frequency, thus covering the four, two, one, and half beat levels. Following this, the wavelet transforms of each spatial component were stacked to form a three-way tensor with dimensions of 65 (frequency) x 1201 (time) x 60 (space) for each subject and stimulus presentation (see Figure 2A).

For the purpose of subsequent tensor decomposition, the absolute values of the wavelet tensor were then averaged across time, yielding a 65 x 60 matrix of wavelet spectra for each subject and stimulus presentation (see Figure 2B). Averaging across time was performed to allow stability and convergence of the subsequent tensor decomposition, as including the time dimension in the data was found to prevent the tensor decomposition from converging to a stable solution (see, e.g., Wang et al., 2018, for a discussion of this issue). Subsequently, these wavelet spectra were concatenated to yield a third-order non-negative tensor $\underline{W} \in \mathfrak{N}_+^{65 \times 60 \times 1168}$, consisting of frequency, space, and trial factors (see Figure 2C). This kind of tensor allowed us to extract how different

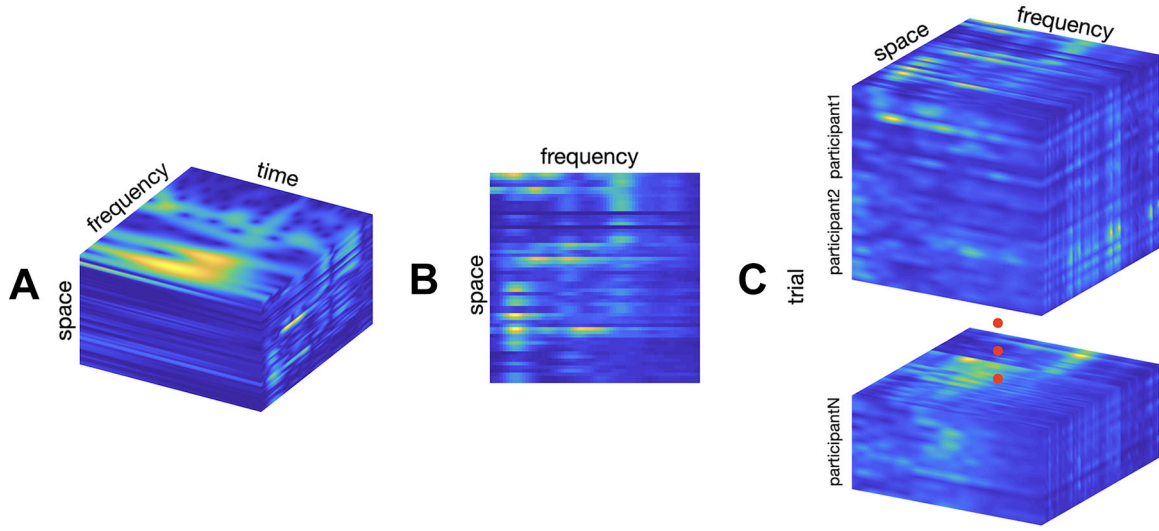


FIGURE 2. Forming of the group wavelet tensor. (A) Space-by-frequency-by-time wavelet tensor of a single participant; (B) space-by-frequency matrix obtained by averaging the wavelet tensor across the time dimension; (C) space-by-frequency-by-trial group tensor obtained by concatenating space-by-frequency matrices for all participants and stimulus presentations.

body parts move in different directions at different frequencies, and how these patterns vary across the different musical stimuli.

Tensor Decomposition

To extract eigenmovements of movement, nonnegative polyadic tensor decomposition (Kim & Park, 2012), as implemented in the Matlab Tensor Toolbox (Bader & Kolda, 2019) and the Nonnegative Matrix and Tensor Factorization Algorithms Toolbox (Kim & Park, 2014) was applied to the data tensor \underline{W} . This decomposition method attempts to find components that minimize the cost function

$$\| \underline{W} - \sum_{i=1}^m \lambda_i \mathbf{f}_i \otimes \mathbf{s}_i \otimes \mathbf{t}_i \| \quad (1)$$

where the vectors $\mathbf{f}_i, \mathbf{s}_i, \mathbf{t}_i \geq 0$ are the frequency, space, and trial factors of component i , respectively, and \otimes denotes the outer product. For each component, the frequency factor represents the amplitude spectrum, the space factor the individual contribution of each joint and movement direction, and the trial factor the amplitude of the respective eigenmovement in each participant-stimulus combination. The tensor decomposition is depicted schematically in Figure 3. The extracted components will be subsequently referred to as eigenmovements.

Model order selection, that is, selection of the number of components used in the decomposition, was determined by maximizing its explained variance and convergence (Hu et al., 2019). To this end, the tensor decomposition was run 1000 times with a range of

model orders², and the maximal model order for which at least 95% of the runs converged was selected. This method suggested 12 components, which were used in all subsequent analyses. The obtained 12-component decomposition yielded a relative error (i.e., cost function divided by the norm of the tensor) of 0.31 and contained 79% of variance in the data.

Intrinsic Dimensionality Estimation

The amount to which a decomposition compresses data can be estimated in various ways. In the present study we used effective dimensionality (Del Giudice, 2020), which we estimated by the Rényi entropy of the eigenvalue spectrum (Pirkl et al., 2012):

$$n = \frac{\left(\sum_{i=1}^N \lambda_i \right)^2}{\sum_{i=1}^N \lambda_i^2} \quad (2)$$

where λ_i is the i 'th eigenvalue. Small effective dimensionality indicates a compact decomposition of data in the sense that a large proportion of the variance in the data is accommodated in a small number of components. A specific example of time- and time-frequency-domain decompositions of movement data and their intrinsic dimensionalities is provided in Appendix B.

² Tensor decomposition, like many other data decomposition methods such as independent component analysis and factor analysis, are based on numerical optimization with random initial conditions, and different decompositions often yield slightly different results. Moreover, if the number of components is too large, the procedure may fail to converge

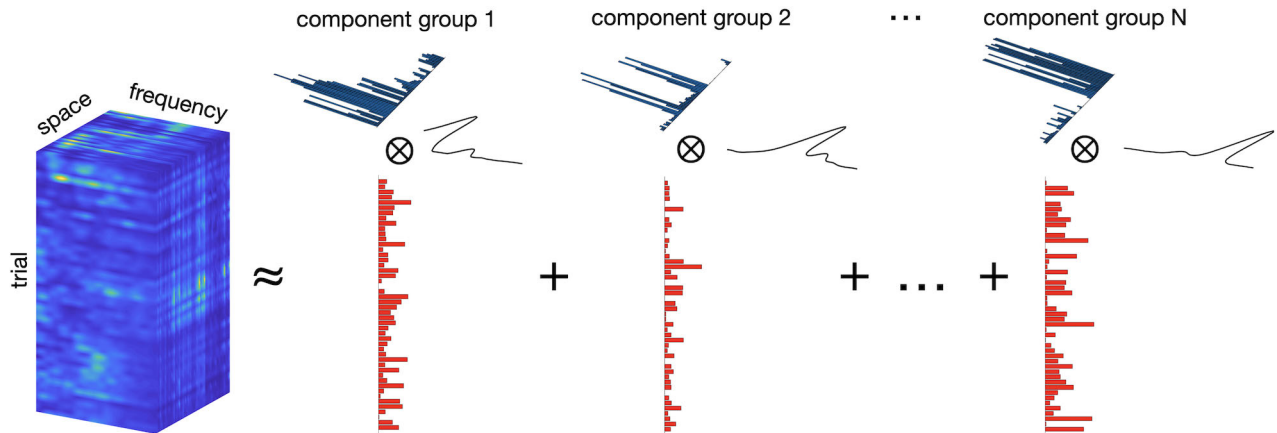


FIGURE 3. Decomposition of the wavelet tensor into a sum of outer products of frequency (black lines), space (black bar graphs), and trial (red bar graphs) factors.

Musical Features

To investigate the relationship between the amplitude of eigenmovements and acoustic properties of the music stimuli, we estimated for each stimulus the intensity of pulsation at different metrical levels using the fluctuation spectrum (Pampalk et al., 2002) using the mirfluctuation function of MIRTtoolbox (Lartillot & Toiviainen, 2007) with the “summary” option. A fluctuation spectrum indicates the strength of periodicities in the music as a function of frequency between 0 and 10 Hz. For the purpose of subsequent analyses, we used fluctuation values at frequencies corresponding to four-, two-, one-, and half-beat levels of each stimulus. Figure 4 shows fluctuation spectra for two stimuli used in the study.

Statistical Analysis on Genre Differences

For statistical analysis of differences in the amplitude of different eigenmovements between genres, the trial factor of each component was reorganized to form a participant-by-stimulus matrix. Subsequently the values for the two stimuli representing each genre were averaged to yield a participant-by-genre matrix (73 x 8) for each of the twelve eigenmovements. These matrices were subjected to Friedman tests with Bonferroni correction to assess the degree to which the respective amplitude values varied across genres.

Results

INTRINSIC DIMENSIONALITY OF TIME-DOMAIN AND TIME-FREQUENCY-DOMAIN DECOMPOSITIONS

We compared the effective dimensionalities of the time-domain and time-frequency-domain representations by

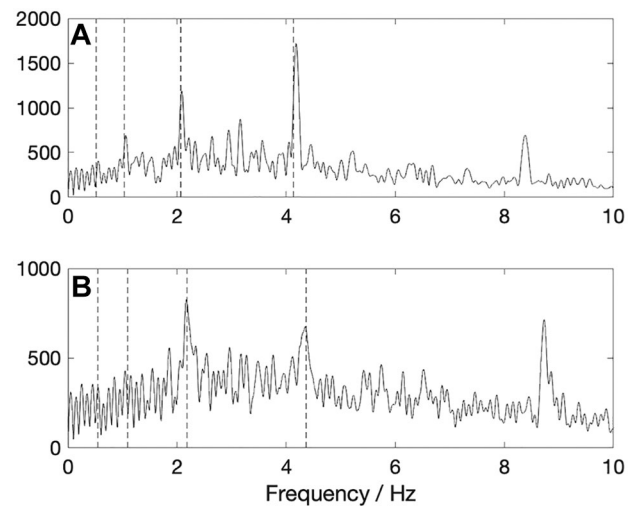


FIGURE 4. Fluctuation spectra of the excerpts of (A) *My Maria* by Brooks & Dunn; and (B) *Redneck* by Lamb of God used in this study. Dashed vertical lines show, from left to right, the frequencies corresponding to four-, two-, one-, and half-beat levels of each stimulus.

performing a PCA, separately for each participant and stimulus, on the time- and frequency-domain representations of the whole data as explained in sections Data Preprocessing and Wavelet Transform, and calculating effective dimensionality of the obtained eigenvalue spectra using Rényi entropy. Figure 5 shows the distribution of the effective dimensionality values for each domain. The mean effective dimensionalities were 14.40 ($SD = 2.61$) and 5.68 ($SD = 1.00$) for time- and time-frequency-domain, respectively, $t(1167) = 126.58$, $p < .0001$. Consequently, the time-frequency-domain representation provides a more compact decomposition

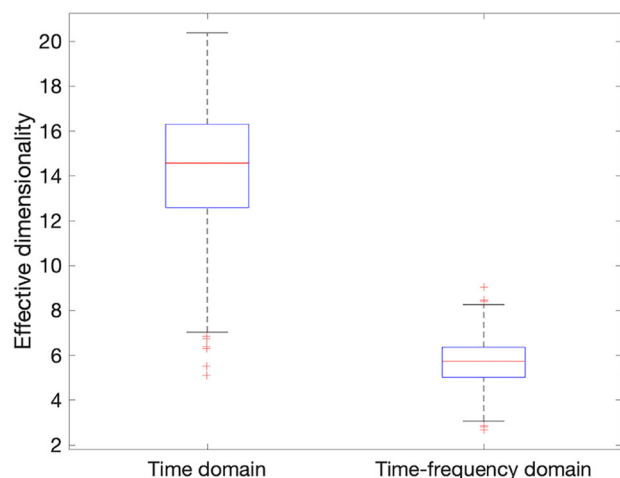


FIGURE 5. Distribution of effective dimensionality values across participants and stimuli for time- and frequency-domain PCA decompositions of the data.

of the data in the sense that more variance is contained in fewer number of components.

EIGENMOVEMENTS

Example animations of extracted eigenmovements are available at the URL <https://jyx.jyu.fi/handle/123456789/74855> and the method for reconstructing them is presented in Appendix C. Each animation comprises twelve eigenmovements as represented in the movements of four different dancers for a given stimulus.

Frequency Factors

The frequency factors of the twelve extracted eigenmovements are depicted in Figure 6A. In the sub-plots, the abscissa is logarithmic and the labels indicate frequency relative to that of the main beat (tactus). Figure 6B shows the peak frequencies of each eigenmovement. As can be seen, 11 out of the 12 eigenmovements (depicted as black bars) are centered around one of the metrical levels, with $\pm 5\%$ tolerance. Thus, the majority of the participants' movement energy is frequency-locked (entrained to the musical beat) at one of the metrical levels. The frequency factors of the components, however, vary in terms of the width of energy distribution, with some frequency components showing narrower distributions than others. Thus, the eigenmovements differ in terms of their degree of frequency-locking (see below).

Spatial Factors

Figure 7 displays the spatial factors of each component, divided into medio-lateral (mp, black bars), antero-

posterior (ap, dark grey bars), and vertical (v, light grey bars) movement directions, and Table 1 summarizes for each eigenmovement the most prominent body part and movement direction, based on the magnitude of the spatial factors, as well as the metrical level in beats corresponding to the peak in the frequency factor. As can be seen, each of the one-, two-, and four-beat metrical levels is associated with both hand- and torso-based eigenmovements with several different movement directions.

Frequency Locking of Eigenmovements

To assess the degree of frequency locking of each eigenmovement, we calculated the relative Shannon entropy of the frequency modes. A frequency spectrum with low entropy has a peaked distribution; that is, high concentration of energy around a frequency, and thus represents high degree of frequency-locking and quasi-periodic structure. The entropies are displayed in Figure 8. In accordance with Figure 1, the eigenmovements differ in terms of their entropy. In particular, eigenmovements 1, 5, and 9 have the smallest entropies and thus are the most frequency-locked. It is notable that these eigenmovements correspond closely to those found in our earlier study (Toiviainen et al., 2010): medio-lateral sway of the torso on four-beat level, anteroposterior movement of hands on two-beat level, and vertical bouncing of the torso on one-beat level.

Eigenmovement Amplitude and Rhythmic Structure

To investigate the relationship between the amplitude of eigenmovements and the rhythmic structure, we averaged the eigenmovement amplitudes across participants and correlated these averages with the fluctuation values of the stimuli at half-, one-, two-, and four-beat levels. Figure 9 shows the obtained correlations. As can be seen, the correlations tend to be positive, indicating that overall the eigenmovements tend to become more salient when the music contains a high amount of fluctuation. In particular, fluctuation at one-beat level has the strongest effect on eigenmovement amplitude. Of the eigenmovements with the highest frequency coupling, modes 1 and 5 tend to show low correlations for all fluctuation levels, suggesting that they are resonating less with the musical structure, and thus might be more serving the function of beat maintenance at their respective metrical levels (4 and 2 beats). Eigenmovement 9, however, displays high correlations with the fluctuation strength at several metrical levels, suggesting that it can rather be considered as a resonance phenomenon.

Eigenmovement Amplitude and Genre

To investigate the differences in eigenmovement amplitudes between musical genres, we performed a Friedman

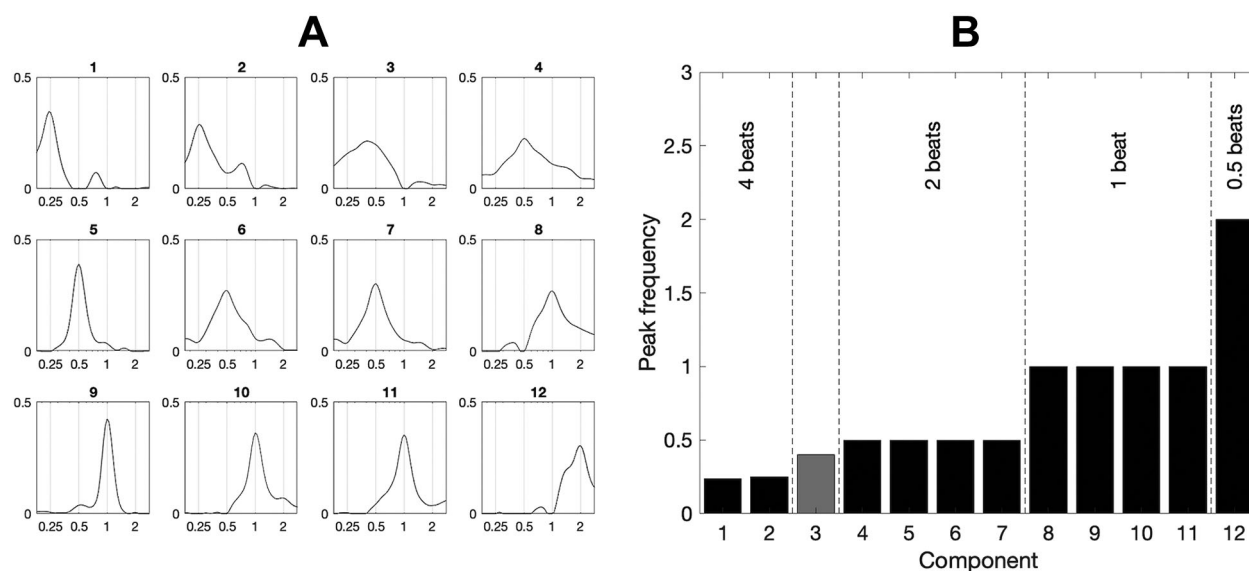


FIGURE 6. A) Frequency factors of eigenmovements. The abscissa is logarithmic and the labels indicate frequency relative to that of the main beat (tactus); B) peak frequencies of each component.

test, that is, a nonparametric one-way ANOVA. To this end, we averaged, for each participant and mode, amplitude values across the two stimuli representing each genre. The result is shown in Figure 10. We used Bonferroni correction to correct for multiple comparisons.

As can be seen in Figure 10, eigenmovements 2, 5, and 8 show highest inter-genre differences for metrical levels on four, two, and one beats, respectively. These eigenmovements are mostly associated with horizontal hand and lower arm movements. Moreover, eigenmovements 2 and 8 have high spectral entropy (see Figure 8), indicating that they manifest a low degree of frequency-locking. This suggests that genre-specific movement patterns could be associated with hand gestures. Additionally, eigenmovement 12 shows significant inter-genre differences, suggesting that genre-specific movements can also be characterized by the presence or absence of fast movements at the half-beat level.

Figure 11 displays the medians and interquartile ranges for amplitudes of eigenmovements 2, 5, 8, and 12. Most notably, jazz tends to have high amplitude values for eigenmovements 5, 8, and 12, suggesting that it is associated with fast movements of hands in particular. Moreover, the country stimuli display an opposite pattern, suggesting that it elicits movement on slower metrical levels. The metal stimuli overall tend to have a low amplitude for all these eigenmovements, suggesting they elicit less of these kinds of gestural movements. The medians and interquartile ranges of all 12 eigenmovements are displayed in Appendix D.

Discussion

In the current paper, we used group-level decomposition paired with time-frequency analysis using discrete wavelet transform to investigate patterns in spontaneous dance movement across eight distinct musical genres. We found that performing decomposition in time-frequency-domain yielded more compact (i.e., lower-dimensional) representations of dance movement over time-domain decomposition. This is because the former groups variables based on their frequency-locking, instead of phase-locking, thus allowing body parts and movement directions with similar frequency but differing phases to be grouped together. Therefore, if the aim of research is to find movement patterns that comprise body parts and movement directions that are frequency-locked but not necessarily phase-locked, performing decomposition in the time-frequency domain can be considered to be superior to time-domain-based decomposition. If, however, we are interested in phase locking of movement, an example of which could be quantifying synchronization accuracy of music-induced movement with a musical stimulus, time-domain-based decomposition methods may prove more efficient.

This novel method allowed us to identify twelve movement primitives (eigenmovements) that appear consistently at the group level, the metrical levels with which each eigenmovement was associated, and the degree to which eigenmovements were rhythmically

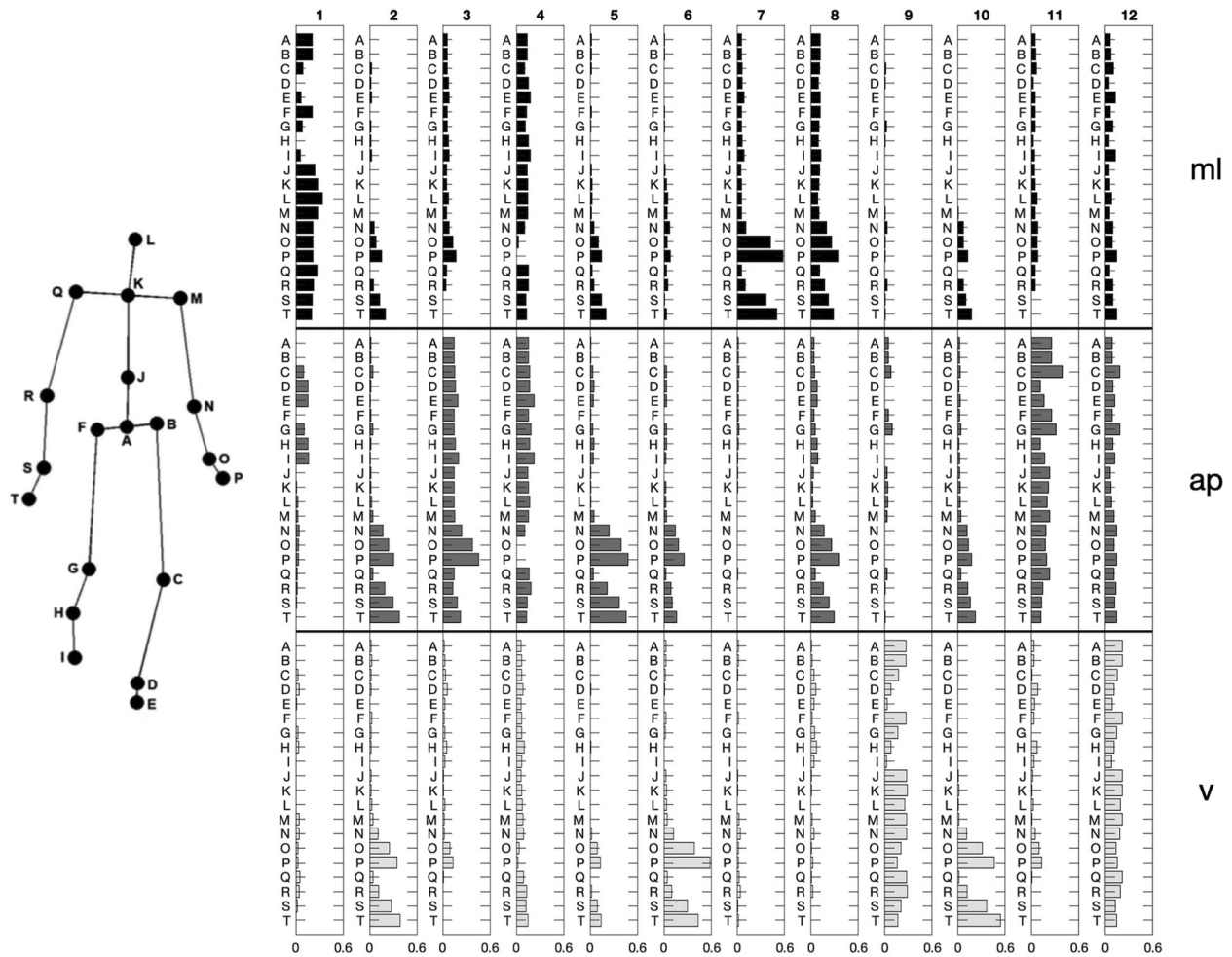


FIGURE 7. Values of spatial factors (horizontal axes) of components 1–12 for each joint and movement direction (vertical axes). ml = mediolateral; ap = anteroposterior; v = vertical.

TABLE 1. Summary of Spectrospatial Properties of the Eigenmovements

| Eigenmovement | Body Part | Direction | Beat level |
|---------------|-----------|-----------|------------|
| 1 | Torso | ml | 4 |
| 2 | Hands | ap/v | 4 |
| 3 | Hands | ap | — |
| 4 | Torso | ml/ap | 2 |
| 5 | Hands | ap | 2 |
| 6 | Hands | v | 2 |
| 7 | Hands | ml | 2 |
| 8 | Hands | ml/ap | 1 |
| 9 | Torso | v | 1 |
| 10 | Hands | v | 1 |
| 11 | Torso | ap | 1 |
| 12 | Torso | v | 0.5 |

Note: Includes the most prominent body part and movement directions, as well as the beat level. ml = mediolateral; ap = anteroposterior; v = vertical.

entrained at these levels. The group-level decomposition method used allows a straightforward analysis of how the amplitude of each eigenmovement depended on the temporal structure of each musical stimulus, and comparison of the amplitude of eigenmovements between genres.

We found that spontaneous movement to music shows a hierarchical organization in the sense that it tends to be simultaneously entrained with several metrical levels. Moreover, we characterized various movement patterns, or eigenmovements, associated with these metrical levels, and showed that the amplitude of these movement patterns depends on the content of the musical stimulus. Out of the twelve identified eigenmovements, two were associated with the four-beat (whole note) level, four were associated with the

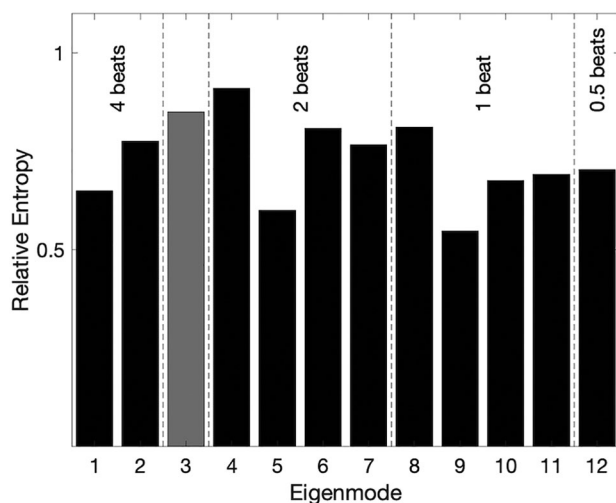


FIGURE 8. Power spectral entropies of the frequency factors of each eigenmovement.

two-beat (half note) level, four with the one-beat (quarter note) level, and one with the half-beat (eight-note) level, leaving only one eigenmovement not associated with a metrical level. This non-rhythmic mode was characterized primarily by anteroposterior movement of the torso, suggesting that it may be related to self-correction movements necessary for maintaining balance during standing and dancing (Day et al., 1997; Johnson et al., 2010). That the majority of eigenmovements did, however, appear to be entrained to metrical levels corroborates the previous findings of Toiviainen et al. (2010) and notably extend their generalizability. This finding furthermore suggests that dance movements which are not metrically entrained (that is, which are more gestural), tend not to be consistent at the group level.

While most of the eigenmovements were entrained at metrical levels, they varied in their degree of phase- and frequency-locking. The eigenmovements showing the highest frequency locking (eigenmovements 1, 5, and 9) are spatially quite similar to those identified in Toiviainen et al. (2010) using time-domain methods; the current eigenmovement 1 to the mediolateral sway of the torso at the four-beat level, eigenmovement 5 to anteroposterior movement of hands on two-beat level, and eigenmovement 9 to the vertical bouncing of the torso on one-beat level. Thus, these particular eigenmovements are predominantly associated with rhythmic entrainment to the musical stimulus at their respective metrical levels, while the other eigenmovements are more loosely associated with the rhythmic structure of the music, and are thus not easily identified

using time-domain analysis only. Most of these newly identified eigenmovements were dominated by hand movement, suggesting the upper limbs tend to show greater flexibility of movement in relation to the beat. Hand movements are associated with expressive gestures during verbal conversation (Goldin-Meadow, 2006; Wong & So, 2018), thus the flexibility of these eigenmovements may reflect expressive functionality, such as the emphasis of meaningful lyrical content. Although the current findings reflect individual movement only, previous work has found hand movement to be associated with responsiveness to a partner in dyadic dance (Carlson et al., 2018), such that in dyadic or group contexts the relative flexibility of these eigenmovements may additionally afford social entrainment (Phillips-Silver et al., 2010).

The current analysis additionally corroborated previous results from Burger et al. (2013) suggesting that a greater amount of rhythmic fluctuation in music tends to elicit more movement. This is particularly notable at the one-beat level, or a frequency of approximately 2 Hz, which has been shown to be the strongest frequency in human locomotion (McDougall & Moore, 2005) as well as corresponding closely to the range of most salient pulse sensations (Fraisse, 1982) and spontaneous tapping rate or preferred tempo (Fraisse, 1982; London et al., 2019; McAuley, 2010).

This association highlights the close relationship between human physiology, culture, and behavior, as it is certainly no accident that we tend to move to music that allows for easy motoric resonance with our preferred tempo. Eigenmovements associated with lower beat levels were less affected by rhythmic fluctuation of the music, specifically those related to mediolateral sway at the four-beat level (eigenmovement 1), and those associated with anteroposterior upper limb sway at the two-beat level (eigenmovement 5), suggesting these may represent more fundamental modes of entrainment, a standard “ground” against which the one- and half-beat level movements of the hands are more free to create a “figure.”

The difference found in eigenmovement amplitude between genres may reflect behaviors influenced by cultural norms as well as acoustic differences between stimuli. Luck et al. (2010) has shown associations between Techno, Latin, and Metal stimuli and genre-stereotypical movement patterns, while Carlson et al. (2020) showed that Metal and Jazz were the most readily identifiable genres based on dancers’ movement patterns. Specific dance moves reflecting cultural associations, such as “headbanging” for Metal or the Charleston in the case of Jazz, may contribute to

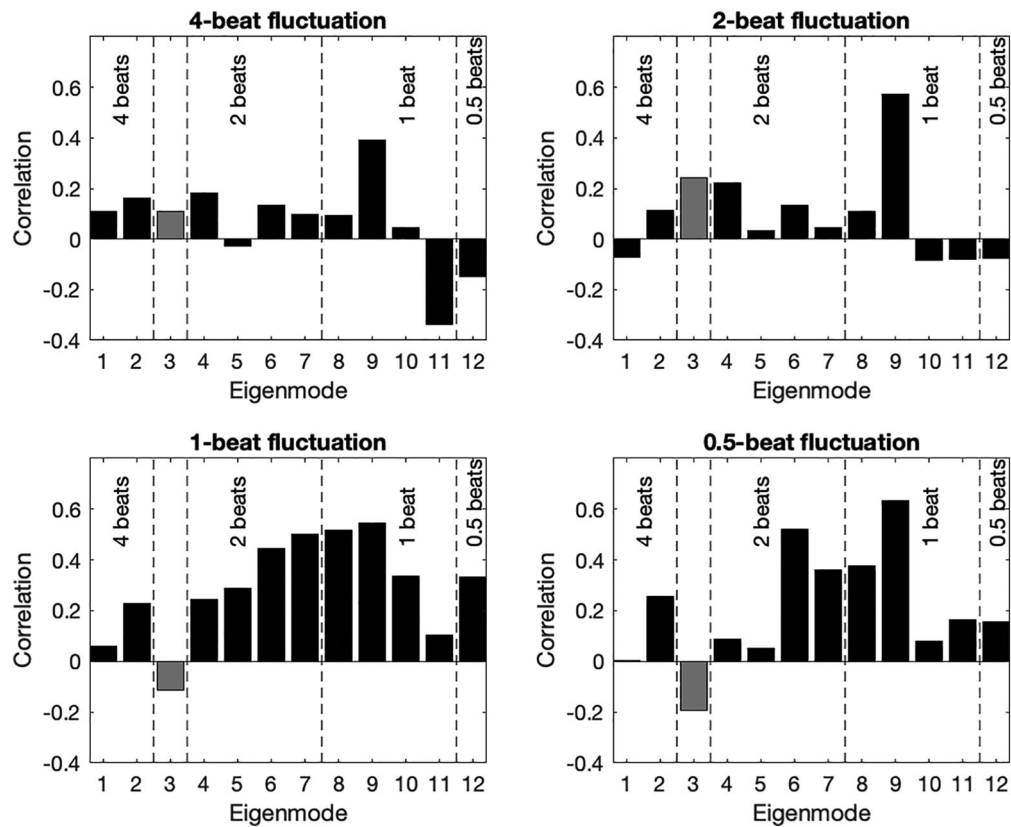


FIGURE 9. Correlations between the amplitude of each eigenmovement and fluctuation strength at different metrical levels.

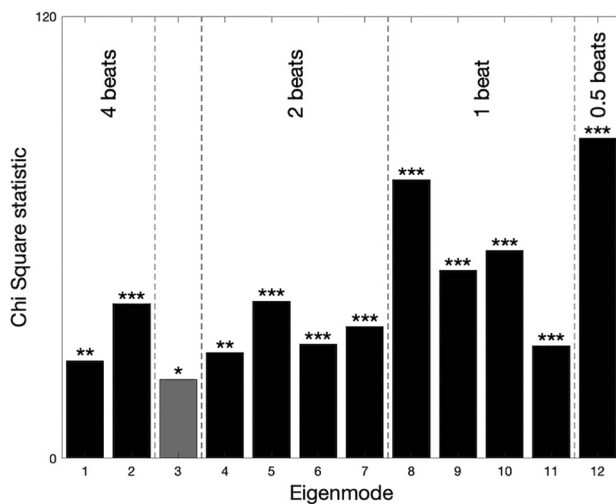


FIGURE 10. Result of Friedman test on the amplitude of each eigenmovement. * $p < .05$, ** $p < .01$, *** $p < .001$, Bonferroni corrected.

differences in the amplitude of different eigenmovements in these genres; the strong amplitude of eigenmovement 12, hand movement at the half-beat

level, would seem to support this notion. However, the influence of genre on individual movement not associated with group-level eigenmovements should be explored in future work.

The current results have both methodological and theoretical implications. Comparison of current results with earlier findings suggests that time-frequency analysis indeed allows for a more fine-grained decomposition of dance movement than do time-domain methods alone. One reason for this is that time-domain methods use correlation between time series, which only allows for signals to be grouped based on having identical or opposite phases, while time-frequency methods allow for the detection of non-zero phase differences. This also allows the degree of frequency locking with the music, or lack thereof, to be easily quantified. The use of group-level tensor decomposition presented here simultaneously identifies eigenmovements common within the group while estimating the strength of the eigenmovement for each participant and stimulus, allowing us a greater degree of confidence in generalizing results beyond the current sample.

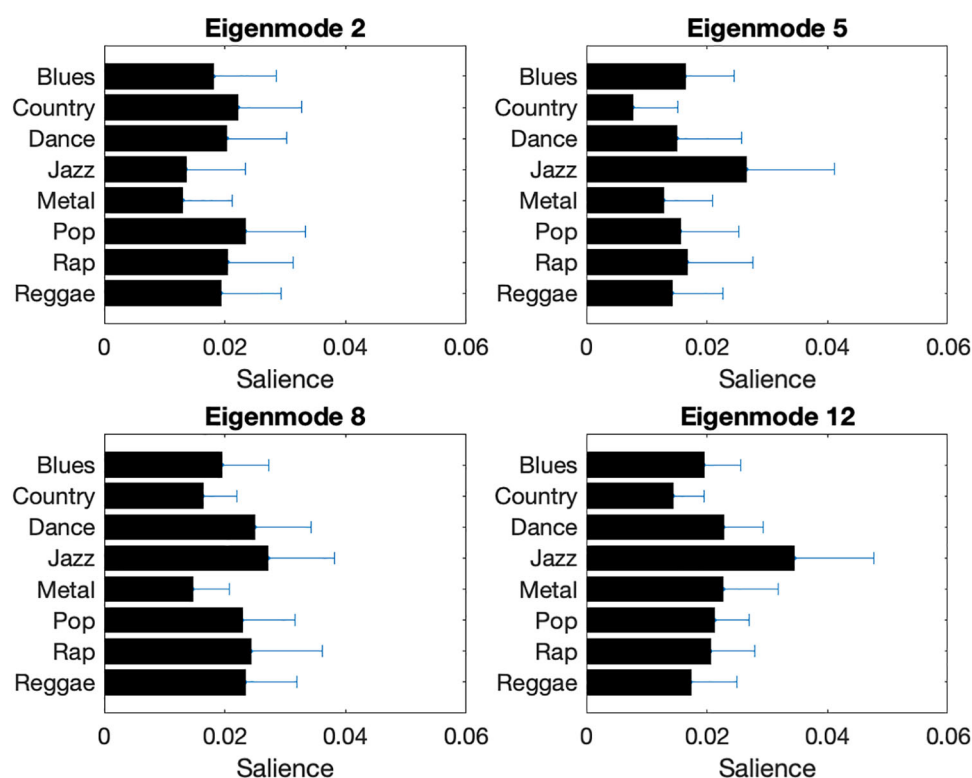


FIGURE 11. Median amplitude values and interquartile ranges per genre for eigenmovements 2, 5, 8, and 12.

The apparently universal presence of metrically hierarchical eigenmovements in spontaneous dance movement (at least within participants representing a generally Western culture) is tempting to interpret this straightforwardly as evidence that music and dance are great equalizers. However, these results should also be interpreted in the light of findings by Carlson et al. (2020), derived from the same data set as used in the current study, that cross-correlation matrixes derived from participants' three-dimensional movements could be used to identify individual dancers with a startling degree of accuracy. Thus, we are forced to understand that, without apparent conflict or paradox, spontaneous dance movement can be understood and interpreted *both* as profoundly individual and profoundly universal. The theoretical framework perhaps best suited to this understanding is that of *floating intentionality*, which has been proposed and developed by Cross (e.g., 2006, 2008, 2013). Music (and by association, dance), Cross argues, is a mode of human communication that privileges emotion and interaction over specificity of meaning; that is, music can be “experienced quite differently by different participants at the same time without the integrity of the music being significantly compromised”

(Cross, 2009, p. 185). Thus, music and dance, unlike regular speech, provides a mode of engagement that allows individuals to simultaneously, indeed synchronously, participate in an interaction that need not have identical meaning or psychosocial affordances. The current findings provide support for this idea in that they show that dancers may move in similar ways, and in time with the music (and, most likely, each other) without sacrificing individuality.

The presented approach has some potential limitations. We used a local coordinate system in the analysis to increase the degree of stationarity of the data. This approach however, ignores certain kinds of movement, such as translation and rotation. By grouping body parts and movement directions based on frequency-locking, the phase information is lost and therefore movement patterns displaying different phase relations are grouped together. For instance, the method does not make a distinction between in-phase and anti-phase hand movement. However, if these kinds of distinctions are important, the phase relations can be recovered using the method explained in Appendix C. Averaging the absolute values of wavelet transforms across time is based on the assumption that the movement patterns be

stationary in terms of their frequency and amplitude (but not necessarily phase). If this condition is not met, such as when the dancer switches from one movement pattern to another during the analysis window, it is not possible to disentangle the two patterns directly from the time-averaged representation. The group-level decomposition, while being straightforward and providing direct measures for statistical comparisons between participants and/or stimuli, disentangling within- and between-subject variation in the obtained decompositions may be difficult. For instance, a high entropy in a frequency factor may be a result of either all participants showing a low degree of frequency-locking in the respective eigenmovement, or different participants showing high degree of frequency-locking but at slightly different frequencies. Finally, the data consisted of only two pieces of music per each genre. Although they were selected using a computational algorithm to maximize their typicality, one should be cautious about drawing too general conclusions about typical dance patterns for these genres.

Of course, the degree to which this may or may not be truly universal cannot be known without future research that includes participants from majority-world cultures. Future research could also expand the

types of musical stimuli used, and assess the degree to which social context, such as dancing with a partner or in a group, affected the presence of eigenmovements in dance movement. The current results are based on spontaneous dance movements only, and thus may not generalize to choreographed or highly stylized types of dance such as Tango or Swing. Investigating the degree to which universal eigenmovements, as well as individual differences, can be captured in such contexts is necessary for developing a more thorough understanding of the relationship between the current findings and cultural norms related to musical styles. The novel methods present in the current study offer an effective means by which many such questions can be addressed and, it is hoped, will find further applications with which to contribute to our understanding of human musicality.

Author Note

This work was supported by the Academy of Finland (project number 332331).

Correspondence concerning this article should be addressed to Petri Toiviainen, P.O. Box 35(M), 40014 University of Jyväskylä, Jyväskylä, Finland. E-mail: petri.toiviainen@jyu.fi

References

- AMELYNCK, D., MAES, P.-J., MARTENS, J. P., & LEMAN, M. (2014). Expressive body movement responses to music are coherent, consistent, and low dimensional. *IEEE Transactions on Cybernetics*, *44*(12), 2288–2301. <https://doi.org/10.1109/TCYB.2014.2305998>
- BADER, B. W., & KOLDA, T. G. (2019). *MATLAB Tensor Toolbox* (version 3.1). Retrieved from <https://www.tensortoolbox.org>
- BISPHAM, J. C. (2018). *The human faculty for music: What's special about it?* University of Cambridge.
- BURGER, B., THOMPSON, M. R., LUCK, G., SAARIKALLIO, S., & TOIVIAINEN, P. (2013). Influences of rhythm- and timbre-related musical features on characteristics of music-induced movement. *Frontiers in Psychology*, *4*, 183. <https://doi.org/10.3389/fpsyg.2013.00183>
- BURGER, B., THOMPSON, M. R., LUCK, G., SAARIKALLIO, S. H. S., & TOIVIAINEN, P. (2014). Hunting for the beat in the body: On period and phase locking in music-induced movement. *Frontiers in Human Neuroscience*, *8*(November), 903. <https://doi.org/10.3389/fnhum.2014.00903>
- BURGER, B., & TOIVIAINEN, P. (2013). Mocal Toolbox - A Matlab toolbox for computational analysis of movement data. In R. Bresin (Ed.), *Proceedings of the 10th Sound and Music Computing Conference*.
- CALHOUN, V. D., LIU, J., & ADALI, T. (2009). A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. *NeuroImage*, *45*(1), S163–S172. <https://doi.org/10.1016/j.neuroimage.2008.10.057>
- CARLSON, E., BURGER, B., & TOIVIAINEN, P. (2018). Dance like someone is watching. *Music and Science*, *1*, 205920431880784. <https://doi.org/10.1177/2059204318807846>
- CARLSON, E., SAARI, P., BURGER, B., & TOIVIAINEN, P. (2017). Personality and musical preference using social-tagging in excerpt-selection. *Psychomusicology: Music, Mind and Brain*, *27*(3), 203–212. <https://doi.org/dx.doi.org/10.1037/pmu0000183>
- CARLSON, E., SAARI, P., BURGER, B., TOIVIAINEN, P., CARLSON, E., SAARI, P., & BURGER, B. (2020). Dance to your own drum: Identification of musical genre and individual dancer from motion capture using machine learning from motion capture using machine learning. *Journal of New Music Research*, *0*(0), 1–16. <https://doi.org/10.1080/09298215.2020.1711778>
- CONG, F., LIN, Q.-H., KUANG, L.-D., GONG, X.-F., ASTIKAINEN, P., & RISTANIEMI, T. (2015). Tensor decomposition of EEG signals: A brief review. *Journal of Neuroscience Methods*, *248*, 59–69. <https://doi.org/10.1016/J.JNEUMETH.2015.03.018>

- CROSS, I. (2006). Music, cognition, culture, and evolution. *Annals of the New York Academy of Sciences*, 930(1), 28–42. <https://doi.org/10.1111/j.1749-6632.2001.tb05723.x>
- CROSS, I. (2008). Musicality and the human capacity for culture. *Musicae Scientiae*, 12(1_suppl), 147–167. <https://doi.org/10.1177/1029864908012001071>
- CROSS, I. (2009). The evolutionary nature of musical meaning. *Musicae Scientiae*, 13(2_suppl), 179–200. <https://doi.org/10.1177/1029864909013002091>
- CROSS, I. (2013). “Does not compute”? Music as real-time communicative interaction. *AI and Society*, 28(4), 415–430. <https://doi.org/10.1007/s00146-013-0511-x>
- DAY, B. L., SÉVERAC CAUQUIL, A., BARTOLOMEI, L., PASTOR, M. A., & LYON, I. N. (1997). Human body-segment tilts induced by galvanic stimulation: A vestibularly driven balance protection mechanism. *Journal of Physiology*, 500(3), 661–672. <https://doi.org/10.1113/jphysiol.1997.sp022051>
- DEL GIUDICE, M. (2020). Effective dimensionality: A tutorial. *Multivariate Behavioral Research*, 56(3), 527–542. <https://doi.org/10.1080/00273171.2020.1743631>
- EEROLA, T., JAKUBOWSKI, K., MORAN, N., KELLER, P. E., & CLAYTON, M. (2018). Shared periodic performer movements coordinate interactions in duo improvisations. *Royal Society Open Science*, 5(2), 171520. <https://doi.org/10.1098/rsos.171520>
- EEROLA, T., LUCK, G., & TOIVIAINEN, P. (2006). An investigation of pre-schoolers’ corporeal synchronization with music. In M. Baroni, A. R. Addressi, R. Caterina, & M. Costa (Eds.), *Proceedings of the 9th International Conference on Music Perception and Cognition*, 472–476. <https://doi.org/10.1.1.324.616>
- FELDMAN, R. (2006). From biological rhythms to social rhythms: Physiological precursors of mother-infant synchrony. *Developmental Psychology*, 42(1), 175–188. <https://doi.org/10.1037/0012-1649.42.1.175>
- FITCH, W. T. (2016). Dance, music, meter and groove: A forgotten partnership. *Frontiers in Human Neuroscience*, 10, 64. <https://doi.org/10.3389/fnhum.2016.00064>
- FRAISSE, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 148–180). Academic Press.
- GOLDIN-MEADOW, S. (2006). Talking and thinking with our hands. *Current Directions in Psychological Science*, 15(1), 34–39.
- HAWKINS, S. (2014). Situational influences on rhythmicity in speech, music, and their interaction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1658). <https://doi.org/10.1098/rstb.2013.0398>
- HERRMANN, E., CALL, J., HERNÁNDEZ-LLOREDA, M. V., HARE, B., & TOMASELLO, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science*, 317(5843), 1360–1366. <https://doi.org/10.1126/science.1146282>
- HU, G., ZHANG, Q., WATERS, A. B., LI, H., ZHANG, C., WU, J., ET AL. (2019). Tensor clustering on outer-product of coefficient and component matrices of independent component analysis for reliable functional magnetic resonance imaging data decomposition. *Journal of Neuroscience Methods*, 325, 108359. <https://doi.org/10.1016/j.jneumeth.2019.108359>
- HUSTER, R. J., PLIS, S. M., & CALHOUN, V. D. (2015). Group-level component analyses of EEG: Validation and evaluation. *Frontiers in Neuroscience*, 9(JUL), 1–14. <https://doi.org/10.3389/fnins.2015.00254>
- JOHNSON, M. B., CACCIATORE, T. W., HAMILL, J., & EMMERIK, R. E. A. V. (2010). Multi-segmental torso coordination during the transition from sitting to standing. *Clinical Biomechanics*, 25(3), 199–205. <https://doi.org/10.1016/j.clinbiomech.2009.11.009>
- KIM, J., & PARK, H. (2012). Fast nonnegative tensor factorization with an active-set-like method. In M. W. Berry, K. A. Gallivan, E. Gallopoulos, A. Grama, B. Philippe, Y. Saad, & F. Saied (Eds.), *High-performance scientific computing: Algorithms and applications* (pp. 311–326). Springer London. https://doi.org/10.1007/978-1-4471-2437-5_16
- KIM, J., & PARK, H. (2014). *Nonnegative matrix and tensor factorization algorithms toolbox* [Software]. Available at: <https://github.com/kimjingu/nonnegfac-matlab>.
- KIRSCHNER, S., & TOMASELLO, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and Human Behavior*, 31(5), 354–364.
- KOLDA, T. G., & BADER, B. W. (2009). Tensor decompositions and applications. *SIAM Review*, 51(3), 455–500. <https://doi.org/10.1137/07070111X>
- LARTILLOT, O., & TOIVIAINEN, P. (2007). MIR in Matlab (II): A toolbox for musical feature extraction from audio. In S. Dixon, D. Bainbridge, & R. Typke (Eds.), *Proceedings of the 8th International Conference on Music Information Retrieval*, 127–130. Österreichische Computer Gesellschaft, Vienna, Austria.
- LEMAN, M. (2008). *Embodied music cognition and music mediation technology*. MIT Press.
- LEMAN, M., & NAVEDA, L. (2010). Basic gestures as spatiotemporal reference frames for repetitive dance/music patterns in Samba and Charleston. *Music Perception*, 28(1), 71–91. <https://doi.org/10.1525/MP.2010.28.1.71>
- LESAFFRE, M., DE VOOGDT, L., LEMAN, M., DE BAETS, B., DE MEYER, H., & MARTENS, J. P. (2008). How potential users of music search and retrieval systems describe the semantic quality of music. *Journal of the American Society for Information Science and Technology*, 59(5), 695–707. <https://doi.org/10.1002/asi.20731>
- LONDON, J. (2001). Metre. In A.-L. Santella (Ed.), *Grove music online*. Oxford University Press. <https://doi.org/10.1093/gmo/9781561592630.article.18519>

- LONDON, J., BURGER, B., THOMPSON, M., HILDRETH, M., WILSON, J., SCHALLY, N., & TOIVIAINEN, P. (2019). Motown, disco, and drumming: An exploration of the relationship between beat salience, melodic structure, and perceived tempo. *Music Perception*, 37(1), 26–41. <https://doi.org/10.1525/mp.2019.37.1.26>
- LUCK, G., SAARIKALLIO, S., BURGER, B., THOMPSON, M. R., & TOIVIAINEN, P. (2010). Effects of the Big Five and musical genre on music-induced movement. *Journal of Research in Personality*, 44(6), 714–720. <https://doi.org/10.1016/j.jrp.2010.10.001>
- MAHON, B. Z. (2015). What is embodied about cognition? *Language, Cognition and Neuroscience*, 30(4), 420–429. <https://doi.org/10.1080/23273798.2014.987791>
- MCAULEY, D. (2010). Tempo and rhythm. In M. Riess Jones, R. Fay, & A. Popper (Eds.), *Music perception. Springer handbook of auditory research* (Vol. 36), 165–199. Springer. https://doi.org/10.1007/978-1-4419-6114-3_6
- MCDUGALL, H. G., & MOORE, S. T. (2005). Marching to the beat of the same drummer: The spontaneous tempo of human locomotion. *Journal of Applied Physiology*, 99, 1164–1173.
- MERCHANT, H., & HONING, H. (2014). Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. *Frontiers in Neuroscience*, 7, 274. <https://doi.org/10.3389/fnins.2013.00274>
- MOBUS, G. E., & KALTON, M. E. (2015). *Principles of systems science*. Springer.
- NAVEDA, L., & LEMAN, M. (2010). The spatiotemporal representation of dance and music gestures using topological gesture analysis (Tga). *Music Perception*, 28(1), 93–111. <https://doi.org/10.1525/MP.2010.28.1.93>
- NETTL, B. (2001). An ethnomusicologist contemplates universals in musical sound and musical culture. In N. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 463–472). MIT Press.
- OGDEN, R., & HAWKINS, S. (2015). Entrainment as a basis for coordinated actions in speech. *ICPhS*, 599, 1–5.
- PAMPALK, E., RAUBER, A., & MERKL, D. (2002). *Content-based organization and visualization of Music Archives*. Retrieved from http://www.ofai.at/elias.pampalk/publications/pam_mm02.pdf
- PATEL, A. D., IVERSEN, J. R., BREGMAN, M. R., & SCHULZ, I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Current Biology*, 1169(10), 459–469. <https://doi.org/10.1111/j.1749-6632.2009.04581.x>
- PHILLIPS-SILVER, A. J., AKTIPIS, C. A., & BRYANT, G. A. (2010). Entrainment: Foundations of rhythmic movement. *Music Perception*, 28(1), 3–14. <https://doi.org/10.1525/mp.2010.28.1.3>
- PHILLIPS-SILVER, J., & TRAINOR, L. J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition*, 105(3), 533–546.
- PIRKL, R. J., REMLEY, K. A., & PATANE, C. S. L. (2012). Reverberation chamber measurement correlation. *IEEE Transactions on Electromagnetic Compatibility*, 54(3), 533–545. <https://doi.org/10.1109/TEM.2011.2166964>
- SATO, N., NUNOME, H., & IKEGAMI, Y. (2015). Kinematic analysis of basic rhythmic movements of hip-hop dance: Motion characteristics common to expert dancers. *Journal of Applied Biomechanics*, 31(1), 1–7. <https://doi.org/10.1123/jab.2014-0027>
- SCHAAL, S., STERNAD, D., OSU, R., & KAWATO, M. (2004). Rhythmic arm movement is not discrete. *Nature Neuroscience*, 7(10), 1136–1143. <https://doi.org/10.1038/nn1322>
- SMITH, L., & HONING, H. (2008). Time-frequency representation of musical rhythm by continuous wavelets. *Journal of Mathematics and Music*, 2, 81–97.
- SMITS-ENGELSMAN, B., VAN GALEN, G., & DUYSSENS, J. (2002). The breakdown of Fitts' law in rapid, reciprocal aiming movements. *Experimental Brain Research*, 145(2), 222–230. <https://doi.org/10.1007/s00221-002-1115-8>
- 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>
- TOMASELLO, M. (2020). The adaptive origins of uniquely human sociality. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 375(1803), 20190493. <https://doi.org/10.1098/rstb.2019.0493>
- VELARDE, G., MEREDITH, D., & WEYDE, T. (2016). A wavelet-based approach to pattern discovery in melodies. In D. Meredith (Ed.), *Computational music analysis* (pp. 303–333). Springer International Publishing. https://doi.org/10.1007/978-3-319-25931-4_12
- WANG, X., LIU, W., TOIVIAINEN, P., RISTANIEMI, T., & CONG, F. (2020). Group analysis of ongoing EEG data based on fast double-coupled nonnegative tensor decomposition. *Journal of Neuroscience Methods*, 330, 108502. <https://doi.org/10.1016/j.jneumeth.2019.108502>
- WANG, D., ZHU, Y., RISTANIEMI, T., & CONG, F. (2018). Extracting multi-mode ERP features using fifth-order non-negative tensor decomposition. *Journal of Neuroscience Methods*, 308, 240–247. <https://doi.org/https://doi.org/10.1016/j.jneumeth.2018.07.020>
- WONG, M. K. Y., & SO, W. C. (2018). Absence of delay in spontaneous use of gestures in spoken narratives among children with Autism Spectrum Disorders. *Research in Developmental Disabilities*, 72(January 2016), 128–139. <https://doi.org/10.1016/j.ridd.2017.11.004>
- ZAHAVI, D. (2001). Beyond empathy. Phenomenological approaches to intersubjectivity. *Journal of Consciousness Studies*, 8(5–7), 5–7.

Appendix A

Musical Stimuli Used in the Study

| Genre | Artist | Track | Tempo (bpm) |
|-------------------|---------------------------------|-----------------------------------|-------------|
| Blues | The Paul Butterfield Blues Band | Mystery Train | 126 |
| Blues | Keb' Mo' | She Just Wants to Dance | 113 |
| Country | Dixie Chicks | Goodbye Earl | 123 |
| Country | Brooks & Dunn | My Maria | 124 |
| Dance/Electronica | M People | Sight For Sore Eyes (Dance Remix) | 122 |
| Dance/Electronica | Lady GaGa | LoveGame (The Gaga Bender Mix) | 128 |
| Jazz | Jimmie Lunceford | Lunceford Special | 120 |
| Jazz | Sidney Bechet | Muskrat Ramble | 96 |
| Metal | Lamb of God | Redneck | 131 |
| Metal | White Zombie | Thunder Kiss | 113 |
| Pop | Christina Aguilera | Come On Over | 118 |
| Pop | Duran Duran | Want You More! | 132 |
| Rap/Hip-Hop | Run-DMC, Jason Nevins | It's Like That | 130 |
| Rap/Hip-Hop | DJ Laz | Move Shake Drop (Remix) | 127 |
| Reggae | Sean Paul | Temperature | 126 |
| Reggae | Shaggy | Oh Carolina | 126 |

Appendix B

Example of Time- and Time-Frequency-Domain Decompositions of Hand Movement Data

The following example aims to help provide insight into the differences of time- and time-frequency-domain representations in terms of their decomposability into latent variables. Figure 12 shows an example of a rather typical movement pattern during spontaneous dance, in which the dancer moves her hands circularly. We applied PCA to the time-domain representation as depicted in Figure 12A, as well as to a time-frequency-domain representation obtained by concatenating the absolute values of the wavelet transforms obtained for the movement data of each hand and direction.

Both representations thus had an equal number of variables (six). Figure 12C displays the proportion of variance contained in each principal component for the two representations. As can be seen, the first eigenvalue of the time-frequency-domain data is remarkably higher than that of the time-domain data. This implies that the former manages to accommodate a higher proportion of variance into the first component. This can be explained by the fact that the

mediolateral and vertical movement directions, due to their circular movement pattern, have a relative phase shift and thus do not covary. Figure 12D shows the loadings for the first three principal components. As is evident, for the time-frequency-domain representation all movement components load significantly to the first component, which for the time-domain representation they are more distributed between the components. Because PCA groups variables based on their mutual covariance, with the time-domain representation of the present data it fails to group the mediolateral and vertical hand movements although they clearly belong to the same gesture. The time-frequency-domain representation, on the other hand, ignores the phase shift and groups them into the same component based on their frequency-locking.

For the present example, the effective dimensionalities of the time-domain and time-frequency-domain representations are 5.00 and 2.95, respectively, suggesting a more compact decomposition of the movement data with the time-frequency representation.

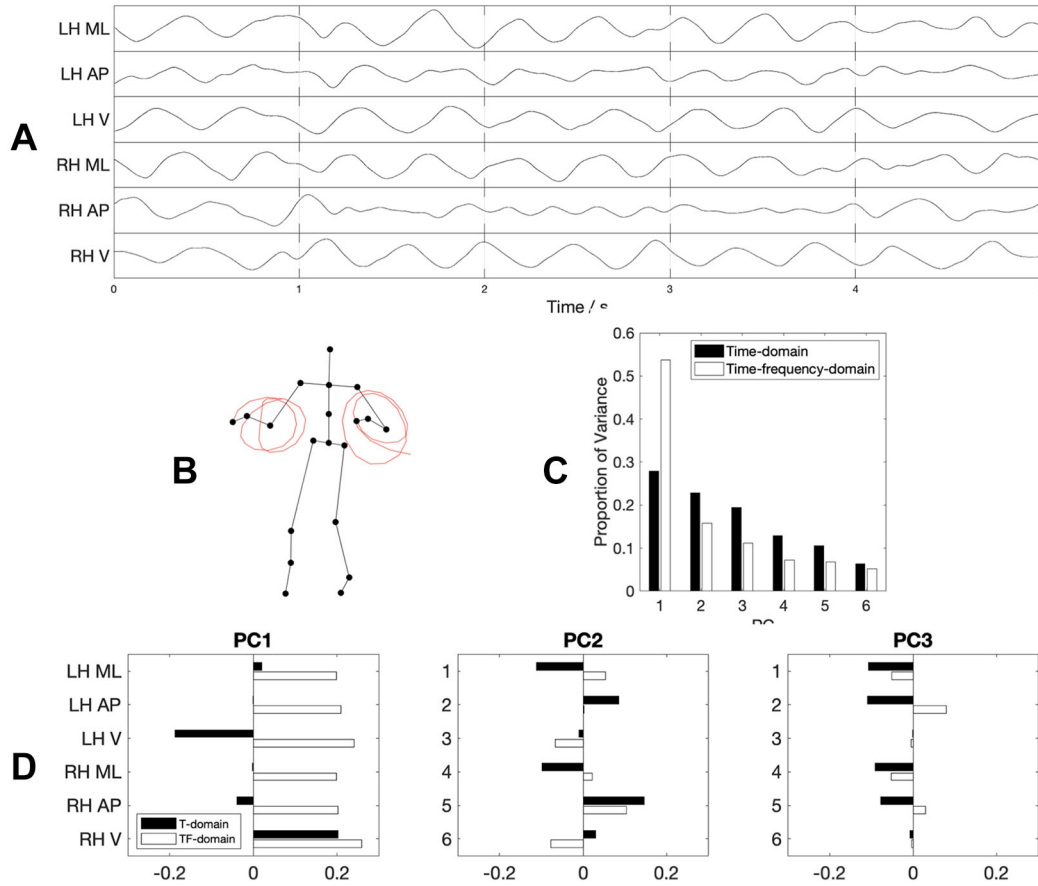


FIGURE 12. (A) Five seconds of velocity data of hands of a dancing participant (LH: left hand; RH: right hand; ML: mediolateral; AP: anteroposterior; V: vertical); (B) visualization of hand movement between 2 and 3 seconds (red traces), showing a circular trajectory; (C) proportion of variance contained by each principal component; and (D) principal component loadings obtained by PCA from time-domain (black bars) and time-frequency-domain representations (white bars).

Appendix C

Reconstruction of Eigenmodes for Visualization

For the purpose of reconstructing eigenmode i of trial j , we define the frequency-by-space-by-time *scaling tensor* Λ_{ji} by

$$\Lambda_{ji} = \lambda_i t_{ji} \mathbf{f}_i \otimes \mathbf{s}_i \otimes \mathbf{e} \quad (3)$$

where λ_i denotes the scaling constant of mode i in the tensor decomposition, t_{ji} the salience of mode i in trial j , \mathbf{e} a vector of ones with a dimensionality equal to the number of time points in the reconstructed data, and \otimes outer product.

The velocity matrix for reconstructed mode i of trial j , \hat{V}_{ji} , is obtained by

$$\hat{V}_{ji} = W^{-1} [\Lambda_{ji} \circ W(\mathbf{V}_j)] \quad (4)$$

where \mathbf{V}_j denotes the velocity data matrix for trial j , W the wavelet transform, W^{-1} the inverse wavelet transform, and \circ the Hadamard product (element-wise product).

Finally, the position data for time point t of eigenmode i of trial j , $\hat{X}_{ji}(t)$, is obtained by temporal integration according to

$$\hat{X}_{ji}(t) = \mathbf{X}_{jt} + \sum_{\tau=1}^t \hat{V}_{ji}(\tau) t \quad (5)$$

where \mathbf{X}_{jt} is the column-wise mean of \mathbf{X}_j , that is, the mean posture across time for trial j , $\hat{V}_{ji}(\tau)$ is the row corresponding to time point τ of \hat{V}_{ji} , and t the sampling interval of motion capture data.

Appendix D

Median Saliency Values and Interquartile Ranges Per Genre for All 12 Eigenmodes

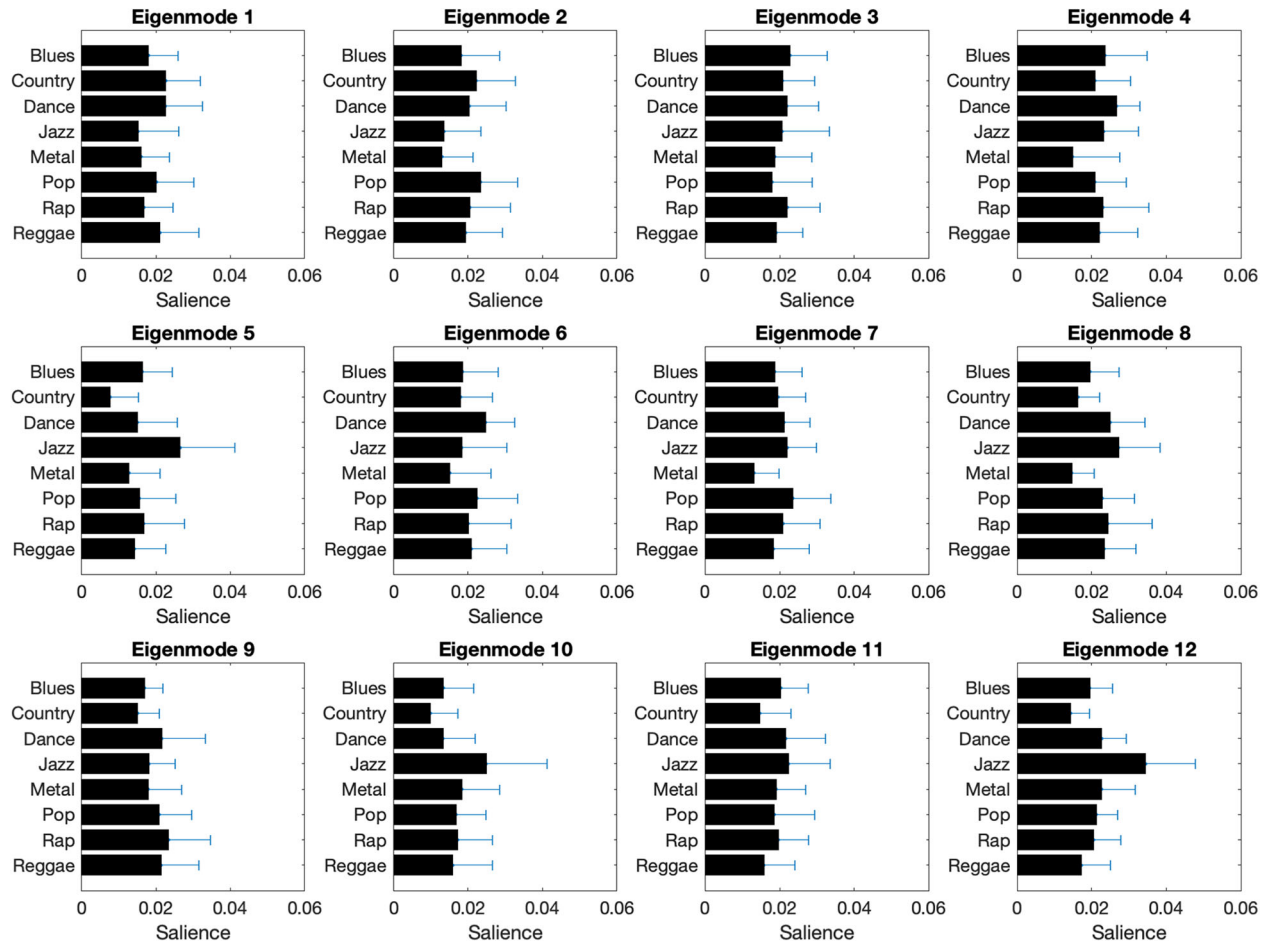


FIGURE 13. Median saliency values and interquartile ranges per genre for all 12 eigenmodes.