The effect of low-frequency equalisation on preference and sensorimotor synchronisation in music

Preprint · August 2021

Estefanía Cano
AudioSoureceRe

Steffen A. Herff
Western Sydney University

Original Article

QJEP

Quarterly Journal of Experimental Psychology
2022, Vol. 75(3) 475–490

DOI: 10.1177/17470218211037145
qjep.sagepub.com
SAGE

The effect of low-frequency equalisation on preference and sensorimotor synchronisation in music

Scott Beveridge[1], Estefanía Cano[1] and Steffen A Herff[2,3]

Abstract

Equalisation, a signal processing technique commonly used to shape the sound of music, is defined as the adjustment of the energy in specific frequency components of a signal. In this work, we investigate the effects of equalisation on preference and sensorimotor synchronisation in music. A total of 21 participants engaged in a goal-directed upper body movements in synchrony with stimuli equalised in three low-frequency sub-bands (0–50, 50–100, and 100–200 Hz). To quantify the effect of equalisation, music features including spectral flux, pulse clarity, and beat confidence were extracted from seven differently equalised versions of music tracks—one original and six manipulated versions for each music track. These music tracks were then used in a movement synchronisation task. Bayesian mixed-effects models revealed different synchronisation behaviours in response to the three sub-bands considered. Boosting energy in the 100–200 Hz sub-band reduced synchronisation performance irrespective of the sub-band energy of the original version. An energy boost in the 0–50 Hz band resulted in increased synchronisation performance only when the sub-band energy of the original version was high. An energy boost in the 50–100 Hz band increased synchronisation performance only when the sub-band energy of the original version was low. Boosting the energy in any of the three sub-bands increased preference regardless of the energy of the original version. Our results provide empirical support for the importance of low-frequency information for sensorimotor synchronisation and suggest that the effects of equalisation on preference and synchronisation are largely independent of one another.

Keywords

Music; movement; equalisation; sensorimotor synchronisation; spectral flux

Received: 24 September 2020; revised: 22 June 2021; accepted: 23 June 2021

Introduction

Music makes us move. This desire to move in response to music can be attributed to specific rhythmic structures and acoustic characteristics, collectively known as "groove" (Pressing, 2002). Prior research has shown that the experience of groove is consistent across individuals (Janata et al., 2012; Madison, 2006) and is largely independent of music style and genre (Madison, 2006). Furthermore, the feeling of groove is positively correlated with movement accuracy and perceived ease of motor entrainment (Janata et al., 2012; Madison et al., 2011). The music factors that represent "groove" are a subject of much research that broadly examines rhythmic (Madison et al., 2011; Pressing, 2002) and acoustic features (Stupacher et al., 2013, 2016). The present investigation predominantly focuses on the role of acoustic features in groove, specifically their effects on sensorimotor synchronisation performance.

In many forms of music, the instruments responsible for driving groove are bass (e.g., bass guitar) and percussion (e.g., drum kit) (Pressing, 2002). It is important to note that these instruments often perform in the lower

[1]Social and Cognitive Computing, Institute of High Performance Computing, Agency for Science, Technology and Research (A*STAR), Singapore
[2]École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
[3]Music Cognition and Action Research Group (MCA), MARCS Institute for Brain, Behaviour & Development, Western Sydney University (WSU), Sydney, NSW, Australia

Corresponding author:
Scott Beveridge, Social and Cognitive Computing, Institute of High Performance Computing, Agency for Science, Technology and Research (A*STAR), Singapore 138632, Singapore.
Email: scott.beveridge@ihpc.a-star.edu.ag

frequency registers (Butterfield, 2010; Pressing, 2002). The preference of using instruments in specific frequency registers to induce rhythmic movements clearly exemplifies the acoustic dimension of groove. Indeed, the link between low frequencies and movements has been evidenced in motor tasks within the context of contemporary dance music, where participants were shown to display a higher degree of tempo entrainment when the dynamic level of the bass drum was increased (Van Dyck et al., 2013). Additional evidence for the connection between low frequencies and movements comes from studies showing that during exercise, boosted bass frequencies can be used to increase arousal and encourage rhythmic movements (Karageorghis et al., 2012). Furthermore, intense low-frequency sounds evoke responses in the vestibular system which – besides generating pleasurable sensations (Todd & Cody, 2000) – plays an important role in musical rhythm perception (Trainor et al., 2009). In a study that artificially manipulated the frequency range of bass and bass drum, tapping variability was lower when the bass instruments contained frequency components in the low bass range (where the played bass line was one octave lower than the original, and the peak frequency of the bass drum was reduced from 140 to 40 Hz) (Stupacher et al., 2016). A "low-tone benefit" has also been observed in studies that require synchronising finger tapping with auditory cues, revealing that pitches in the lower registers improved synchronisation accuracy compared with pitches in higher registers (Hove et al., 2014). Taken together, these studies emphasise a connection between low-frequency components in music and motor activity.

Music listeners as well as professional audio engineers have the means to directly manipulate low-frequency components of music. This is often times achieved through equalisation. Equalisation (EQ) is a signal processing technique that allows a listener to adjust the energy of specific frequencies within the audible bandwidth (Huber & Runstein, 2005). In consumer audio playback devices, equalisation is commonly applied to specific frequency ranges known as sub-bands. Rhythm-carrying instruments, for example, typically inhabit the lowest sub-bands, with the low presence of the bass kick drum in the 50–100 Hz range and the electric bass guitar in the 41.2–343.2 Hz (E1 to F4) range (Hepworth-Sawyer & Hodgson, 2016; Huber & Runstein, 2005). Equalisation provides substantial control over frequency content, allowing listeners to boost or attenuate these sub-bands based on their requirements or preferences. However, excessive manipulation of frequency content can also be detrimental and can result in loss of clarity between instruments. Among professional music engineers and audiophiles alike, there exist some common "rules-of-thumb" concerning equalisation of low-frequency content. In the range of 16–60 Hz, which gives the bass a sense of power (Izhaki, 2018), excessive boosting can make the music sound *muddy*. This can lead to a loss of definition between instruments (Owsinski, 1999). In the 60–250 Hz frequency range, which encompasses the fundamental frequencies of the rhythm section (Izhaki, 2018) and gives the bass sound body and depth, excessive boosting may overdrive the reproduction loudspeaker. This can make music sound what is termed as *boomy* (Owsinski, 1999). It is important to note, that while the goal of equalisation is to change the perceptual sound qualities of a song, it can also implicitly produce changes in acoustic features (e.g., spectral flux, pulse clarity, etc.) relevant to aspects of groove and sensorimotor synchronisation.

The frequency content and spectral characteristics of a given song can be described with numerous acoustic features. Spectral flux (i.e., a measure of the change in the spectral content of a signal between different time instances), in particular, has proven to be a useful characterisation of groove and hence a good "anchor" for sensorimotor synchronisation tasks. In a number of behavioural studies, specific relationships between spectral flux measured in the low frequencies and music-induced movement have been investigated, most notably those conducted by Burger and colleagues (2013, 2014, 2018). In the earliest of these studies, where participants were asked to move to music, a positive correlation between head movement speed and spectral flux in the 50–100 Hz frequency sub-band was reported. In Burger et al. (2014) using the same movement instructions, participants demonstrated periodic body movements in the mediolateral (sideways) and anteroposterior (forward-back) planes that correlated with spectral flux in the same sub-band (50–100 Hz). These movements were found to synchronise to different metrical levels simultaneously and interchangeably, suggesting a complex relationship between periodic movement, low-frequency spectral flux, and meter (Burger et al. 2014). In a later study examining music-induced synchronisation ability, Burger et al. (2018) reported improved synchronisation ability to music with high spectral flux in the 100–200 Hz sub-band. More specifically, these results showed an improvement in synchronisation ability in vertical feet and hip movement at the beat level. These studies guide the present investigation in using spectral flux to characterise low-frequency content in music.

It is important to note that the above studies also suggest that synchronisation ability may be influenced by tempo (Burger et al., 2018). Generally, tempo synchronisation can be more easily achieved when music is played at a tempo that affords easy synchronisation (Madison et al., 2011). Previous research suggests synchronisation to music is most easily achieved at a moderate tempo corresponding to beat periods (i.e., inter-beat-intervals) between 450 and 700 ms (Fraisse, 1982; Parncutt, 1994; van Noorden & Moelants, 1999), with a preference for tempi around 120 BPM, with an inter-beat-interval around 500 ms (Moelants, 2002).

Pulse clarity is another acoustic feature that has been studied in the context of sensorimotor synchronisation. Pulse clarity attempts to characterise the high-level dimension that conveys how easily a listener can perceive the metrical pulsation in a given musical piece (Lartillot, Eerola, et al., 2008). Results of previous research show correlations between the amount of whole-body movements and pulse clarity (Burger et al., 2013) with highest correlation exhibited between superior-inferior movements at the beat level (Burger et al., 2014). Pulse clarity was also examined in an experiment involving a goal-directed (walking) movement (Prithvi et al., 2019), and it was observed that reduced pulse clarity resulted in poorer synchronisation ability. Following these findings, we also include two separate measures of pulse clarity in our study and investigate their effect on sensorimotor synchronisation.

Sensorimotor synchronisation performance is most commonly assessed using the finger tapping paradigm. In such a paradigm, performance is assessed by comparing the alignment between a specific movement event (e.g., a finger tap) and the rhythmic events in the reference signal (e.g., a metronome click or the beat of music). The time difference between the movement event and the reference rhythmic event is known as the synchronisation error or *asynchrony*. Sensorimotor synchronisation ability has been assessed in a variety of contexts and for a variety of movements including finger tapping in motor control research (Repp, 2005; Repp & Su, 2013), upper-limb movement in the analysis of skilled movements in musicians (Beveridge et al., 2020; Fujii et al., 2009), lower limb exercise performance (Buhmann et al., 2018), and lower limb assessment of neurological disorders (Stegemöller et al., 2009; Verheul & Geuze, 2004).

Given our study design, it is important to consider aesthetic preference as an established mediating variable. For example, the "loudness-war" refers to a specific mixing trend that maximises perceived loudness on the assumption that loudness correlates with preference (Vickers, 2010) (also see Hove et al., 2019). Indeed, in a task that provided participants with the direct opportunity to apply audio manipulations to musical songs, they preferred mixes with more energy in the bass frequencies (Dobrowohl et al., 2019), and listeners tend to prefer a higher loudness level proportional to their liking of the music (Cullari & Semanchick, 1989). Furthermore, prior research has shown a relationship between musical preference and performance in simultaneously executed tasks (North & Hargreaves, 1999), engagement with the environment (North & Hargreaves, 1996), and willingness to move (Witek et al., 2014). In addition, the music-induced desire to synchronise movement is often considered enjoyable. Indeed, prior studies have shown that higher aesthetic preference, predicts higher perceived ratings of groove (Senn et al., 2019). As a result, should an effect of low-frequency equalisation be observed, a possible explanation would be that a particular low-frequency manipulation may have influenced aesthetic preference, which in turn, increases task performance and willingness to move in the synchronisation task.

Specifically, this study investigates the relationship between sensorimotor synchronisation performance and low-frequency equalisation in music. We conducted a music movement experiment in which participants were asked to perform a goal-directed movement in time with manipulated versions of novel music tracks. Track manipulation involved boosting or attenuating three low-frequency sub-bands (0–50, 50–100, and 100–200 Hz). This was achieved by applying different types of filters to the stimuli that increase or decrease energy in these sub-bands. The manipulation procedure yielded seven versions of each track (one original, unmodified version, one attenuated at each sub-band, and one boosted at each sub-band). By applying these filters we are in effect causing spectral modifications to the stimuli. While these spectral modifications have a clear impact on musical features (like spectral flux, pulse clarity, and beat confidence), they do not modify tempo, style, or instrumentation of the different versions of the stimuli. This described manipulation is analogous to the equalisation process. To control for potential confounding effects of tempo, this study only uses stimuli with a tempo of 120 BPM. In the synchronisation task, participants performed an arm abduction/adduction movement commonly used as part of motor rehabilitation regimes. Synchronisation performance was assessed by means of synchronisation error.

To account for the influence of aesthetic preference, we assess whether the relationship between preference and low-frequency equalisation is similar to the relationship between low-frequency equalisation and sensorimotor synchronisation. In a separate procedure, we measured self-report measures of preference in response to a different set of music tracks manipulated in the same manner as in the synchronisation task. A preference task immediately followed a movement synchronisation task. Each participant completed a total of seven movement synchronisation tasks, each one followed by a preference task. The goal of alternating movement and preference tasks was to reduce possible carry over effects in the participants when synchronising to stimuli with the same tempo (120 BPM in our case).

This study addresses two main research questions. First, we investigate how sensorimotor synchronisation ability is affected by low-frequency equalisation. Second, we study how the same low-frequency manipulations affect preference and whether this effect is similar to that observed in the sensorimotor synchronisation task. Following the notion that groove is "carried" by rhythmic instruments in the low-frequency register, and the reported correlation between higher groove and ease of motor

Table 1. List of tracks from the MUSDB18 dataset used as stimuli in the experiment.

Movement set	Preference set
Buitraker—Revo X	Al James—Schoolboy Fascination
Georgia Wonder—Siren	Arise—Run Run Run
Girls Under Glass—We Feel Alright	Forkupines—Semantics
James Elder & Mark M Thompson—The English Actor	Little Chicago's Finest—My Own
Phre The Eon—Everybodys Falling Apart	Mu—Too Bright
The Easton Ellises (Baumi)—SDRNR	Nerve 9—Pray For The Rain
Zeno—Signs	Speak Softly—Broken Man

entrainment, we hypothesised that increased low-frequency energy would result in improved sensorimotor synchronisation performance and preference, irrespective of band manipulation.

Materials and methods

Participants

A total of 21 participants took part in the study (9 female, 12 male, $M=31$ years, $SD=7.56$ years). Three participants were left-handed and 18 right-handed. Eight participants had received formal musical education. All participants gave their informed consent before the start of the study, and were free to withdraw from the study at any point. Ethical approval for this study was obtained by the Human Biomedical Research Council Institutional Review Board (IRB Ref 2018-002), Agency for Science, Technology, and Research (A*STAR), Singapore.

Stimuli

Seven songs were selected for the synchronisation task, and another seven songs were selected for the preference task. No songs were shared between the two tasks to avoid potential carry over or order effects, as music in particular induces resilient memory traces (Herff et al., 2018, 2019). For the same reason, song order was randomised in both tasks. For each song, there were seven differently modified versions.

All stimuli were selected from the MUSDB18 corpus (Rafii et al., 2017) (see Table 1 for the list of audio stimuli). MUSDB18 is a freely available dataset comprising 150 multi-track recordings of mostly rock and pop music, compiled for use in research in the music sound source separation community. Besides instrumental stems, MUSDB18 also provides a mix for each track in the corpus. All tracks are non-commercial recordings, and so are completely unfamiliar to the participants. Selected segments were 15 s long, were in a time signature of 4/4, and had similar instrumentation, including drums, bass, accompanying instruments and a lead singer (female or male). All segments were extracted from the mixed track included in the corpus. The set of tracks used for the preference task have tempi in the 87–151 BPM range, and were used without any tempo modifications. The set of tracks used in the synchronisation task originally had tempi in the 113–127 BPM range. However, the tempo of all the tracks in the synchronisation set was modified to 120 BPM using the default "Change Tempo" effect in Audacity[1] (version 2.3.0, *high quality stretching* option enabled). This tempo modification ensured synchronisation tracks were at the preferred tempo for sensorimotor synchronisation (Moelants, 2002). Tracks in the preference task were left unmodified to reduce carry over effects from repeated synchronisation tasks with tracks at 120 BPM. As the preference task was web-based (see the "Music preference task" section), this alternating task arrangement had the additional benefit of providing respite from the physical activity of the movement synchronisation task. Tracks presented in the synchronisation task were preceded with an extra audio segment that represented a four beat count in. This count was provided by a tone similar to that of a metronome. The tracks in the preference set were kept in their original stereo format, while the tracks in the movement set were down-mixed from stereo to mono (single-channel) to remove possible effects of the panning in the mix on sensorimotor synchronisation. The process of down-mixing to mono essentially removes any spatial/location cues from the different musical instruments in the mix. All the tracks have a sampling frequency of 44,100 Hz.

To modify the spectral characteristics of the stimuli, and assess its impact on sensorimotor synchronisation, a set of filters were designed with the goal of attenuating or boosting the energy in certain frequency bands. While there are potentially many ways of modifying the music tracks, we argue that using filters commonly used in audio processing (e.g., in the equalisation process) allows us to seamlessly modify the music tracks without creating audible distortions or changing the auditory experience of the participants beyond a simple equalisation. To make the modifications of the music files as transparent and congruent as possible with the feature extraction process, the cut-off frequencies of the filters were chosen such that they coincide with those of the sub-bands used in the spectral flux calculation (see the "Music feature extraction" section). In particular, the boundaries of the three lowest

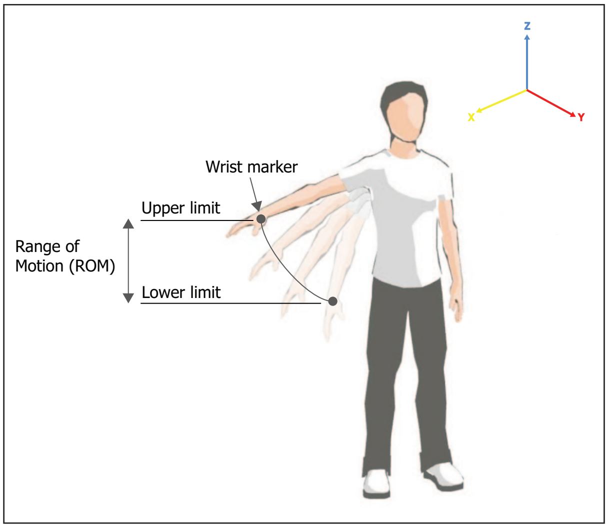

Figure 1. Experimental task showing wrist marker position and range of motion (ROM) bounded by upper and lower arm limits.

spectral bands in the spectral flux calculation, namely, 50, 100, and 200 Hz, were used for the filter design.

To achieve the attenuation of the spectral bands, three high-pass filters were implemented with 50, 100, and 200 Hz as cut-off frequencies, respectively. The three filters were second-order elliptical filters with 3 dB passband ripple, and 40 dB stopband attenuation. Similarly, to boost the energy of the spectral bands, three second-order bandpass filters with 50 Hz bandwidth, 3 dB gain, and with 50, 100, and 200 Hz as their higher cut-off frequencies, respectively, were implemented. The choice of these filters was made to ensure that the filtered versions of the tracks were distinct from each other, while maintaining a natural sound without distortions, saturation, or unpleasant audible effects. After the application of the filters and before feature calculation, peak normalisation was applied to all the audio files. This was done to guarantee that the feature calculations are not influenced by different energy levels in the songs. The same filtering and normalisation procedure was applied to all the tracks in the movement and preference sets. In total, seven versions of each track were used in the study: (1) Gain 50 Hz, (2) Gain 100 Hz, (3) Gain 200 Hz, (4) Original—No filters, (5) Attenuation 50 Hz, (6) Attenuation 100 Hz, and (7) Attenuation 200 Hz. This amounts to a total of 49 stimuli (7 songs × 7 versions) for each set.

Apparatus

Participants' movements were recorded with a Microsoft Kinect™ V2 camera (Microsoft, USA) controlled by custom software written in Unity (Unity Technologies, USA) and using the Kinect SDK Unity plugin.[2] Skeleton joint position data in three axes (X = mediolateral, Y = anteroposterior, Z = vertical) were captured at a frame rate of 50 Hz. Joint position data are estimated with the Kinect, which uses both infrared (IR) and RGB cameras for anatomical landmark identification (Menna et al., 2011). Music was delivered by headphone (KNS-6400 Studio Monitor Headphones, KRK Systems, USA) and synchronised with the movement data in the custom software.

Procedure

Each participant completed a total of seven movement synchronisation tasks, each one followed by a music preference task.

Movement synchronisation task. In each synchronisation task, participants performed a directed movement to the beat of a music track. Participants were positioned 2 metres from the Kinect motion capture camera. The main movement involved an abduction/adduction of the right arm in the frontal (X–Z) plane (Figure 1). This arm abduction/adduction movement was chosen as it is a component in so-called "reaching tasks," a common assessment for motor impairment (Chen et al., 2016; Roby-Brami et al., 2003; Thaut et al., 2002). Arm abduction/adduction is also a fundamental component in a number of functional movements (e.g., dressing, hair combing). Upper and lower limits of motion for each participant were defined prior to the beginning of the main task. Participants were instructed to move their arm between upper and lower limits in synchrony with the music stimuli. All tracks in the movement set were modified to have a tempo of 120 BPM, which is equivalent to an inter-beat interval (IBI) of 500 ms, and in the preferred range for beat induction and synchronisation (Moelants, 2002). Participants were encouraged to begin attempting to synchronise immediately upon hearing the priming (count in) tones. No visual feedback was given to the participant during the procedure. This task can be tiring for the arm muscles, and as a result, the task was limited to seven trials in total per participant and interleaved with the preference task to avoid fatigue effects. The order in which the tracks were presented was fully randomised for each participant. Each participant listened to each track once, each subjected to a different filter (see the "Music feature extraction" section). Which filters is applied to which track is counterbalanced across participants, to ensure that we obtain equal numbers of observations of each track-filter pair across the sample. To avoid carry-over effects between the preference and the synchronisation task, different songs (from the same corpus) were utilised in the two tasks. This step also ensures that potential similarities between the effect of a given manipulation in the two tasks is indeed driven by the audio manipulation rather than a specific song being used in both tasks.

Music preference task. The music preference task was conducted as a web-based listening test where participants were asked to rank in order of preference the different filtered versions of the same original song. The user interface

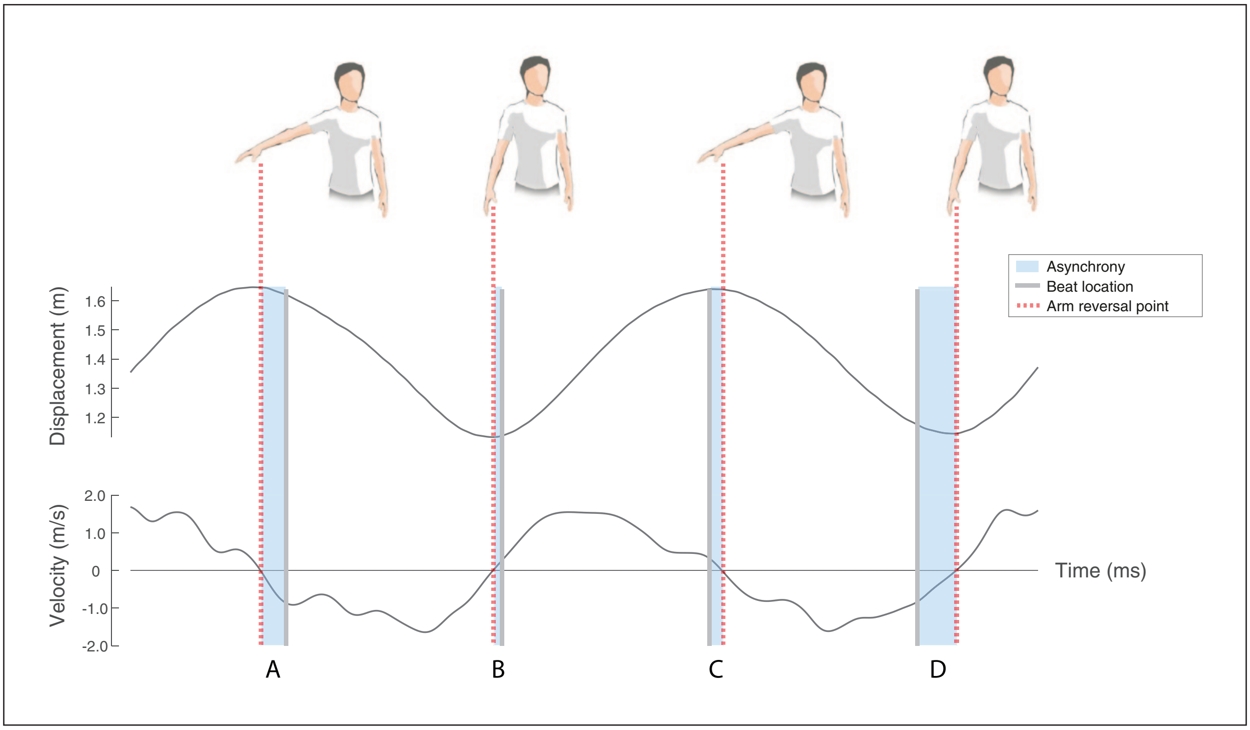

Figure 2. Schematic of the movement task showing asynchronies (synchronisation error) between beat locations and arm reversal points (denoted by points in the movement where velocity = 0). In A and B the reversal point precedes the beat location. In C and D the reversal point lags the beat location.

used for the preference task is shown in the Supplementary Materials section, which is a modification of the web-MUSHRA framework (Schoeffler et al., 2018). The seven versions of each song in the preference set (six modified versions plus the original song) were presented to the participant in randomised order. Participants were free to listen to the seven versions in any order, and as many times as necessary to be able to rank them according to preference (1 = *most preferred*, 7 = *least preferred*). The user interface only allows the participant to assign a ranking to a track once it has been played at least once. The order in which the seven preference tests were presented to each user was also randomised.

Data processing and synchronisation analysis

Using MATLAB (Mathworks, USA) we calculated the velocity of the wrist marker in the (vertical) *Z*-axis (see Figure 1). As each participant was instructed to synchronise their movement between upper and lower limits, we consider the arm reversal point as the synchronisation point. The reversal point is defined as the zero-crossing point of the velocity trajectory of the projected wrist marker (Figure 2). These points will be referred as *reversal points* in the remainder of this article.

To extract the reference beat positions, the *Madmom* beat tracker algorithm proposed in Böck et al. (2016) was applied to the original version of each song. The Madmom algorithm has reported an *F*-measure score on beat tracking performance of 0.9 on Western pop music, music similar to that used in this study. To further verify the accuracy of the beat positions extracted with Madmom, one professional musician manually annotated the beat positions for all the original tracks. A maximum time difference (Δ) between automatic and manual annotations of 50 ms was defined as the threshold for annotation accuracy. All the extracted beat positions had Δs below the threshold, and hence the annotations obtained with Madmom were used as reference beat positions for all tracks. Given that the modifications applied to each original songs only change the timbral characteristics of the tracks (and not the rhythmic ones), the beat positions extracted from the original songs are also valid for all versions of the tracks.

As a measure of synchronisation error (or asynchrony), we calculate the absolute value of the difference between each arm reversal point and its corresponding beat position in ms. We refer to this measure as *AbsAsynchrony*, and formally define it as:

$$AbsAsynchrony_i = \left| arm\ reversal\ point_i - beat\ position_i \right| \quad (1)$$

with i the index of the corresponding beat position and reversal point. It is important to note that *AbsAsynchrony* is calculated for each individual (abduction/adduction) movement as indicated by the i subscript in its definition. The absolute value in the calculations signifies that we do not discriminate between asynchronies that precede or succeed the beat position, that is, we do not discriminate between being early and being late. *AbsAsynchrony* is given in units of ms. The result is a different time series of *AbsAsynchrony* values for each participant and stimulus combination.

Music feature extraction

Spectral flux. Spectral flux is a measure of the frame-wise change in the spectral content of an audio signal. In this work, we focus our analysis on the three lowest bands of the spectral flux calculation as obtained by the sub-band ("Subband") decomposition of the MIRToolbox (version 1.7.2) (Lartillot, Toiviainen, & Eerola, 2008). The spectral flux can be used to reveal periodicities in a signal such as beats or bars, or to measure the stability of the signal over time. In this study, the spectral content in the spectral flux calculation is taken from the short time Fourier spectrum calculated on time windows of a given length. The distance (or change) between two spectral frames is obtained by calculating the Euclidean distance between the frames.

The spectral flux of the three lower spectral bands (subbands) is defined as follows for the remainder of this work:

- SFBand1: spectral flux calculated in the 0–50 Hz frequency range.
- SFBand2: spectral flux calculated in the 50–100 Hz frequency range.
- SFBand3: spectral flux calculated in the 100–200 Hz frequency range.

It is important to note that our experimental design does not guarantee a constant increase/decrease in terms of spectral flux for all of the tracks when applying a given filter. Spectral flux represents a measurement of how much the spectral content in a certain spectral band changes in time. However, the amount of energy in each of the spectral bands in consideration highly depends on the original song: while a given song might have heavy percussion and a bright bass sound evident in the spectral flux of the lower bands, some other song might have lighter percussion and less prominent bass sounds. When applying the filters to these two distinct songs, the results will naturally be very different. The filter implementation, however, does bring a guarantee for the stimuli in our dataset: the spectral flux for each band will be the lowest for the attenuated version, highest for the boosted version, and in between the filtered versions for the original track. This effect can be seen in Figure 3, where the spectral flux for the original, attenuated and boosted version for the first three spectral bands is shown for each song in the Movement Set (see Table 1).

Pulse clarity. Pulse clarity refers to the high-level musical dimension that conveys how easily listeners can perceive the metrical pulsation in a given musical piece (Lartillot, Eerola, et al., 2008). In practice, the calculation of pulse clarity is based on an *onset detection function* (ODF). The ODF indicates the main events in the music piece that may contribute to the evocation of pulsation. Pulse clarity is then defined in terms of the autocorrelation of the ODF. In this work, we use an onset detection function based on *spectral flux*, and the global maximum of the autocorrelation curve ("MaxAutoCor") to calculate pulse clarity using the MIRToolbox (version 1.7.2) (Lartillot, Toiviainen, & Eerola, 2008).

Beat tracking confidence measure. As an additional measure of beat clarity we include the beat tracking confidence measure proposed in (Zapata et al., 2012). The idea behind the confidence measure is to provide a numerical value that indicates how reliably beat information can be algorithmically extracted for a given song. The measure is defined as the mean mutual agreement (MMA) between a committee of beat trackers. To calculate the confidence measures, the implementation available in Essentia[3] library for music analysis was used.

Statistical approach. The present design considers beat-wise arm movement data as well as song-wise preferences rankings. However, within a song and within a participant, these data are non-independent. Both aggregating and under-utilising the data provided, as well as analysing the data without accounting for the underlying structural dependencies, can lead to flawed conclusions (Meteyard & Davies, 2020). To make use of the full data structure, while accounting for hierarchical dependencies, we use Bayesian mixed-effects models (Snijders & Bosker, 2011). These models can account for cross-random effects between participants and stimuli (Baayen et al., 2008), while also controlling for fatigue effects that correlate with an increasing trial number. All models are implemented in R using the brms package (Bürkner, 2017).

For each analysis, we describe the architecture of the model and report the results of models with 1,000 warm-ups, 10,000 iterations, on 4 chains, provided with a weakly informative prior $t(3,0,1)$ (Gelman et al., 2008). We report coefficient estimates (β), and estimated error in the coefficients (EE). Evidence ratios (Odds) for the individual hypotheses are reported, and for convenience, we indicate effects that can be considered significant at an $\alpha = .05$ level with *. This corresponds to an odds ratio $\geqslant 19$ (Milne & Herff, 2020).

Results

Synchronisation

We refer to spectral flux calculated in the 0–50 Hz frequency range as SFBand1, to spectral flux calculated in

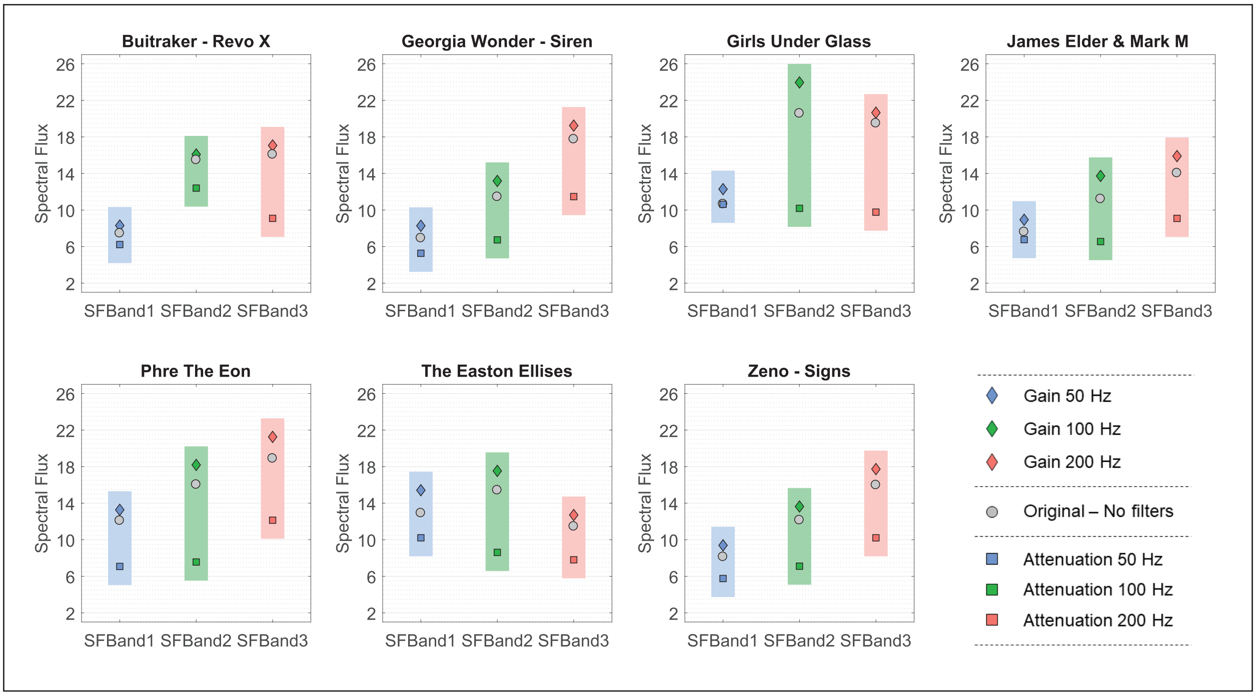

Figure 3. Spectral flux for all the tracks in the synchronisation study.
It can be observed that for each spectral band (*SFBand1*, *SFBand2*, *SFBand3*), the spectral flux of the tracks with the attenuation filter is the lowest, the spectral flux of the tracks with gain filters if the highest, and the spectral flux of the original version (without filtering) is in between the attenuation and gain versions.

the 50–100 Hz frequency range as SFBand2, and to spectral flux calculated in the 100–200 Hz frequency range as SFBand3. We denote the original song the 0 condition, and define their sub-band spectral flux as: *SF0Band1*, *SF0Band2*, and *SF0Band3*. For example this means that in a trial in which a participant is listening to the original version of a stimulus, *SFBand* and *SF0Band* would be identical for all three frequency bands (*SFBand1*=*SF0Band1, SFBand2*=*SF0Band2*, etc.). However, in a trial where participants are listening to a version in which *SFBand1* is boosted, then in this trial, *SFBand1* would be larger than *SF0Band1*. This allows us to model the effect of boosting and attenuating frequency bands, as well as their possible interactions with the spectral flux of the original version. Considering the interaction between the current song version, as well as the original song version is important. For example, it could be that boosting *SFBand1* of a song that has high *SF0Band1* to begin with may affect both preference and synchronisation behaviour differently, compared with boosting *SFBand1* of a song that has low *SF0Band1* to begin with. This is because low or high *SF0Band1* could be indicative of a deliberate choice of the audio engineer that mixed the original version of a song. This choice may have been made to highlight or suppress a specific frequency band that contains a particular instrument or an important part of the song.

All spectral flux values were normalised by subtracting the mean spectral flux value of the respective band and dividing by its standard deviation. A Bayesian mixed-effects model was used to predict *AbsAsynchrony*. In the context of this work, lower *AbsAsynchrony* indicates better synchronisation performance. To predict *AbsAsynchrony*, the model was provided with *SF0Band1*, *SF0Band2*, *SF0Band3*, *SFBand1*, *SFBand2*, *SFBand3*, as well as the interaction terms *SFBand1* × *SF0Band1, SFBand2* × *SF0Band2*, and *SFBand3* × *SF0Band3*. To account for possible sources of noise in the experimental design, the model was provided with a random intercept for *TrialNumber, ArmReversalPointNumber* within a trial, *Participant* and *Song*.

Effect of the original spectral flux on synchronisation performance. Results do not show strong evidence that *SF0Band1*, $\beta_{SF0Band1}$=–.07, $EE_{SF0Band1}$=.38, *Odds* ($\beta_{SF0Band1}$ < 0)=1.46, or *SF0band2*, $\beta_{SF0Band2}$=.26, $EE_{SF0Band2}$=.40, $Odds(\beta_{SF0Band2} > 0)$=3.21, predict *AbsAsynchrony*. However, we obtained strong evidence that greater values in *SF0Band3*, $\beta_{SF0Band3}$=–.76, $EE_{SF0Band3}$=.36, $Odds(\beta_{SF0Band3} < 0)$=41.86*, predict lower *AbsAsynchrony*. This suggests that in general, original track versions are easier to synchronise to, the higher the *SF0Band3*, whereas *SF0Band1* and *SF0Band2* do not show an effect. This can be seen in the marginal effects plots in Figure 4, where flat lines are observed for *SF0Band1* and *SF0Band2*, and a steep downwards slope is observed for *SF0Band3* in the original stimuli.

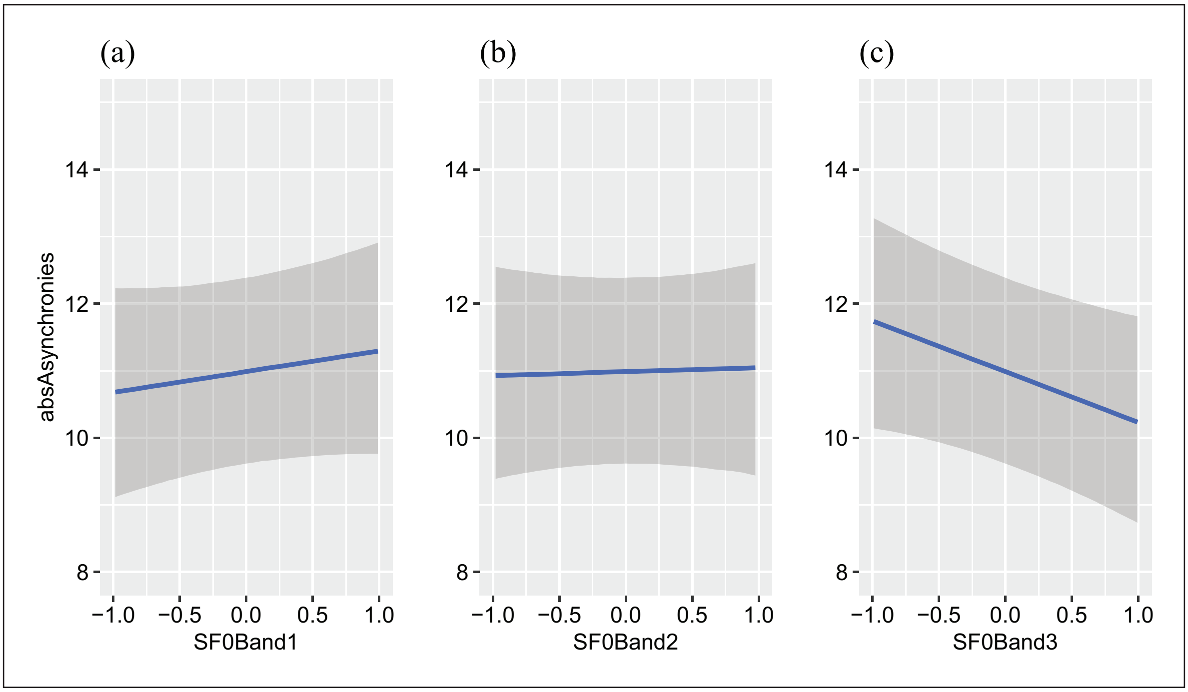

Figure 4. Marginal effects plots for the change in predicted *AbsAsynchrony* based on the (a) *SF0Band1*, (b) *SF0Band2*, and (c) *SF0Band3*. All SF values are given in SDs. Grey areas indicate 95% CIs. We only see strong evidence for an effect in *SF0Band3*. Specifically, the higher *SF0Band3* of an original song, the better synchronisation performance. This is indicated by the downwards slope in the Figure 4c.

Effect of modified spectral flux on synchronisation performance. We obtained no compelling evidence that higher *SFBand1*, $\beta_{SFBand1}=-.02$, $EE_{SFBand1}=.20$, $Odds(\beta_{SFBand1} < 0)=.86$, or *SFBand2*, $\beta_{SFBand2}=-.07$, $EE_{SFBand2}=.20$, $Odds(\beta_{SFBand2} < 0)=1.8$, predicts changes in *AbsAsynchrony*. However, we obtained strong evidence that higher *SFBand3* predicts higher *AbsAsynchrony*, $\beta_{SFBand3}=.28$, $EE_{SFBand3}=.12$, $Odds(\beta_{SFBand3} > 0)=141.29$*. This can be seen in Figure 5 with the flat lines for *SFBand1* and *SFBand2*, and the steep upwards slope for *SFBand3*. Higher spectral flux in *SFBand3* is indicative of lower synchronisation performance in the modified stimuli.

Interaction of original spectral flux and modified spectral flux on synchronisation performance. We observe strong evidence that *SF0Band1* × *SFBand1* predicts *AbsAsychrony*, $\beta_{SF0Band1*SFBand1}=-.41$, $EE_{SF0Band1*SFBand1}=.08$, $Odds(\beta_{SF0Band1*SFBand1} < 0)=>9999$*, as does *SF0Band2* × *SFBand2*, $\beta_{SF0Band2*SFBand2}=.30$, $EE_{SF0Band2*SFBand2}=.08$, $Odds(\beta_{SF0Band2*SFBand2} > 0)=>9999$*. However, we only obtain weak evidence for an *SF0Band3* × *SFBand3*, $\beta_{SF0Band3*SFBand3}=-.17$, $EE_{SF0Band3*SFBand3}=.12$, $Odds(\beta_{SF0Band3*SFBand3} < 0)=10.82$, interaction. As shown in Figure 6, the interaction between *SF* and *SF0* is very different for each of the three bands. If an original song has low spectral flux in band 1 (*SF0Band1*), further decreasing it increases synchronisation performance. This is indicated by upwards slope of the blue line in Figure 6a. Similarly, if an original song has high spectral flux in band 1 (*SF0Band1*), further increasing it also increases synchronisation performance. This is indicated by the downwards slope of the red line in Figure 6a. The interaction in band 2 follows the opposite pattern. If an original track has low spectral flux in band 2 (*SF0Band2*), increasing it improves synchronisation performance. This is indicated by the blue line in Figure 6b. In contrast, if an original track has high spectral flux in band 2 (*SF0Band2*), increasing it decreases synchronisation performance (red line, Figure 6b). We do not observe compelling evidence for an interaction in band 3. In this case, regardless of the amount of spectral flux in band 3 of the original tracks (*SF0Band3*), increasing the spectral flux always decreases performance. This is indicated by the upwards slopes in all lines in Figure 6c. However, tracks that have high spectral flux in band 3 (*SF0Band3*) to begin with, show better synchronisation performance. This is indicated by the red line in Figure 6c consistently showing the lowest *absAsychronies*.

In *beat clarity and beat tracking confidence*, we considered two computational methods for extracting periodicity and beat confidence of a given stimulus. We deployed a separate Bayesian mixed-effects models predicting *absAsynchronies* based on *pulseClarity* and

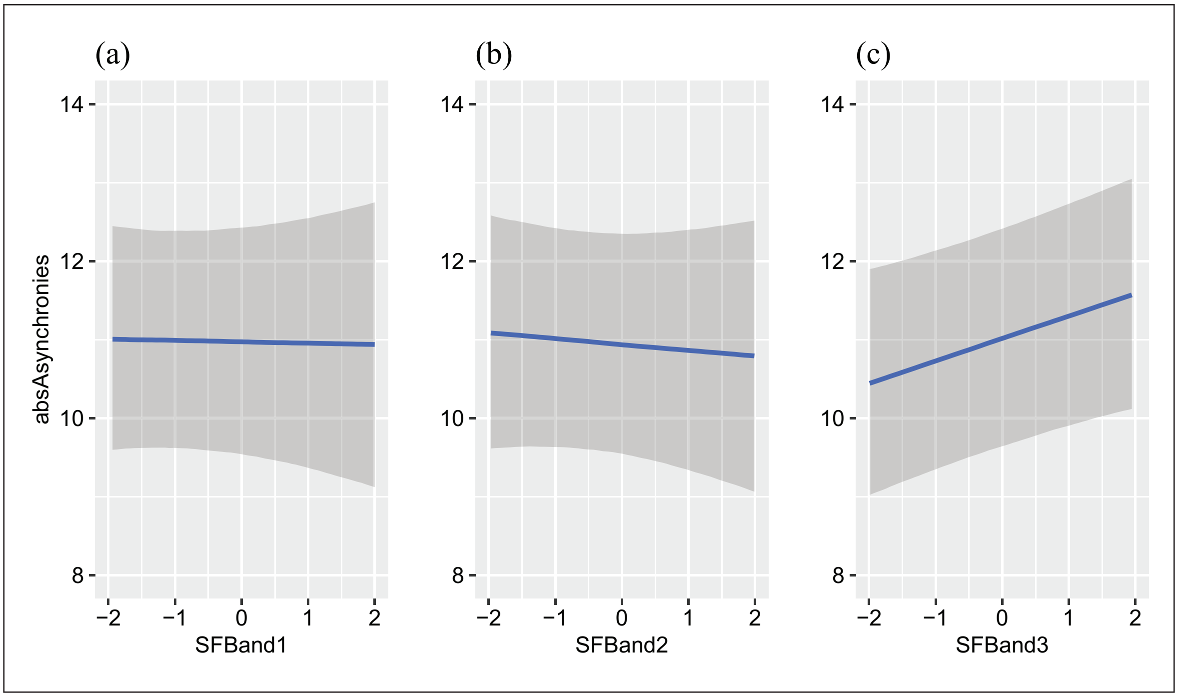

Figure 5. Marginal effects plots for the change in predicted *AbsAsynchrony* based on the (a) *SFBand1*, (b) *SFBand2*, and (c) *SFBand3* of a song version.
All spectral flux values are given in SDs. Grey areas indicate 95% CIs. We only see strong evidence for an effect in *SFBand3*. Specifically, the higher *SFBand3*, the worse synchronisation performance, as indicated by higher *AbsAsynchrony*.

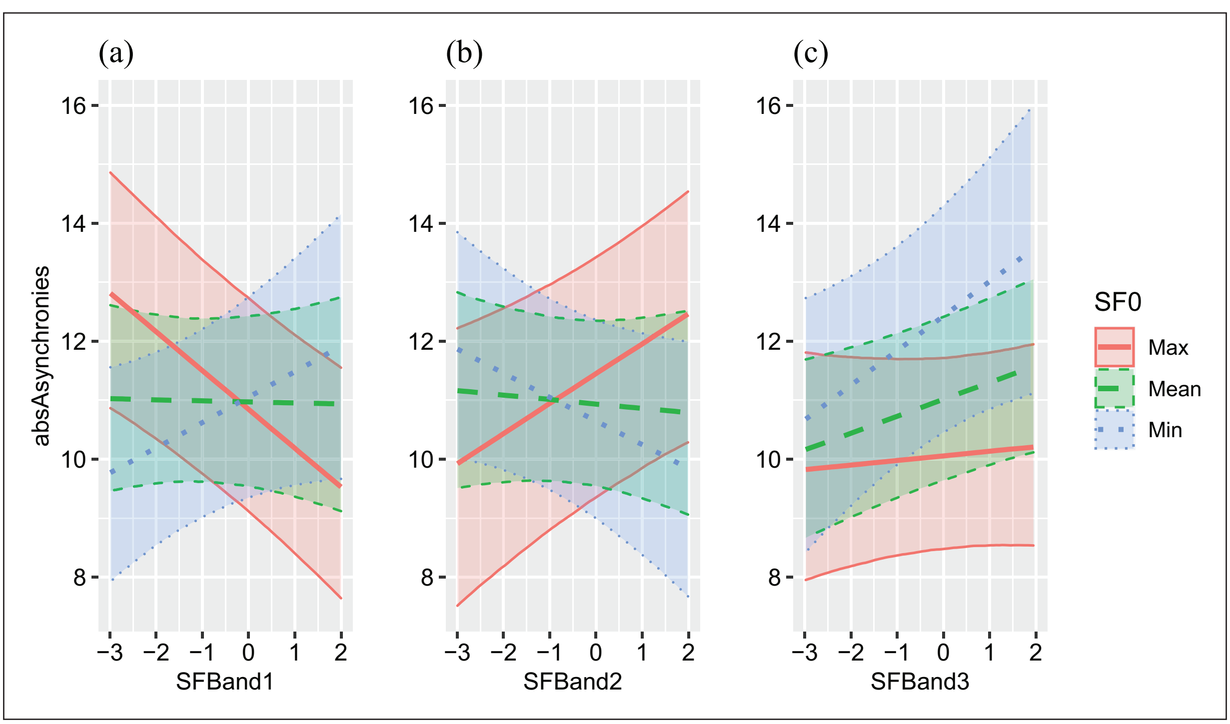

Figure 6. Marginal effects plots for the change in predicted *AbsAsynchrony* based on the interaction between *SFBand1* (a), *SFBand2* (b), and *SFBand3* (c) of a song version, and *SF0* of the respective band in the original version.
Lower *absAsynchronies* indicate better synchronisation performance. The interaction between *SF0* and *SF* shows very different behaviour for each of the three bands. While increasing *SFBand3* always decreases performance regardless of *SF0Band3*, modifications of *SFBand1* and *SFBand2* can have both positive and negative effects on performance depending on *SF0* in the respective band. All *SF* values are given in SDs. The shaded areas between thin lines indicate 95% CIs around the bold effect lines.

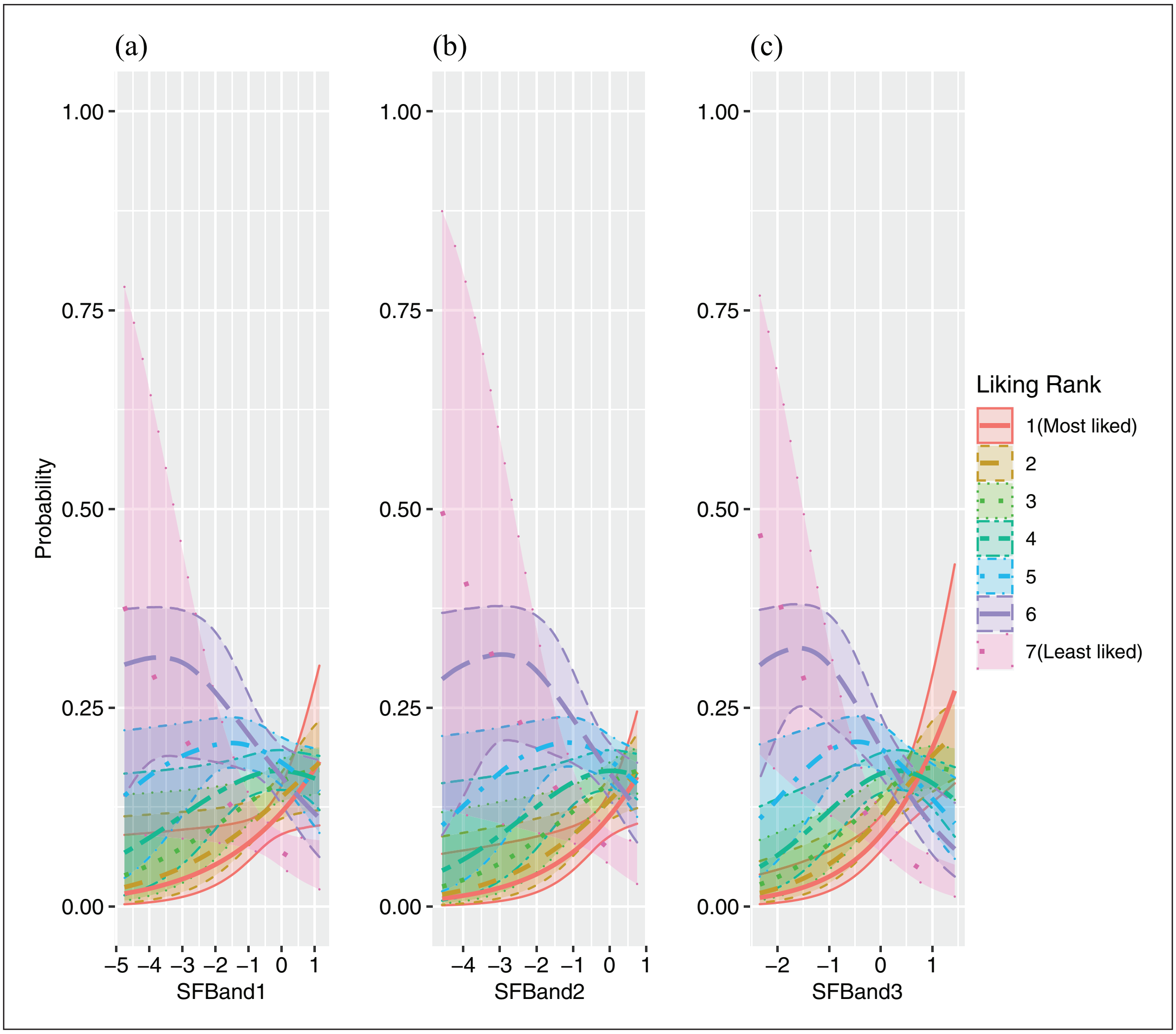

Figure 7. Liking rank probability based on relative spectral flux in bands 1, 2, and 3 in relation the spectral flux in the original songs (Liking rank 1 = *most liked*, 7 = *least liked*). Reducing (a) *SFBand1*, (b) *SFBand2*, or (c) *SFBand3* results in substantially less liking. The shaded areas between thin lines indicate 95% CIs around the bold effect lines.

beatConfidence. The model was provided with the same random effects structure as the models before. Contrary to the hypothesis, we observed little to no evidence that *pulseClarity*, $\beta_{pulseClarity} = -.04$, $EE_{pulseClarity} = .74$, $Odds(\beta_{pulseClarity} < 0) = 2.38$, or *beatConfidence*, $\beta_{beatConfidence} = .65$, $EE_{beatConfidence} = .05$, $Odds(\beta_{beatConfidence} > 0) = 11.08$, predict *absAsynchronies*.

Preference

A Bayesian mixed-effects model was used to predict the ordinal scaled liking ratings (1 = *most liked*, 7 = *least liked*) using *SF0Band1*, *SF0Band2*, *SF0Band3*, and *SFBand1*, *SFBand2*, *SFBand3* as predictors. As done in the synchronisation analysis (see Synchronisation section), interactions between the spectral flux of the original track (*SF0*), and the spectral flux of a given track version (*SF*) were defined for each band: *SFBand1* × *SF0Band1, SFBand2* × *SF0Band2, SFBand3* × *SF0Band3*. The model was provided with a random slope over *participant* for all predictors, as well as their interactions.

We obtained strong evidence that reducing *SFBand1* predicts less liking, $\beta_{SFBand1} = -.44$, $EE_{SFBand1} = .21$, $Odds(\beta_{SFBand1} < 0) = 56.23$*. We also obtained strong evidence that decreasing *SFBand2*, $\beta_{SFBand2} = -.56$, $EE_{SFBand2} = .23$, $Odds(\beta_{SFBand2} < 0) = 120.62$*, as well as decreasing *SFBand3*, $\beta_{SFBand3} = -.93$, $EE_{SFBand3} = .26$, $Odds(\beta_{SFBand3} < 0) = 3999$*, predict less liking. This can be seen in the three panels of Figure 7. Figures 7a (*SFBand1*), 7b (*SFBand2*), and 7c (*SFBand3*) show that the probability of high liking (e.g, rank 1 in red) increases with higher *SFBand1, SFBand2*, and *SFBand3* values, respectively, while the probability for lower liking (e.g., rank 7 in purple) increases when song versions are modified to have less spectral flux compared with the original version in any of the three bands. All other predictors and interactions yielded no to weak evidence (all evidence ratios <6.26).

Taken together this suggests that decreasing spectral flux in bands 1, 2, and 3 relative to the original versions, decreases the relative liking of the modified version.

Discussion

This study investigated the effect of low-frequency equalisation on sensorimotor synchronisation performance and preference.

Our first step was to analyse synchronisation performance while participants listened to the original versions of the stimuli. In this analysis, we found that higher levels of spectral flux in sub-band 3 (*SF0Band3* [100–200 Hz]) led to better synchronisation performance among our participants. In contrast, we found no effect for spectral flux, pulse clarity, or beat confidence in sub-band 1 (*SF0Band1* [0-50 Hz]) and sub-band 2 (*SF0Band2* [50-100 Hz]).

Our second step was to examine synchronisation performance in response to modified versions of the stimuli. Interestingly, we discovered that an increase in spectral flux in sub-band 3 (*SFBand3*), decreased synchronisation performance. We found no evidence of an effect between SMS performance and spectral flux in sub-band 1 (*SFBand1*), spectral flux in sub-band 2 (*SFBand2*), pulse clarity, or beat confidence. These results are interesting when contrasted with the analysis of the original version of the stimuli. While original track versions with higher SF in sub-band 3 show better synchronisation performance, results with the modified versions have shown that simply increasing the SF in sub-band 3 (increased by proxy of filtering in this study) will not necessarily result in better synchronisation performance. In fact, results show quite the opposite effect, with increased SF in sub-band 3 being detrimental to synchronisation performance.

To further elucidate the effects of sub-band manipulation on SMS performance, we examined the interaction effects between the original and modified version of the stimuli. For sub-band 1 we found that if an original track had low spectral flux then moving to a version with higher spectral flux decreased performance. Conversely, if an original track had a high level of spectral flux then moving to a version with higher spectral flux increased performance. In sub-band 2 we found the opposite. If an original track had a low level of spectral flux then moving to a version with higher spectral flux increased performance. If the original track had a high level of spectral flux then moving to a version with a higher level of spectral flux decreased performance. In sub-band 3 we found no interaction effect, and moving to versions with higher spectral flux always decreased performance. A possible explanation for these findings may come from the equalisation "rules of thumb" for mixing popular music of the type used in this study. Sub-band 1 is within the 16–60 Hz frequency region where sound is felt more than heard. It can provide a track with a subjective measure of "power" but excessive boosting of this band can lead to a "muddy" sound in the final mix. Here, "muddy" is defined as a lack of clarity or definition of the sounds/instruments within this band. Our findings suggest that tracks originally mixed with low sub-band 1 energy benefit from a further reduction in this band with respect to SMS performance. It is possible that no useful rhythmic acoustic information exists in these mixes (or in the original frequency content of the song), and boosting these frequencies results in "muddying" the mix or losing clarity. However, mixes produced with well defined spectral activity in sub-band 1 benefit from a boost in this band. It is possible they have been originally mixed with "good power" and a boost just helps synchronisation.

Sub-band 2 is within the 60–250 Hz range where instruments like the electric bass guitar have their fundamental frequencies. If frequencies within this range are boosted too much, the mix can sound subjectively "boomy." Boomy is a term linked to the capabilities of a loudspeaker transducer itself. Too much energy in sub-bands 2 and 3 can overdrive the speaker and wash out the frequencies in this band. Our findings suggest that tracks originally mixed with low energy in sub-band 2 benefit from gain in this sub-band with respect to SMS performance. Tracks orignally mixed with high energy in sub-band 2 lose synchronisation power from a boost. This suggests that tracks intentionally mixed with high energy in sub-band 2 might already be at the physical limits of frequency reproduction, and further increasing the energy in this band can "wash out" the spectral content useful for synchronisation due to boominess.

In sub-band 3 we found that no matter what level of sub-band energy the orignal tracks versions were produced with, moving to a version with boosted sub-band 3 (and hence, boosted SFBand3) is detrimental to synchronisation. It is very likely that mixing engineers instinctively know the optimal amount of energy required in sub-bands 2 and 3 to maximise the capabilities of the speaker systems before it becomes too much. If tracks are further equalised (for example, by the music consumer) to have more energy in sub-band 2 or sub-band 3 (and hence more spectral flux in these bands), they might be pushed over the limit, losing clarity and having a negative effect on SMS performance. From a mechanistic perspective, it could be that what we observe here, is manipulating the ease through which neural populations can entrain to the rhythmic properties within the music (Herff et al., 2020; Nozaradan et al., 2011). This explanation could be further tested in the future through neuroscientific studies that measure neural entrainment to beat frequency, as function of the audio manipulations explored in this work.

Overall our results point to the importance and skill of the mixing engineer in the production of the original music tracks. It seems that there is a very fine margin before further manipulations become detrimental for synchronisation performance: sub-band 3 was balanced in the mixes such that any increase from the original levels would cause lower synchronisation performance. Our results also highlight that while benefits for synchronisation can be obtained by means of equalisation (which in turn modifies sub-band spectral flux), there is no golden rule that can be applied to every song. The extent to which the synchronisation potential of a song can be improved by means of equalisation (or by increasing the sub-band spectral flux) will greatly depend on both the original musical content of the song, and on the way the song was originally mixed.

Results from the music preference task showed a different facet of the effects of low-frequency equalisation: any attenuation in the energy of any of the sub-bands (which in turn results in a decrease of the sub-band spectral flux) resulted in less liking. These results are interesting as they suggest that a given equalisation technique may not categorically show the same effect on synchronisation as it does on liking. Theoretically, this could lead to scenarios in which listeners synchronise better to songs that they do not enjoy the most. As a consequence, this means that mixing a song to increase liking, might require different steps compared with mixing a song to increase synchronisation.

Our initial hypothesis stated that increased low-frequency energy would result in improved sensorimotor synchronisation performance, irrespective of band manipulation. When examining the original unmodified tracks we found that higher spectral flux in sub-band 3 was indicative of higher SMS ability. These findings are consistent with (Burger et al., 2013) albeit in a different context (music-induced whole-body movement vs. directed abduction/adduction of the arm). In addition, when we began manipulating the original tracks by means of sub-band filtering, we found that additional sub-bands may also influence SMS performance. This points to a somewhat more complex relationship between low-frequency energy, spectral flux, and sensorimotor synchronisation performance. The full nature of which is still an open question, and given that only one sub-band was manipulated at any one time in this study, future research will focus on these interactions.

Although our original stimuli are drawn from the same corpus, it is possible that our results could by influenced by our track selection and manipulation procedures. To avoid carry over effects between the preference and synchronisation tasks, seven different tracks were selected for each task. Despite exhibiting similar distributions, it is possible that this selection process affected comparability between the tracks and influenced our results. Furthermore, tracks in the synchronisation task were artificially modified to set a uniform tempo of 120 BPM. Our motivation for this tempo manipulation was twofold: First, it ensured tracks were optimised for synchronisation (Moelants, 2002). Second, it allowed us to control for tempo in our experimental design. When controlling for tempo in addition to meter (all stimuli had 4/4 time signature), music genre and to a certain extent instrumentation, we still observed a difference in synchronisation ability. In a future study, we hope to further curate our corpus or modify our experimental design to ensure comparability. Furthermore, it is worth noting that the sensorimotor task only allows conclusions on a population level, whereas the liking task also allows interpretation on an individual level. This is because not all participants listened to all version-song combinations in the sensorimotor task, instead, the version song combinations were counterbalanced across the sample. This was necessary to address muscle fatigue within the present task.

Another important limitation to mention is the playback system. Headphone frequency responses are nearly never flat, and the specific model used here shows a relatively little energy below 50 Hz.[4] However, as manipulating SFBand1, which mainly considers energy below 50 Hz, affected behaviour in both tasks, we can conclude that the headphones used here were adequate to represent the experimental manipulations. Nevertheless, precise effect sizes in this study should be interpreted with care and not be used as the basis of equalisation practise. Future studies should attempt to replicate the present studies using different playback systems, to ensure that the present results are not specific to the headphone pair used. Indeed, a systematic investigation of the interaction between playback system and synchronisation as well as preferences responses would provide actionable insights for audio engineers.

Finally, another possible limitation of our study is the format for our movement synchronisation task. In this study, we measure the sensorimotor synchronisation performance based on an abduction/adduction of the right arm in the frontal plane. This choice of movement was motivated by a specific use case of tele-rehabilitation, in which arm abduction/adduction is commonly used to assess the range-of-motion (ROM) and is a constituent part of many more complex functional tasks (i.e., dressing, hair combing). Despite its contextual relevance in this study, this movement is not typical in the assessment of SMS and may have influenced the results. In a future study, we aim to replicate our results using a more common movement paradigm (e.g., finger tapping).

Conclusion

This study examined the relationship between low-frequency equalisation, preference, and sensorimotor synchronisation. Our findings indicate that the frequency components in the original versions of music were finely balanced. Balancing the low-frequency components that convey rhythm is much more nuanced than broad sub-band boost/attenuation, and there appears to be no "golden rule" for balancing/mixing that is applicable to all songs. This is even more interesting when considering our finding of independence between sensorimotor synchronisation performance and preference. An overall boost of low frequencies is likely to increase liking but not necessarily synchronisation. Further research should examine the effects of more fine-grained equalisation techniques and the interaction between instrumentation, equalisation, and synchronisation.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Scott Beveridge https://orcid.org/0000-0003-1576-3476
Steffen A Herff https://orcid.org/0000-0001-8396-1306

Notes

1. https://www.audacityteam.org/
2. https://assetstore.unity.com/packages/3d/characters/kinect-v2-examples-with-ms-sdk-and-nuitrack-sdk-18708
3. https://essentia.upf.edu/reference/std_BeatTrackerMultiFeature.html
4. https://github.com/jaakkopasanen/AutoEq/tree/master/results/headphonecom/headphonecom_harman_over-ear_2018/KRK%20SYSTEMS%20KNS%206400

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

Beveridge, S., Herff, S. A., Buck, B., Madden, G. B., & Jabusch, H. C. (2020). Expertise-related differences in wrist muscle co-contraction in drummers. *Frontiers in Psychology*, *11*, Article 1360. https://doi.org/10.3389/fpsyg.2020.01360

Böck, S., Korzeniowski, F., Schlüter, J., Krebs, F., & Widmer, G. (2016). Madmom: A new python audio and music signal processing library. In *Proceedings of the 24th ACM international conference on multimedia* (pp. 1174–1178). https://arxiv.org/abs/1605.07008

Buhmann, J., Moens, B., Van Dyck, E., Dotov, D., & Leman, M. (2018). Optimizing beat synchronized running to music. *PLOS ONE*, *13*(12), Article e0208702.

Burger, B., London, J., Thompson, M. R., & Toiviainen, P. (2018). Synchronization to metrical levels in music depends on low-frequency spectral components and tempo. *Psychological Research*, *82*(6), 1195–1211.

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*, Article 183.

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*, Article 903.

Bürkner, P. C. (2017). brms: An r package for Bayesian multilevel models using stan. *Journal of Statistical Software*, *80*(1), 1–28.

Butterfield, M. (2010). Participatory discrepancies and the perception of beats in Jazz. *Music Perception*, *27*(3), 157–176.

Chen, J. L., Fujii, S., & Schlaug, G. (2016). The use of augmented auditory feedback to improve arm reaching in stroke: A case series. *Disability and Rehabilitation*, *38*(11), 1115–1124.

Cullari, S., & Semanchick, O. (1989). Music preferences and perception of loudness. *Perceptual and Motor Skills*, *68*, Article 186.

Dobrowohl, F. A., Milne, A. J., & Dean, R. T. (2019). Timbre preferences in the context of mixing music. *Applied Sciences*, *9*(8), Article 1695.

Fraisse, P. (1982). Rhythm and tempo. *The Psychology of Music*, *1*, 149–180.

Fujii, S., Kudo, K., Ohtsuki, T., & Oda, S. (2009). Tapping performance and underlying wrist muscle activity of non-drummers, drummers, and the world's fastest drummer. *Neuroscience Letters*, *459*(2), 69–73.

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y. S. (2008). A weakly informative default prior distribution for logistic and other regression models. *Annals of Applied Statistics*, *2*(4), 1360–1383.

Hepworth-Sawyer, R., & Hodgson, J. (2016). *Mixing music*. Taylor & Francis.

Herff, S. A., Herff, C., Milne, A. J., Johnson, G. D., Shih, J. J., & Krusienski, D. J. (2020). Prefrontal high gamma in ECoG tags periodicity of musical rhythms in perception and imagination. *Eneuro*. https://doi/org/10.1523/ENEURO.0413-19.2020 https://www.eneuro.org/content/early/2020/06/25/ENEURO.0413-19.2020

Herff, S. A., Olsen, K. N., Anic, A., & Schaal, N. K. (2019). Investigating cumulative disruptive interference in memory for melodies, words, and pictures. *New Ideas in Psychology*, *55*, 68–77.

Herff, S. A., Olsen, K. N., & Dean, R. T. (2018). Resilient memory for melodies: The number of intervening melodies does not influence novel melody recognition. *Quarterly Journal of Experimental Psychology*, *71*(5), 1150–1171.

Hove, M. J., Marie, C., Bruce, I. C., & Trainor, L. J. (2014). Superior time perception for lower musical pitch explains why bass-ranged instruments lay down musical rhythms. *Proceedings of the National Academy of Sciences of the United States America*, *111*(28), 10383–10388.

Hove, M. J., Vuust, P., & Stupacher, J. (2019). Increased levels of bass in popular music recordings 1955–2016 and their relation to loudness. *The Journal of the Acoustical Society of America*, *145*(4), 2247–2253.

Huber, D. M., & Runstein, R. E. (2005). *Modern recording techniques* (6th ed.). CRC Press.

Izhaki, R. (2018). *Mixing audio: Concepts, practices and tools*. Focal Press.

Janata, P., Tomic, S. T., & Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *Journal of Experimental Psychology: General*, *141*(1), 54–75.

Karageorghis, C. I., Terry, P. C., Lane, A. M., Bishop, D. T., & Priest, D. (2012). The BASES expert statement on use of music in exercise. *Journal of Sports Sciences*, *30*(9), 953–956.

Lartillot, O., Eerola, T., Toiviainen, P., & Fornari, J. (2008). Multi-feature modeling of pulse clarity: Design, validation, and optimization. In *Proceedings of the international symposium on music information retrieval*. https://ismir2008.ismir.net/papers/ISMIR2008_145.pdf

Lartillot, O., Toiviainen, P., & Eerola, T. (2008). A Matlab toolbox for music information retrieval. In C. Preisach, H. Burkhardt, L. Schmidt-Thieme & R. Decker (Eds.), *Data analysis, machine learning and applications* (pp. 261–268). Springer.

Madison, G. (2006). Experiencing groove induced by music: Consistency and phenomenology. *Music Perception*, *24*(2), 201–208.

Madison, G., Gouyon, F., Ullén, F., & Hörnström, K. (2011). Modeling the tendency for music to induce movement in humans: First correlations with low-level audio descriptors across music genres. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(5), 1578–1594.

Menna, F., Remondino, F., Battisti, R., & Nocerino, E. (2011). Geometric investigation of a gaming active device. In *Videometrics, range imaging, and applications* XI (Vol. 8085). International Society for Optics and Photonics. https://www.spiedigitallibrary.org/conference-proceedings-of-spie/8085/80850G/Geometric-investigation-of-a-gaming-active-device/10.1117/12.890070.short?SSO=1

Meteyard, L., & Davies, R. A. (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language*, *112*, Article 104092.

Milne, A. J., & Herff, S. A. (2020). The perceptual relevance of balance, evenness, and entropy in musical rhythms. *Cognition*, *203*, Article 104233. https://doi.org/10.1016/j.cognition.2020.104233

Moelants, D. (2002). Preferred tempo reconsidered. In *Proceedings of the 7th international conference on music perception and cognition* (pp. 1–4). https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.465.8019&rep=rep1&type=pdf

North, A. C., & Hargreaves, D. J. (1996). The effects of music on responses to a dining area. *Journal of Environmental Psychology*, *16*(1), 55–64.

North, A. C., & Hargreaves, D. J. (1999). Music and driving game performance. *Scandinavian Journal of Psychology*, *40*(4), 285–292.

Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the neuronal entrainment to beat and meter. *Journal of Neuroscience*, *31*(28), 10234–10240.

Owsinski, B. (1999). *The mixing engineers handbook*. Mix Books.

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, *11*(4), 409–464.

Pressing, J. (2002). Black atlantic rhythm: Its computational and transcultural foundations. *Music Perception*, *19*(3), 285–310.

Prithvi, K., Alecu, S. R., & Dahl, S. (2019, October 14–18). *The effect of auditory pulse clarity on sensorimotor synchronization* [Conference session]. 14th International Symposium on Computer Music Multidisciplinary Research, Marseille, France.

Rafii, Z., Liutkus, A., Stöter, F. R., Mimilakis, S. I., & Bittner, R. (2017). *The MUSDB18 corpus for music separation*. https://doi.org/10.5281/zenodo.1117372

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*, *12*(6), 969–992.

Repp, B. H., & Su, Y. H. (2013). Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonomic Bulletin & Review*, *20*(3), 403–452.

Roby-Brami, A., Feydy, A., Combeaud, M., Biryukova, E., Bussel, B., & Levin, M. (2003). Motor compensation and recovery for reaching in stroke patients. *Acta Neurologica Scandinavica*, *107*(5), 369–381.

Schoeffler, M., Bartoschek, S., Stöter, F. R., Roess, M., Westphal, S., Edler, B., & Herre, J. (2018). webMUSHRA: A comprehensive framework for web-based listening tests. *Journal of Open Research Software*, *6*(1), Article 8.

Senn, O., Bechtold, T. A., Hoesl, F., & Kilchenmann, L. (2019). Taste and familiarity affect the experience of groove in popular music. *Musicae Scientiae*, *25*(1), 45–66.

Snijders, T. A., & Bosker, R. J. (2011). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. SAGE.

Stegemöller, E. L., Simuni, T., & MacKinnon, C. D. (2009). The effects of Parkinson's disease and age on syncopated finger movements. *Brain Research*, *1290*, 12–20.

Stupacher, J., Hove, M. J., & Janata, P. (2016). Audio features underlying perceived groove and sensorimotor synchronization in music. *Music Perception: An Interdisciplinary Journal*, *33*(5), 571–589.

Stupacher, J., Hove, M. J., Novembre, G., Schütz-Bosbach, S., & Keller, P. E. (2013). Musical groove modulates motor cortex excitability: A TMS investigation. *Brain and Cognition*, *82*(2), 127–136.

Thaut, M., Kenyon, G., Hurt, C., McIntosh, G., & Hoemberg, V. (2002). Kinematic optimization of spatiotemporal patterns in paretic arm training with stroke patients. *Neuropsychologia*, *40*(7), 1073–1081.

Todd, N. P. M., & Cody, F. W. (2000). Vestibular responses to loud dance music: A physiological basis of the rock and roll threshold? *The Journal of the Acoustical Society of America*, *107*(1), 496–500.

Trainor, L. J., Gao, X., Lei Jj Lehtovaara, K., & Harris, L. R. (2009). The primal role of the vestibular system in determining musical rhythm. *Cortex*, *45*(1), 35–43.

Van Dyck, E., Moelants, D., Demey, M., Deweppe, A., Coussement, P., & Leman, M. (2013). The impact of the bass drum on human dance movement. *Music Perception: An Interdisciplinary Journal*, *30*(4), 349–359.
van Noorden, L., & Moelants, D. (1999). Resonance in the perception of musical pulse. *Journal of New Music Research*, *28*(1), 43–66.
Verheul, M. H., & Geuze, R. H. (2004). Inter-limb coupling in bimanual rhythmic coordination in Parkinson's disease. *Human Movement Science*, *23*(3–4), 503–525.
Vickers, E. (2010). The loudness war: Background, speculation, and recommendations. In: *Audio engineering society convention 129*. Audio Engineering Society. https://www.aes.org/e-lib/browse.cfm?elib=15598
Witek, M. A., Clarke, E. F., Wallentin, M., Kringelbach, M. L., & Vuust, P. (2014). Syncopation, body-movement and pleasure in groove music. *PLOS ONE*, *9*(4), Article e94446.
Zapata, J. R., Holzapfel, A., Davies, M., Oliveira, J., & Gouyon, F. (2012). Assigning a confidence threshold on automatic beat annotation in large datasets. In *13th International Society for Music Information Retrieval conference* (pp. 157–162). International Society for Music Information Retrieval.