

Aesthetics of musical timing: Culture and expertise affect preferences for isochrony but not synchrony

Kelly Jakubowski^{a,*}, Rainer Polak^{b,c,**}, Martín Rocamora^d, Luis Jure^e, Nori Jacoby^f

^a Department of Music, Durham University, UK

^b Music Department, Max Planck Institute for Empirical Aesthetics, Germany

^c RITMO Centre for Interdisciplinary Studies in Rhythm, Time and Motion, University of Oslo, Norway

^d School of Engineering, Universidad de la República, Uruguay

^e School of Music, Universidad de la República, Uruguay

^f Research Group Computational Auditory Perception, Max Planck Institute for Empirical Aesthetics, Germany

ARTICLE INFO

Keywords:

Timing
Rhythm
Cross-cultural
Expertise
Aesthetic preference

ABSTRACT

Expressive communication in the arts often involves deviations from stylistic norms, which can increase the aesthetic evaluation of an artwork or performance. The detection and appreciation of such expressive deviations may be amplified by cultural familiarity and expertise of the observer. One form of expressive communication in music is playing “out of time,” including asynchrony (deviations from synchrony between different instruments) and non-isochrony (deviations from equal spacing between subsequent note onsets or metric units). As previous research has provided somewhat conflicting perspectives on the degree to which deviations from synchrony and isochrony are aesthetically relevant, we aimed to shed new light on this topic by accounting for the effects of listeners’ cultural familiarity and expertise. We manipulated (a)synchrony and (non-)isochrony separately in excerpts from three groove-based musical styles (jazz, candombe, and jembe), using timings from real performances. We recruited musician and non-musician participants ($N = 176$) from three countries (UK, Uruguay, and Mali), selected to vary in their prior experience of hearing and performing these three styles. Participants completed both an aesthetic preference rating task and a perceptual discrimination task for the stimuli. Our results indicate an overall preference toward synchrony in these styles, but culturally contingent, expertise-dependent preferences for deviations from isochrony. This suggests that temporal processing relies on mechanisms that vary in their dependence on low-level and high-level perception, and emphasizes the role of cultural familiarity and expertise in shaping aesthetic preferences.

1. Introduction

Expressivity is an important part of human aesthetic experience. The appeal of renowned artworks or performances is often associated with idiosyncratic patterns of deviation from established nominal or cognitive templates (e.g., a straight line or an isochronous rhythm) (Martindale, 1990; Stamkou, van Kleef, & Homan, 2018; Van de Cruys & Wagemans, 2011). Examples of such patterned variations include a painter’s particular brushstroke, a poet’s rhythmic feel, or a singer’s personal style of intonation; such nuanced capacities typically require extensive training to develop (e.g., Clarke, 1993; Lisboa, Williamson,

Zicari, & Eiholzer, 2005). At the same time, aesthetic engagement with artworks or performances also requires skill on the side of their audiences, who determine the value of artistic communication by realizing and appreciating its meanings. To become art connoisseurs or musical fans, for instance, often requires repeated exposure through which these audiences become sufficiently and aptly sensitive to expressive variations (Leder, Belke, Oeberst, & Augustin, 2004).

The process of expressive communication thereby involves both skilled producers and receivers who share a common ground of interactional codes, goals, and attentional foci

(Camurri, Mazzarino, Ricchetti, Timmers, & Volpe, 2003;

* Correspondence to: K. Jakubowski, Department of Music, Durham University, Palace Green, Durham DH1 3RL, United Kingdom.

** Correspondence to: R. Polak, RITMO Centre for Interdisciplinary Studies in Rhythm, Time and Motion, University of Oslo, P.O. Box 1133, Blindern, 0318, Oslo, Norway.

E-mail addresses: kelly.jakubowski@durham.ac.uk (K. Jakubowski), polak@mail.de (R. Polak).

¹ Joint first authors.

Gabrielsson & Juslin, 1996; Widmer & Goebel, 2004). In order for this communication to succeed, both the patterns and magnitudes of expressive variation need to be perceptually salient so that they can be processed by the receiver. On the other hand, exceedingly idiosyncratic patterns and/or large degrees of variation from an established norm are often recognized as exaggerated, or even vulgar (Van de Cruys & Wagemans, 2011; see also Berlyne, 1970). Thus, expressive communication requires a nuanced balance between a reference template and patterns and degrees of variation from this template.

Obtaining this balance often requires a level of cultural familiarity and/or expertise, through which producers and receivers of artistic expression learn the established norms and expectations for negotiating (following or questioning) these norms through variations. The idea that aesthetic judgments are shaped not only by properties of a stimulus but also observer-specific features including cultural background and expertise has been highlighted in existing models of aesthetic preference (Hekkert & Leder, 2009; Jacobsen, 2006) and demonstrated empirically in studies of visual art (Chokron & De Agostini, 2000; Leder et al., 2018; Masuda, Gonzalez, Kwan, & Nisbett, 2008; Silvia & Barona, 2009). Another domain in which artistic expressive communication has received much attention is music (Clarke, 1985; Fabian, Timmers, Schubert, & Eds., 2014). It has been shown, for instance, that a listener's cultural familiarity with a musical style tends to increase emotion recognition accuracy as well as the range of the recognizable emotions (e.g., Fritz et al., 2009; Laukka, Eerola, Thingujam, Yamasaki, & Beller, 2013; Thompson & Balkwill, 2012). In addition, despite popular notions that expressivity stems from individual artists' idiosyncratic genius, it appears that expressive features of performances are actually constrained by aspects of the musical structure (Repp, 1990, 1997a, 1998). Indeed, some experiments have shown that averaged versions of expressive patterns from multiple performers are preferred over individually performed ones (Repp, 1997b; Wolf, Kopiez, Platz, Lin, & Mütze, 2018), indicating that aesthetic responses to music are governed by expectations for particular expressive norms/prototypes within a musical tradition.

One common example of expressivity in music is playing “out of time” (Keil, 1987, p. 275) in various ways. For instance, the rhythmic patterns in a swing jazz piece can vary in the degree to which they are “swung” (with notes on the beat subdivision level played slightly unevenly in duration); or an accompanist in a rock band may play slightly behind the beat, or “laid back,” whilst the soloist plays slightly ahead. Two of the most prevalent ways of playing “out of time” are asynchrony (i.e., deviations from perfect synchrony between instruments) and non-isochrony² (i.e., deviations from equal spacing, or *isochrony*, between subsequent metric units, such as beats or beat subdivisions).

Previous literature presents conflicting ideas on the extent to which timing deviations from a nominal or cognitive reference structure (e.g., a music score or a perceptual prototype) are simply perceived by music listeners as “errors” or imprecisions driven by constraints of the human motor system (e.g., Wing, 1993), or whether these features contribute to the aesthetic appreciation of a performance. In the context of groove-based musical genres (e.g., American jazz), theoretical claims have been made that small-scale timing variations³ (including deviations from synchrony and isochrony) increase the rhythmicity or groove of the music in comparison to music that lacks such deviations, which has been assumed to sound mechanical and uninteresting to human ears (Keil, 1987; Prögler, 1995). In music cognition research, “groove” has been defined primarily as a pleasant urge to move along with beat-based rhythmic music, but also involves the prosocial experience of sharing a feeling of being immersed in the music together (Cámara & Danielsen,

2020; Janata, Tomic, & Haberman, 2012; Senn et al., 2019). However, empirical research has provided contradictory evidence for the notion that deviations from synchrony and isochrony within a musical performance increase listeners' experience of groove. Specifically, several experiments measuring ratings of desire to move along and enjoyment have found that stimuli utilizing timing profiles from actual/averaged musical performances elicited no significant difference (Cameron et al., 2019; Senn, Kilchenmann, von Georgi, & Bullerjahn, 2016) or even lower ratings than stimuli utilizing a greater degree of synchrony and/or isochrony (Datseris et al., 2019; Davies, Madison, Silva, & Gouyon, 2013; Hofmann, Wesolowski, & Goebel, 2017; Kilchenmann & Senn, 2015).

The perceptual studies cited above have focused primarily on responses to Western (i.e., Euro-American) music styles, such as jazz and funk, and used predominantly Euro-American participants. However, recent large-scale corpus analyses of Western and non-Western music styles have revealed cultural variations in terms of patterns of synchronization between instruments (Clayton et al., 2020; Jacoby, Polak, & London, 2021) and the consistent usage of non-isochronous metrical patterns at the beat subdivision level (Polak, Jacoby, & London, 2016; Rocamora, 2018). Similarly, previous experiments have revealed cross-cultural differences in synchronization abilities and the perception of non-isochronous rhythmic patterns, which appear to be shaped to some extent by the musical practices of one's culture (Hannon, Soley, & Ullal-Gupta, 2012; Jacoby et al., 2021; Jacoby & McDermott, 2017; Polak et al., 2018; Wittek et al., 2020). Given these differences in both usages and sensitivities to (a)synchrony and (non-)isochrony across cultures, one might expect cross-cultural differences to also manifest at the level of aesthetic evaluation.

In addition to the need to consider potential differences driven by cultural familiarity with a musical style, there is also evidence to suggest expertise differences, even *within* a particular culture, may affect both perception and aesthetic evaluations of rhythm (e.g., Senn, Kilchenmann, Bechtold, & Hoesl, 2018; Yates, Justus, Atalay, Mert, & Trehub, 2017). Musicians have been found to perform better than non-musicians in a variety of both rhythm production and perception tasks (e.g., synchronizing to auditory beats with greater precision and more consistency, Repp, 2010), which seems to depend on general rather than instrument-specific musical experience (Matthews, Thibodeau, Gunther, & Penhune, 2016). This is consistent with findings from Neuhoff, Polak, and Fischinger (2017) that both Malian musicians and dancers reliably discriminated and preferred the non-isochronous subdivision timing patterns that are typical of certain Malian music styles (Polak et al., 2016; Polak & London, 2014) over isochronous versions of these patterns. This indicates that rhythm perception abilities are enhanced not only by expertise in physically performing these drum patterns but also of dancing to them. Recently, Danielsen et al. (2021) showed that expert musicians/producers of different music styles (jazz, Nordic folk dance music, electronic dance music/hiphop) from within the Western (Euro-American) sphere perceive and synchronize differently with musical sounds, suggesting that musical expertise is to some degree style-specific and thus also influenced by musical enculturation.

In the present research, we thereby aimed to gain a more comprehensive understanding of the aesthetic evaluation of variations in (a) synchrony and (non-)isochrony in music performance by taking into account potential differences driven by cultural familiarity and expertise. In addition to broadening previous perspectives on how timing variations affect aesthetic evaluations of music, our work speaks to a wider debate around the presence or absence of “universal” features across musical cultures. For instance, cross-cultural research has revealed both convergence on geographically widespread (and perhaps culturally universal) rhythmic categories defined by prototypes at the simplest integer ratios (1:1 and 2:1), as well as considerable cultural variation in the usage of more complex ratios such as 3:2 or 4:3 (Jacoby & McDermott, 2017; Jacoby, Polak, Grahn, et al., 2021; Polak, Jacoby, et al., 2018). In this context, if our results show preferences for

² Also referred to as “anisochrony”.

³ These have been referred to using a variety of terms including “expressive timing,” “microtiming” or “microrhythm,” “participatory discrepancies,” and “systematic variation of durations.”

synchrony and/or isochrony *do not* systematically vary across culture and expertise groups, this would provide an indication that such preferences *may* be governed by culturally universal and/or biologically predetermined cognitive structures (cf., [Savage, Brown, Sakai, & Currie, 2015](#)). Alternatively, our study has the potential to reveal a more complex, culture- and expertise-dependent pattern of results in comparison to previous empirical findings of preferences for synchrony and isochrony (or only minimal variations thereof) in Western participant samples (e.g., [Datseris et al., 2019](#); [Davies et al., 2013](#); [Senn et al., 2016](#)). It may be, for instance, that previous results have been skewed by a bias toward investigations of Western music practices, where a historically contingent preference for synchrony and isochrony may have developed in the context of musical practices under the influence of musical literacy and technologies (e.g., metronomes, digital music production).

1.1. Study overview and hypotheses

We tested the influence of deviations from synchrony and isochrony on aesthetic preferences for groove-based music in a fully-balanced, cross-cultural design, as outlined in [Fig. 1](#). As previous research in this domain has focused on African-American styles of music (e.g., jazz and funk), we extended this paradigm to compare one of these Western styles (jazz music) to music with similar properties (Uruguayan candombe music, Malian jembe music). These three styles have some overlap in their (African and African-diasporic) histories and typological similarities in musical properties; this allowed us to study conceptually equivalent stimuli and constructs, which is a critical criterion for valid cross-cultural research ([He & van de Vijver, 2012](#)). For each music style, a representative excerpt was selected and subjected to separate manipulations of (a)synchrony between instruments (8 levels) and (non-)isochrony at the beat subdivision level (5 levels). Both manipulations used the timings from real musical performances, and both introduced variations in the magnitude and distribution of the asynchronies/non-isochronous pattern (see *Stimuli* in the *Method* section for full details). We selected participants from three countries (UK, Uruguay, Mali) who we anticipated would be culturally familiar with one of the musical styles (jazz, candombe, jembe, respectively) and less familiar with the other two styles. We also examined the influence of musical expertise by comparing musicians and non-musicians within each country.

Our primary research question on aesthetic preferences for the stimuli across culture and expertise groups was assessed via liking ratings for the full stimulus set. To supplement these results, we included a secondary task—same/different discrimination between pairs of stimuli. This allowed us to examine the degree to which the timing manipulations we imposed could be discriminated from one another, which was important given that we presumed differences between manipulated versions needed to be perceptually salient in order to elicit preference differences.

We hypothesized that aesthetic preferences for deviations from synchrony and isochrony would be modulated by cultural familiarity and expertise. We predicted that Western (UK) participants would show results similar to previous studies (e.g., [Datseris et al., 2019](#); [Senn et al., 2016](#)), specifically, a preference for fully synchronous and isochronous stimuli. Given the existence of cross-cultural differences in the usage and sensitivity to particular patterns of asynchrony and non-isochrony, we predicted that Uruguayan and Malian participants would show differences in their aesthetic preferences from the UK participants. In particular, previous research on Malian musicians ([Neuhoff et al., 2017](#)) suggests certain non-Western groups may be more sensitive to and place greater aesthetic value on non-isochronous metrical patterns, at least in music from their own culture. Asynchrony preferences have not been investigated cross-culturally, and thus our aims were more exploratory in this regard. We also anticipated that any cross-cultural differences would be amplified in participants with greater expertise (musicians), given their increased exposure to these musical styles (accrued via both producing and listening). In sum, this experiment aimed to reveal new insights on the aesthetic experience of rhythm and the degree to which top-down factors such as enculturation and expertise modulate such experiences.

2. Method

2.1. Stimuli

2.1.1. Musical excerpts

For each music style, we selected one excerpt (8 metric beats, 3.2–3.7 s) from a live performance recorded with multi-track equipment in a studio context. All recording sets were produced by music researchers specialized in the respective styles, in collaboration with expert musicians. The same researchers and musicians were responsible for the selection of stylistically representative excerpts for the present study. The recordings of candombe and jembe music came from our own research archives ([Jure & Rocamora, 2016](#); [Jure, Rocamora, Tarsitani, & Clayton, 2020](#); [Polak et al., 2016](#); [Polak, Tarsitani, & Clayton, 2018](#); [Rocamora et al., 2015](#)); the jazz recording was provided by Olivier Senn and Lorenz Kilchenmann (Lucerne University of Applied Sciences and Arts) from materials used in previous timing and groove studies ([Kilchenmann & Senn, 2015](#); [Senn et al., 2016](#)).

For each excerpt, we measured timings by manually marking all event (instrumental) onsets. To check that the selected excerpts were representative of the styles in question with respect to their non-isochronous timing, we first calculated the mean locations of metric positions (beats and beat subdivisions) in each excerpt by averaging the timings of all events that fell on the same metric position. We found a mean non-isochronous beat subdivision of 77:23 for the jazz excerpt, 24:23:22:31 for the candombe excerpt, and 25:33:42 for the jembe

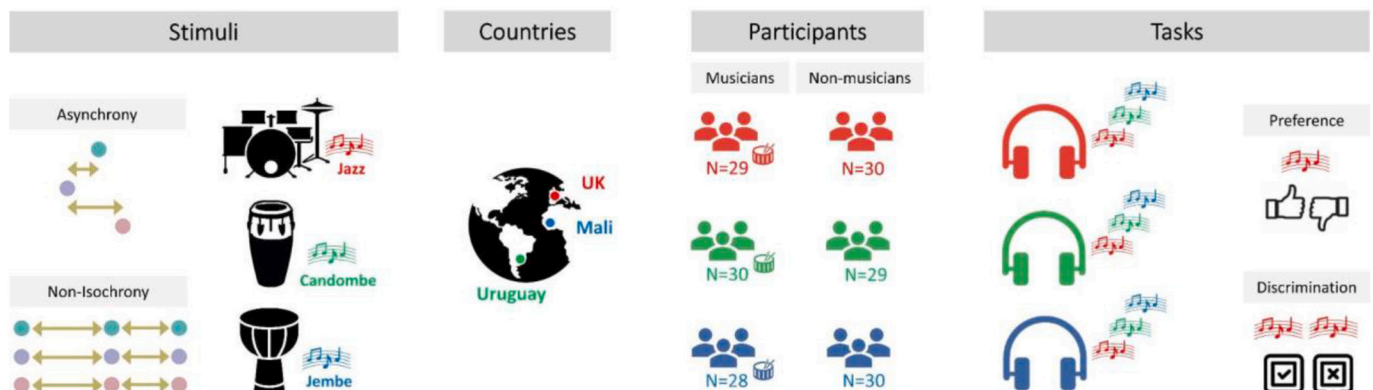


Fig. 1. Schematic representation of the study design.

excerpt. We then computed the mean asynchronies between event onsets in the same metric position by taking the root mean square (RMS) of pairwise asynchronies and then taking the RMS of all pairs (Rasch, 1988). The mean asynchrony values were 19 ms for the jazz excerpt, 14 ms for the candombe excerpt, and 13 ms for the jembe excerpt. These non-isochronous subdivision and asynchrony values are all representative of the styles in question when compared against large-scale corpus analyses of sets of expert-curated recordings (Clayton et al., 2020; Friberg & Sundström, 2002; Jure & Rocamora, 2016; Prögler, 1995; Rocamora, 2018). Fig. 2A plots the exact timing of the event onsets and the metric reference structure of beat and subdivision locations for each of the selected excerpts. For further details of the selected recordings, see Supplementary Materials 1.

2.1.2. Manipulations

We performed two independent manipulations (synchrony and isochrony) on each musical excerpt, using the timings from the original performance (see Fig. 2B). In both cases, the magnitude of the manipulations was proportional to the magnitude of variations in the original performances. These manipulations introduced variations to the magnitude and distribution of the asynchronies/non-isochronous pattern (see Supplementary Materials 2 for a detailed explanation of this methodological decision).

The synchrony manipulation introduced alterations to both the magnitude and distribution of asynchronies between different instruments realizing events in the same metric positions (see Table 1). We included both the original magnitude and distribution of asynchronies as performed (Asynchrony Condition 1-ori) and a fully synchronized (“quantized”) version with all asynchronies reduced to zero (Asynchrony Condition 0-qua). We also linearly increased the magnitude of asynchrony, to create versions that doubled and tripled (Asynchrony Conditions 2-ori and 3-ori) the magnitude of the original asynchronies. In addition, we altered the distribution of the asynchronies in two ways. In the first, we retained the magnitude but inverted the sign of the asynchrony of each event from the respective metric position (Asynchrony Condition 1-inv); for example, an event that originally occurred 15 ms after the beat would now occur 15 ms before the beat. In the second type of distribution manipulation, we shuffled the original asynchronies within an excerpt across all of the instruments and events involved, thus creating a non-patterned, random distribution of asynchronies. We applied this manipulation to the original magnitude (Asynchrony Condition 1-ran) as well as the doubled and tripled magnitudes of asynchrony (Asynchrony Conditions 2-ran and 3-ran).

Our second manipulation concerned the (non-)isochrony of the metric structure, that is, the relative duration of beats and beat subdivisions in the metric cycle (see Table 2). Since all three music styles are based on a very stable isochronous beat, our manipulation targeted the beat subdivision level. For this manipulation, we reduced all asynchronies to zero, to focus solely on the effects of different levels of deviation from isochrony on participant preferences. We included both the original pattern (Non-Isochrony Condition 1-ori) and an isochronous pattern (Non-Isochrony Condition 0-iso) that was plausible from a culturally informed, music-theoretical perspective (detailed in the following paragraph) as stimuli. In analogy to the synchrony manipulation, we also doubled the magnitude of the distance between the isochronous and original pattern (Non-Isochrony Condition 2-ori) and manipulated the distribution of the non-isochronous patterns, namely, by inverting the original and doubled patterns (Non-Isochrony Conditions 1-inv and 2-inv).

The definition of the isochronous pattern to which each excerpt was compared was straightforward in the cases of candombe and jembe, where each beat subdivision is realized by rhythmic events performed by one or more instruments in the ensemble (see Fig. 2A). However, in the case of swing jazz, although the pattern used here realizes two events per beat, this pattern is primarily discussed (by practitioners, as well as in music theory and pedagogy) with reference to a ternary subdivision of

the beat (33:33:33), of which only events 1 and 3 are sounded (67:33) (Benadon, 2006; Spring, 2014). It is in this sense that we speak of “isochronous” metric subdivision in the case of jazz.

2.1.3. Stimulus generation

We synthesized the stimuli from audio samples of single instrument sounds. The sound samples of the candombe and jembe ensemble instruments were studio recordings of professional players. The drum sounds for the jazz stimuli were taken from the sample library *jazz/funk kit* by the manufacturer Orange Tree Samples. The bass track was synthesized from slices of the original audio, which was perfectly clean so that it did not introduce any artifacts to the stimuli (cf. Kilchenmann & Senn, 2015, who took the same approach). We compiled soundbanks comprising four variants of each single sound, which were randomly triggered by a MATLAB MIDI sequencer on the grounds of the timbral/melodic information from the original performances and the manipulated timing information. This allowed us to isolate the timing manipulations and keep all other acoustic parameters constant, while creating stimuli that still sounded quite realistic and musical. In order to control for any potential effects of the random triggering from the four samples per sound in the soundbanks or the random shuffling of the asynchronies across instruments in the random asynchrony conditions, we created 10 versions of each stimulus that were presented in a counterbalanced order across participants. The jazz stimuli were presented at a tempo of 150 bpm, candombe at 120 bpm, and jembe at 130 bpm, which are all in the typical range of tempi for these music styles and patterns (Dittmar, Pfeleiderer, Balke, & Müller, 2017; Jacoby, Polak, & London, 2021; Rocamora, 2018). All stimuli can be accessed on the Open Science Framework here: osf.io/uebdk

2.2. Participants

We recruited 58 to 59 participants in each of three countries (UK, Uruguay, Mali), comprising two sub-groups delineated on the basis of expertise (hereafter referred to as musicians and non-musicians). Our primary objective in sampling participant groups was to operationalize the level of cultural familiarity with the musical stimuli. Specifically, we assumed that the inhabitants of the three countries (UK, Uruguay, Mali) would be primarily familiar with one of the tested music styles (jazz, candombe, jembe, respectively). Sociocultural heterogeneity and access to mass media in most parts of the world today make it plausible that these participants have been exposed to a variety of local, regional and international music styles. Similarly, most musicians today, even if specializing in a specific style, have experience in playing other styles. Our main consideration was thus to ensure that both musicians and non-musicians were *relatively* more familiar with the music of their country than the music of the other countries, whilst acknowledging they might have some familiarity with the other styles used (to a lesser degree). As such, self-report ratings of prior familiarity with each music style (on a 1–7 scale) were taken at the start of the experiment by presenting the style labels and their respective culture-geographic origins (Euro-American jazz, Uruguayan candombe, Malian jembe) to the participants. Given the ordinal dependent variable (familiarity ratings), we fit a cumulative link mixed model with Country of residence (UK/Uruguay/Mali), Expertise (musician/non-musician), Cultural Familiarity of the music excerpt (own culture/other culture), and a random effect of Participant as predictors of familiarity ratings.⁴ From this model, we extracted two sets of pairwise contrasts (with Bonferroni correction). In the first, we found that all six Country × Expertise groups gave significantly higher familiarity ratings for the music from their own culture than music from other cultures (all $ps < 0.001$). In the second, we found that musicians gave higher familiarity ratings than non-musicians for

⁴ Names of independent variables are capitalized here and in subsequent reporting of analyses to highlight the main factors under investigation.

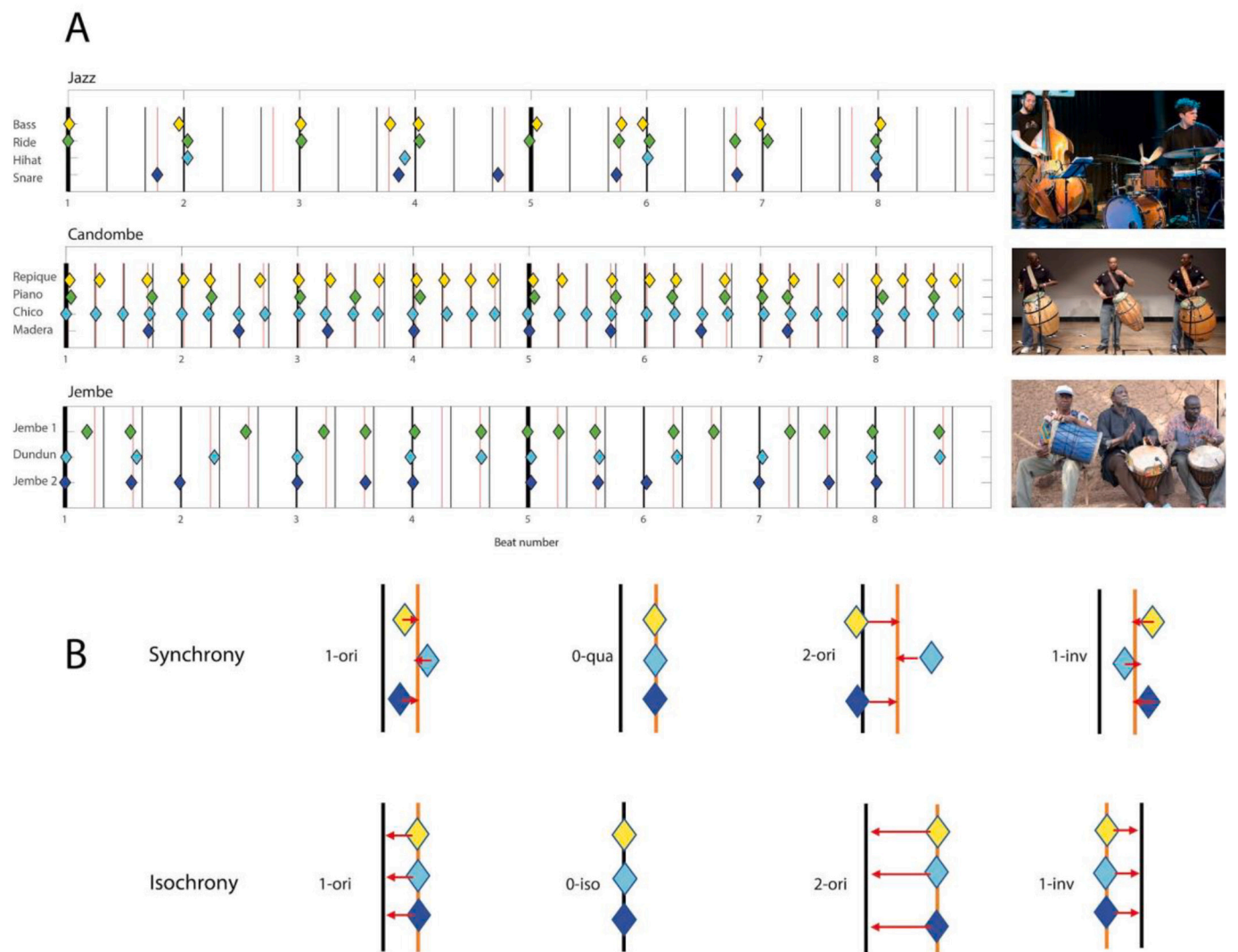


Fig. 2. Stimuli and manipulations. A. The three excerpts selected for the experiment. Black vertical lines mark the isochronous metric positions; thickness indicates the metric level: thick = 4-beat cycle, medium = beat, thin = beat subdivision. Thin red lines indicate the averaged non-isochronous subdivisions as performed in the recordings from which the excerpts were taken. B. Schematic examples of the synchrony and isochrony manipulations. Each example shows event onsets for only one metric position. Black and red vertical lines schematically mark the isochronous and averaged performed metric positions, respectively. Red arrows indicate deviations of onsets from averaged onset timings in the synchrony manipulation and deviations of onsets from an isochronous metric location in the isochrony manipulation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Manipulations of synchrony.

Asynchrony Condition	Definition
0-qua	Quantized by reducing all asynchronies to zero
1-ori	Original magnitude and distribution of asynchronies as performed
2-ori	Doubled magnitude of the original distribution of asynchronies
3-ori	Tripled magnitude of the original distribution of asynchronies
1-inv	Original magnitude with inverted distribution of asynchronies (reversed sign of deviation from the mean metric position for each onset)
1-ran	Original magnitude with a random (shuffled) distribution of asynchronies
2-ran	Doubled magnitude with a random (shuffled) distribution of asynchronies
3-ran	Tripled magnitude with a random (shuffled) distribution of asynchronies

Note. The Asynchrony Condition naming format first considers the magnitude (values 0, 1, 2, 3) and then specifies the type of distribution (qua = quantized, ori = original, inv = inverted, ran = random) of the asynchronies.

their own country’s music in all three countries (all $ps < 0.001$) (see Fig. S1 in Supplementary Materials).

The main criterion for classifying participants as musicians (in the UK, Uruguay, and Mali, respectively) was an extensive and current experience in performing in the music styles (jazz, candombe, or jembe, respectively) used in the experiment. Only instrumental musicians (no vocalists) were included, since all our stimuli comprised instrumental excerpts. The non-musicians, by contrast, reported little or no experience in performing any style of music; most were current university students (see Table 3). In independent-samples Wilcoxon rank-sum tests, all musician groups reported significantly more years of musical training than the corresponding non-musician group for that country (all $ps < 0.001$, with Bonferroni correction). We also took a more objective measure of musicianship by calculating participants’ asynchronies (difference between stimulus and response) while tapping to an isochronous beat and computing the standard deviation of these asynchronies for each participant; in previous studies, musicians have been shown to perform less variably on such a task than non-musicians (Jacoby, Polak, Grahn, et al., 2021; Polak, Jacoby, et al., 2018; Repp,

Table 2

Manipulations of isochrony, and timing pattern for each condition in percent of beat duration by excerpt.

Non-Isochrony Condition	Definition	Jazz	Candombe	Jembe
	Full isochrony of metric subdivisions	33:33:33	25:25:25:25	33:33:33
0-iso	Rhythmic realization of isochronous subdivisions	67:33	25:25:25:25	33:33:33
1-ori	Original magnitude and distribution of non-isochrony as performed	77:23	24:23:22:31	25:33:42
1-inv	Original magnitude of non-isochrony with inverted distribution	56:44	26:27:28:19	42:34:25
2-ori	Doubled magnitude of original distribution of non-isochrony	87:13	23:21:19:37	17:33:51
2-inv	Doubled magnitude of inverted distribution of non-isochrony	46:54	27:29:31:13	50:34:16

Note. The Non-Isochrony Condition naming format first considers the magnitude (values 0, 1, 2) and then specifies the type of distribution (iso = isochronous, ori = original, inv = inverted) of the pattern.

2010). In each country, the musician group tapped less variably than the non-musician group (all $ps < 0.016$ in independent-samples t -tests, with Bonferroni correction). There were some differences in age and education between the groups, although these are relatively representative of general cultural differences, such as the fact that jembe and candombe music are not traditionally taught at schools/conservatories, unlike jazz music in the UK (see additional analyses in Supplementary Materials 3).

2.3. Experimental tasks

For the preference task, we created stimuli of 6–8 s in duration by repeating the eight-beat excerpts twice in a seamless loop. In this task, participants were asked to rate how much they liked each stimulus from each music style and condition (for both the synchrony and isochrony manipulations) on a 4-point scale (dislike a lot/dislike a little/like a little/like a lot). These rating scales were accompanied by visual depictions (two thumbs down/one thumb down/one thumb up/two thumbs up; see Fig. 3B) that were confirmed to be comprehensible in all three countries.

For the discrimination task, we created stimuli of 8–10 s in duration, which comprised two of the eight-beat excerpts (3–4 s) separated by 2 s of silence. Half of the trials comprised two presentations of the same stimulus, and half consisted of two different stimuli. After hearing both stimuli, participants were asked to judge whether the two stimuli were the same or different. The response options were also accompanied by visual depictions (OO for same, OX for different; see Fig. 3B). We did not test the full matrix of all possible “different” pairings, as this would have significantly increased the duration of the experiment, with likely fatigue effects. For the synchrony manipulation, “different” trials

comprised comparisons between the quantized stimulus (Asynchrony Condition 0-qua) and all other synchrony manipulations. For the isochrony manipulation, “different” trials comprised comparisons between each manipulated version of the stimulus and Non-Isochrony Condition 1-ori (the original, non-isochronous timing).

Two tapping tasks were also administered. The first was an isochronous tapping task, in which participants were asked to tap in synchrony with a 150 bpm isochronous sequence of noise bursts of approximately 33 s in duration (see results from this task in the *Participants* section). In the second tapping task, participants heard a looped version of each of the original stimuli for each music style of approximately 30 s in duration and were asked to tap along to the beat; results of this task will not be reported here as they are outside the scope of the present research questions.

The preference and discrimination tasks were run via OpenSesame (Mathôt, Schreij, & Theeuwes, 2012), and responses to these tasks were made using a computer keyboard. Visual icons displaying the possible response options for the preference (thumbs up/down symbols) and discrimination (OO and OX) tasks were attached as sticker labels to the computer keyboard. The stickers were distributed such that there was at least one key between each sticker position (with functionless keys between the relevant keys), to minimize the possibility of accidentally striking a wrong key. These measures made it easier for any participants who were not used to completing psychological experiments or using computers. Practice trials helped to verify that participants understood and were capable of performing the tasks.

Tapping data were collected via a device with a soft surface and microphone installed in the interior, which participants held with one hand either on their lap or on a table in front of their seats while tapping with the other hand. The tapping setup and automatic onset extraction of taps were identical to that reported in Jacoby and McDermott (2017). Tapping stimuli were presented and recorded in either Cubase or Audacity software.

All participants wore headphones and were able to request volume adjustments during the practice trials, after which a constant volume was maintained throughout the main experiment. All tasks were conceived of and first designed in English; all materials relevant to the procedure (information, consent, task instructions, rating scales, etc.) were translated to Spanish for sessions run in Uruguay and French as well as Bambara (official language and lingua franca, respectively) for sessions run in Mali.

2.4. Procedure

The study received ethical approval from the Durham University Music Department Ethics Committee and the Columbia University Institutional Review Board (IRB protocol: IRB-AAAR3726). After providing informed consent and demographic information, each participant first completed two trials of the isochronous tapping task. They then completed the preference task, in which they were asked to rate how much they liked each musical stimulus across all music styles and conditions (for both manipulation types). Preference stimuli were

Table 3

Demographic information for musician and non-musician participant groups from the UK, Uruguay, and Mali.

	UK Musicians	UK Non-Musicians	Uruguayan Musicians	Uruguayan Non-Musicians	Malian Musicians	Malian Non-Musicians
Total Number of Participants	29	30	30	29	28	30
Mean Age in Years (SD)	30 (13)	30 (13)	45 (13)	25 (3)	38 (12)	26 (3)
Gender	2 Female, 27 Male	21 Female, 9 Male	1 Female, 29 Male	15 Female, 14 Male	1 Female, 27 Male	15 Female, 15 Male
Mean Years of Formal Education (SD)	16 (2)	16 (3)	9 (3)	17 (2)	6 (5)	16 (2)
Total Number who Play an Instrument or Sing	29	6	30	0	28	6
Mean Years of Regular Music Practice (SD)	17 (10)	1 (2)	35 (14)	0 (0)	24 (13)	1 (4)
Mean Standard Deviation of Asynchrony for Isochronous Tapping (SD)	17.17 (5.49)	26.63 (7.93)	15.48 (3.87)	23.60 (13.99)	13.38 (2.73)	21.35 (10.21)

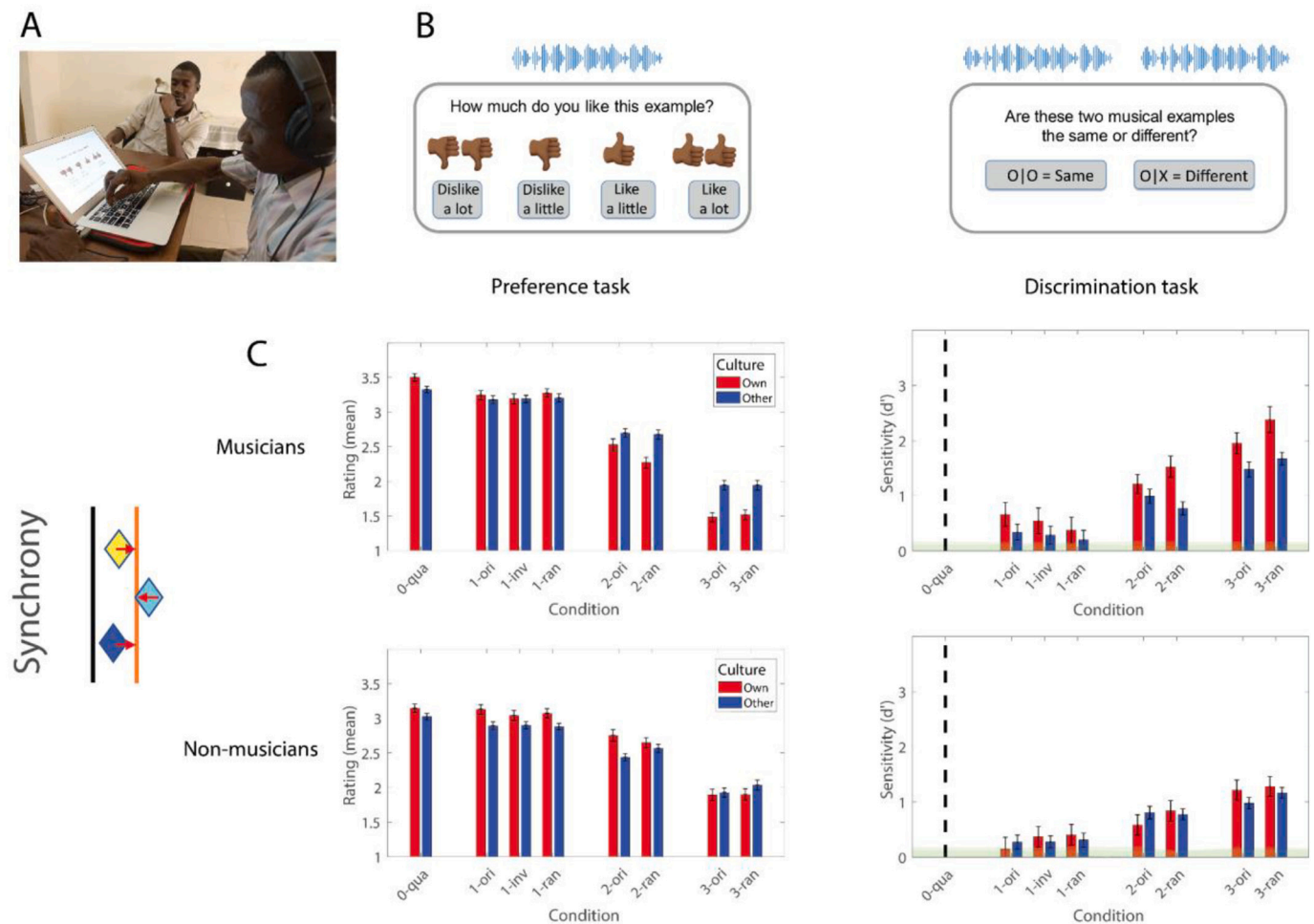


Fig. 3. Synchrony manipulation task and results. A. Procedure (photograph from a session in Mali). B. Task questions and rating scales. C. Results of the preference and discrimination tasks for the synchrony manipulation. Results are separated by task (left column: preference, right column: discrimination), expertise (upper tier: musicians, lower tier: non-musicians), and cultural familiarity (red: own culture, blue: other culture). Error bars represent one standard error of the mean across participants for the preference task and one standard deviation of the d' estimated via generating 1000 bootstrapped datasets with replacement for the discrimination task. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

blocked by music style and manipulation type (synchrony/isochrony) and counterbalanced across participants, with each block prefaced by two practice trials. Each stimulus was rated twice, resulting in 78 total ratings in the main preference task. Next, participants completed two trials of tapping to the beat of each of the original musical stimuli. Finally, they completed the pairwise discrimination (same/different) task. The discrimination task was always completed at the end of the experiment in order to avoid biasing ratings in the preference task, since the reference stimulus (0-qu in synchrony and 1-ori in isochrony manipulations) was presented more often than the comparison stimuli in the discrimination task. We counterbalanced both the blocking of stimuli by music style and manipulation type and the order of presentation within pairs (AB vs. BA) across participants. Each block of the discrimination task began with four practice trials with feedback (correct/incorrect). Participants were able to listen again to each practice stimulus if desired after receiving the feedback. No feedback was provided during the main discrimination task, and each same or different pairing was presented once per participant, resulting in 72 trials total.

2.5. Analysis

For the preference task, given the ordinal dependent variable (preference ratings on a 4-point scale) and repeated-measures nature of the synchrony/isochrony manipulations, the data were analyzed via

cumulative link mixed models using the 'ordinal' package in R (Christensen, 2019), and the statistical significance of the fixed effects was assessed via likelihood ratio χ^2 tests using the 'RVAideMemoire' package (Hervé, 2022). Separate models were fitted for the synchrony and isochrony manipulations. For the synchrony manipulation, the model included fixed effects of Country of residence (UK/Uruguay/Mali), Expertise (musician/non-musician), Cultural Familiarity of the music excerpt (own culture/other culture), and Asynchrony Condition (8 levels, see Table 1), with a random effect of Participant, as predictors of preference ratings. We also included all two-, three-, and four-way interactions of the fixed effect variables. The analysis of the isochrony manipulation followed the same format, but with Non-Isochrony Condition (5 levels, see Table 2) instead of Asynchrony Condition. Rather than using the raw Music Style (jazz/candombe/jembe) factor as a predictor, we utilized Cultural Familiarity as a more specific variable that classified each style as being from one's own culture (i.e., jazz in UK, candombe in Uruguay, jembe in Mali) or another culture (i.e., all other Country/Music Style combinations). This approach is more aligned with testing our hypotheses that cultural familiarity plays a role in synchrony/isochrony preferences, although the full breakdown of the dataset by Music Style can also be observed within the Supplementary Materials (see Figs. S2, S3, S4, and S5). An analogous approach to this was also used to test the effects of Country, Expertise, and Cultural Familiarity on familiarity ratings for each of the music styles via a

cumulative link mixed model; see these results in the *Participants* section of the *Method*.

From these initial, full factorial models, we then extracted two sets of post hoc comparisons. First, we extracted pairwise comparisons (with Bonferroni correction) for all Asynchrony Conditions, in order to examine the extent to which the manipulations of both the magnitude and distribution of asynchronies impacted preference ratings. Analogous pairwise comparisons were extracted for all Non-Isochrony Conditions from the model for that data. Second, as a central question within this research program was to investigate whether the original performance timing was preferred over a fully synchronous/isochronous version and whether such preferences varied in relation to cultural familiarity and expertise, and given the significant four-way interactions that emerged in the full factorial models, we then extracted a set of planned contrasts from these initial models. Specifically, these contrasts comprised the pairwise comparisons of preference ratings for the original (as performed) stimulus (1-ori) versus the synchronous/isochronous version (0-qua/0-iso) for each Country, Cultural Familiarity, and Expertise combination. This analysis thereby allowed us to examine the effects of cultural familiarity and expertise on preferences for the original versus fully synchronous/isochronous separately for each country. Contrast analyses were performed using the ‘emmeans’ package in R (Lenth, 2022).

The discrimination task results were then used to support and further interpret the results of each of the preference tasks (i.e., to examine parallels between preferences for a stimulus and its perceptual discriminability). For the discrimination tasks, we could not compute participant-wise sensitivity analyses (d -prime: d'), since each participant performed only a small number of trials per condition. Instead, we aggregated the data from all participants and computed d' values for the aggregated data (as if the data came from a single participant). To compute error bars, we generated 1000 bootstrapped datasets (with replacement). The null distribution was obtained by computing 1000 bootstrapped datasets where the responses were shuffled (so that the match between a stimulus and response was permuted). To compute the statistical significance of d' values relative to the null distribution we used a similar procedure but instead used 10,000 bootstrapped datasets (the additional bootstrapping was needed to account for multiple comparisons).

All data can be accessed via the Open Science Framework: osf.io/uebdk.

3. Results

3.1. Synchrony manipulation

Table 4 shows the results of the likelihood ratio χ^2 tests for the main effects and interactions within the cumulative link mixed model predicting preference ratings for the synchrony manipulation. A full breakdown of all effects for all levels of the factors is also provided in the Supplementary Materials (Table S1). A comparison of this fitted model to a null model (intercept-only model, with a random effect of Participant) produced a Nagelkerke pseudo-R-squared value of 0.34.

As shown in Table 4, almost all predictors and interactions included in the model were statistically significant. The main effect of Country was driven by Malians giving somewhat higher preference ratings than participants from other countries, while the main effect of Expertise showed that musicians gave higher preference ratings overall than non-musicians. Of more direct relevance to the current research questions, a strong, significant main effect of Asynchrony Condition was also found. In pairwise contrasts extracted from the initial model with Bonferroni correction for 28 comparisons, we found that, overall, the quantized stimuli were preferred over both the original version (1-ori: $p = .005$) as well as the other two manipulations using the original magnitude of asynchrony (1-ran: $p = .005$; 1-inv: $p < .001$). All manipulations using the original magnitude of asynchrony were preferred over those in

Table 4

Cumulative link mixed model results for the effects of country, expertise, cultural familiarity, and asynchrony condition on preference ratings.

Predictor	Likelihood Ratio χ^2	df	p-value
Country	16.73	2	< 0.001***
Expertise	5.89	1	0.02*
Cultural Familiarity	0.50	1	0.48
Asynchrony Condition	2700.53	7	< 0.001***
Country × Expertise	2.28	2	0.32
Country × Cultural Familiarity	52.55	2	< 0.001***
Expertise × Cultural Familiarity	45.84	1	< 0.001***
Country × Asynchrony Condition	50.70	14	< 0.001***
Expertise × Asynchrony Condition	130.83	7	< 0.001***
Cultural Familiarity × Asynchrony Condition	83.37	7	< 0.001***
Country × Expertise × Cultural Familiarity	24.15	2	< 0.001***
Country × Expertise × Asynchrony Condition	36.57	14	0.001**
Country × Cultural Familiarity × Asynchrony Condition	76.99	14	< 0.001***
Expertise × Cultural Familiarity × Asynchrony Condition	30.02	7	< 0.001***
Country × Expertise × Cultural Familiarity × Asynchrony Condition	30.51	14	0.006**

Note. * = $p < .05$, ** = $p < .01$, *** = $p < .001$; “Participant” was included as a random effect in the model; the standard deviation of this random effect was 0.68.

which the magnitude of asynchrony was doubled (all $ps < 0.001$), and those with a doubled magnitude of asynchrony were preferred over those with tripled asynchrony ($ps < 0.001$). No significant differences were found between Asynchrony Conditions of the same magnitude in which we varied the *distribution* of the asynchronies (e.g., 1-ori vs. 1-ran vs. 1-inv) (all $ps > 0.99$).

With the exception of Country × Expertise, all interactions in the model were statistically significant, and Fig. 3C gives an overview of the relationship between Expertise, Cultural Familiarity, and Asynchrony Condition (see also Fig. S2 for data from all Music Styles by Country, Expertise group, and Asynchrony Condition). In particular, it is notable that the Expertise by Asynchrony Condition interaction is primarily driven by the musicians showing more polarized preference responses across the range of magnitudes of the asynchrony manipulation than non-musicians. The interaction of Cultural Familiarity and Asynchrony Condition is primarily driven by more polarized preference differences across the range of asynchrony manipulations for the music of one’s own culture; this is particularly apparent in the musician group.

Fig. 5A shows the results of the pairwise contrasts extracted from the initial model in which we compared preference ratings for the original stimulus (1-ori) versus the fully synchronous/quantized (0-qua) version by Cultural Familiarity and Expertise for each Country. This reveals a relative lack of any statistically significant differences, with no group showing a significant preference for the original timing pattern over the quantized version. The UK musicians even showed a preference for the quantized version of culturally familiar music (jazz, in their case) over the original performance ($p = .008$ with Bonferroni correction), and the UK non-musicians showed a small but significant preference for the quantized version of culturally unfamiliar music ($p = .046$ with Bonferroni correction).

To allow for comparison between the preference task results and performance on the discrimination task, Fig. 3C also shows the discrimination task results for the synchrony manipulation as a function of Expertise and Cultural Familiarity (see also Fig. S3). The nominal d'

values for conditions with small magnitudes of asynchrony were relatively small (in the range of 0.14–0.68 for the 1-ori, 1-inv, and 1-ran conditions). Of these Asynchrony Conditions, the d' values were significantly above the null distribution (green line in Fig. 3) for four conditions (musicians listening to excerpts from their own and other cultures for Condition 1-ori: $p = .001$ and $p = .035$, musicians listening to excerpts from their own culture for Condition 1-inv: $p = .011$, and non-musicians listening to excerpts from other cultures for Condition 1-ran: $p = .036$). However, d' values were larger (in the range of 0.59–2.37) for the larger magnitudes of the manipulation (2-ori, 2-ran, 3-ori, 3-ran), and significantly above the null distribution (all $ps > 0.003$, via bootstrapping). Taken together, the results in Fig. 3C show that participants were both highly sensitive and showed more pronounced preference differences for the more exaggerated versions of the manipulation, whereas preference differences and sensitivities were relatively small for manipulations with a magnitude of asynchrony that is similar to the original performance. This suggests that reduced differences in preferences between stimuli co-occur with a reduced ability to perceptually discriminate between them.

3.2. Isochrony manipulation

Table 5 displays the results of the cumulative link mixed model analysis for preference ratings of the isochrony manipulation (see also Table S2, for a breakdown of all effects for all levels of the factors). When compared against a null model (intercept-only model, with a random effect of Participant), the Nagelkerke pseudo-R-squared value for this model was 0.27.

The significant main effect of Cultural Familiarity showed that, on the whole, music from other cultures was given higher preference ratings than music from one's own culture. The main effect of Expertise again showed that musicians gave higher preference ratings overall than non-musicians, and the main effect of Country was driven by Uruguayans giving lower overall ratings than the other groups. The significant main effect of Non-Isochrony Condition was explored in post hoc

Table 5

Cumulative link mixed model results for the effects of country, expertise, cultural familiarity, and non-isochrony condition on preference ratings.

Predictor	Likelihood Ratio χ^2	df	p-value
Country	10.18	2	0.006**
Expertise	5.35	1	0.02*
Cultural Familiarity	29.97	1	< 0.001***
Non-Isochrony Condition	680.48	4	< 0.001***
Country $\times$ Expertise	4.23	2	0.12
Country $\times$ Cultural Familiarity	56.20	2	< 0.001***
Expertise $\times$ Cultural Familiarity	84.67	1	< 0.001***
Country $\times$ Non-Isochrony Condition	149.45	8	< 0.001***
Expertise $\times$ Non-Isochrony Condition	48.05	4	< 0.001***
Cultural Familiarity $\times$ Non-Isochrony Condition	83.42	4	< 0.001***
Country $\times$ Expertise $\times$ Cultural Familiarity	9.26	2	0.01*
Country $\times$ Expertise $\times$ Non-Isochrony Condition	42.00	8	< 0.001***
Country $\times$ Cultural Familiarity $\times$ Non-Isochrony Condition	287.98	8	< 0.001***
Expertise $\times$ Cultural Familiarity $\times$ Non-Isochrony Condition	35.60	4	< 0.001***
Country $\times$ Expertise $\times$ Cultural Familiarity $\times$ Non-Isochrony Condition	32.77	8	< 0.001***

Note. * = $p < .05$, ** = $p < .01$, *** = $p < .001$; "Participant" was included as a random effect in the model; the standard deviation of this random effect was 0.65.

pairwise contrasts extracted from the initial model with Bonferroni correction for 10 comparisons. These showed that, overall, the isochronous pattern (0-iso) was preferred over both the original, non-isochronous pattern (1-ori) ($p < .001$) and the inverted version of the original pattern (1-inv) ($p < .001$). The original pattern was preferred over the doubled (2-ori) ($p < .001$) and doubled-inverted pattern (2-inv) ($p < .001$). The original pattern was also preferred over the inverted pattern of the same magnitude ($p < .001$), but no overall differences in preferences were found between the doubled pattern and doubled-inverted pattern ($p > .99$).

The main effects and interactions of Expertise, Cultural Familiarity, and Non-Isochrony Condition can be seen in Fig. 4C (these data are further broken down by Music Style and Country in Fig. S4). The Expertise by Non-Isochrony Condition interaction was again driven by the musicians showing more polarized preference differences across the range of manipulations. The Cultural Familiarity by Non-Isochrony Condition interaction was particularly driven by a divergence in ratings for one's own versus other cultures' music between the original timing pattern (1-ori) and the isochronous version (0-iso). That is, overall, preference ratings seemed to favor the isochronous version over the original timing pattern for culturally unfamiliar, but not culturally familiar, music; this general pattern is unpacked in more detail in the contrast analysis reported below.

In analogy to the post hoc analyses performed for the synchrony manipulation, we extracted from the initial model Bonferroni-corrected contrasts comparing the original stimulus (1-ori) against the isochronous (0-iso) version by Cultural Familiarity and Expertise within each Country; results are presented in Fig. 5B. For UK musicians there was a clear preference for isochronous versions of both culturally familiar (jazz; $p < .001$) and culturally unfamiliar music ($p < .001$), with UK non-musicians exhibiting a weaker preference toward isochrony that was only statistically significant for culturally unfamiliar music ($p < .001$). However, musicians in Uruguay and Mali preferred the original, performance-based version over the isochronous variant for their own music styles (Mali: $p = .005$; Uruguay: $p < .001$), suggesting that these expert musicians have internalized style-specific timing prototypes as aesthetic ideals. Non-musicians showed no significant preferences for the original patterns over isochronous versions, and Malian non-musicians even showed a small preference toward the isochronous version of music from their own country (jembe; $p = .032$).

Fig. 4 also shows the discrimination task results for the isochrony manipulation as a function of Expertise and Cultural Familiarity (see also Fig. S5). Although the difference in preference ratings between the isochronous and original patterns was relatively small, participants were able to distinguish the difference between these two conditions, as the discrimination task results showed relatively good discrimination ability (d' values were all > 1.11 and 0.97 for musicians and non-musicians, respectively). Note also that d' values for all Non-Isochrony Conditions were significantly above chance ($ps < 0.001$ via bootstrapping, comparing the d' values from our experiment to the ones obtained for the randomly shuffled data), indicating that all manipulations of isochrony were successfully discriminated from the original (1-ori) stimulus by both musicians and non-musicians for both culturally familiar and unfamiliar music.

3.3. Post hoc replication in US jazz musicians

In light of the finding that UK musicians did not show a preference for the non-isochronous timing pattern based on an actual performance of music from their culture (jazz), in contrast to Uruguayan and Malian musicians, we collected an additional, post hoc set of data from a group of jazz musicians in a different location (the US). As jazz music is arguably a more globally widespread style, with many different groups of musicians throughout the world performing in many different sub-traditions, this allowed us to test whether our UK results generalized to another cultural group with similar experience in performing this

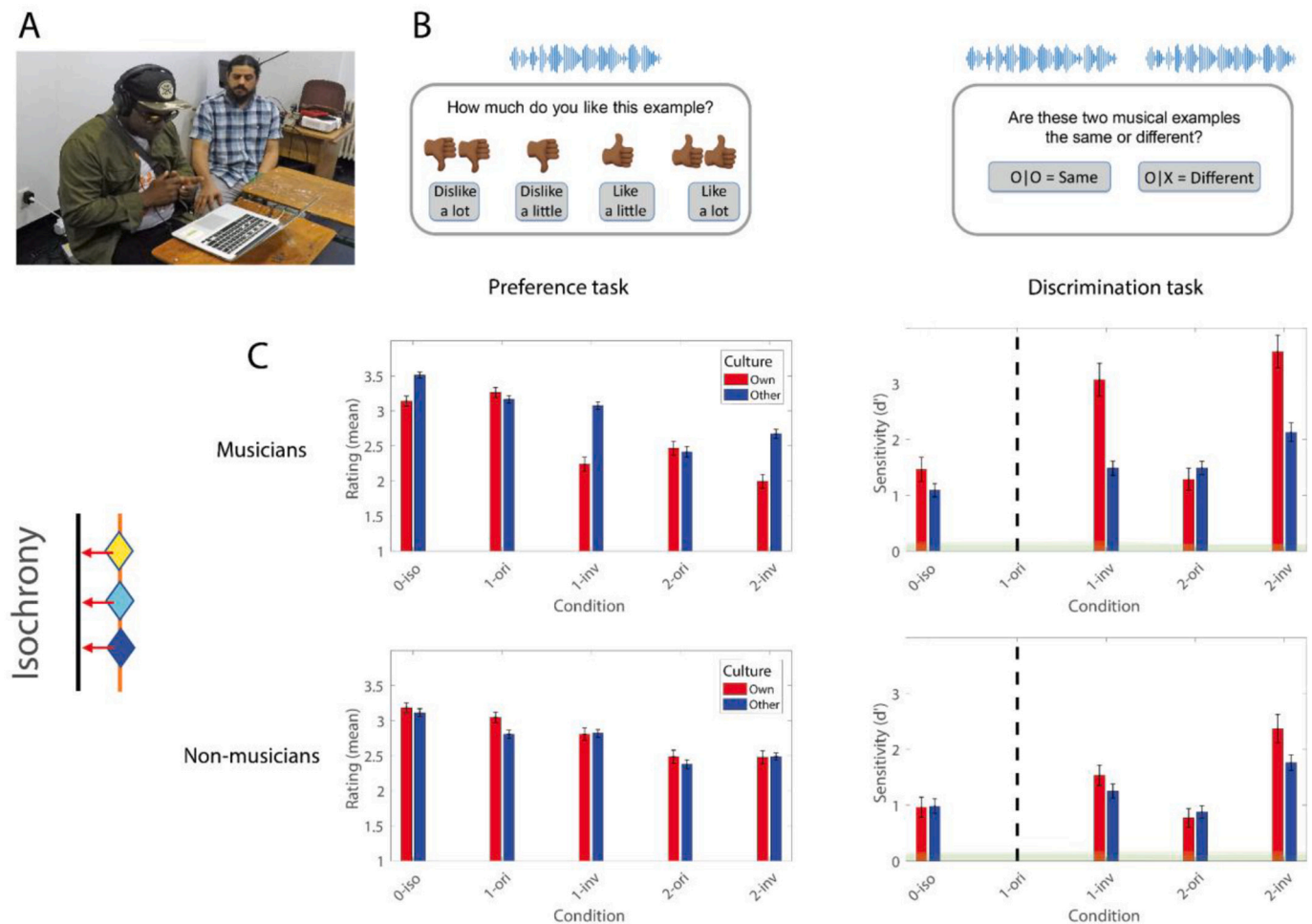


Fig. 4. Isochrony manipulation task and results. A. Procedure (photograph from a session in Uruguay). B. Task questions and rating scale. C. Results of the preference and discrimination tasks for the isochrony manipulation. Results are separated by task (left column: preference, right column: discrimination), expertise (upper tier: musicians, lower tier: non-musicians), and cultural familiarity (red: own culture, blue: other culture). Error bars represent one standard error of the mean across participants for the preference task and one standard deviation of the d' estimated via generating 1000 bootstrapped datasets with replacement for the discrimination task. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

style. Specifically, we collected data from 24 US jazz musicians based in New York City. The US musicians ranged in age from 19 to 67 years ($M = 37$, $SD = 14$; 22 male, 2 female) and had, on average, been engaging in regular practice on an instrument for 24 years ($SD = 11$). These participants completed only the preference task (for both synchrony and isochrony manipulations).

For the synchrony preference rating task, the US musicians showed a similar response pattern to the UK participants, as well as all other participant groups (see Fig. S2). For consistency with the main experiment analyses, we fit a cumulative link mixed model on the data from the US musicians, including main effects and interactions of Cultural Familiarity and Asynchrony Condition, with a random effect of Participant. We then extracted contrasts (with Bonferroni correction) from this model comparing preference ratings for the quantized (0-qua) to the original timing pattern (1-ori) by Cultural Familiarity. Similarly to the UK musicians (see Fig. 5A), the US musicians showed a significant preference for the quantized version over the original timing pattern for music from their own culture (jazz; $p < .001$), with no significant difference in preferences between these two versions for culturally unfamiliar music ($p > .99$).

For the isochrony manipulation, results for US musicians were also similar to the UK participants (see Fig. S4). An analogous cumulative link mixed model and contrast analysis was performed to that described above for the synchrony manipulation. When comparing the

isochronous (0-iso) to the original timing pattern (1-ori), the US musicians tended to prefer the isochronous version, although this difference was only statistically significant for the culturally unfamiliar music ($p < .001$) and not the culturally familiar (jazz) music ($p > .99$).

4. Discussion

The main aim of the present study was to test the degree to which aesthetic preferences for deviations from synchrony between instruments and isochrony of the beat subdivision in music are modulated by a listener's culture and expertise. Supplemental measures of pairwise discrimination were also taken to examine the convergence between subjective preference ratings and perceptual sensitivity.

We found similar patterns of (a)synchrony preferences across all groups, regardless of the listeners' levels of expertise and cultural familiarity with the music. Preference ratings consistently decreased with greater magnitudes of asynchrony, while the distribution of the asynchronies (e.g., which instrument played ahead or behind) did not systematically affect preferences. In addition, no Country $\times$ Expertise group showed a preference for the stimulus containing the asynchronies from the original performance over a quantized version, in either their own music or music of other cultures, and UK (and US) jazz musicians even showed a significant preference for the quantized jazz excerpt over the original timing. This suggests that the performers' deviations from

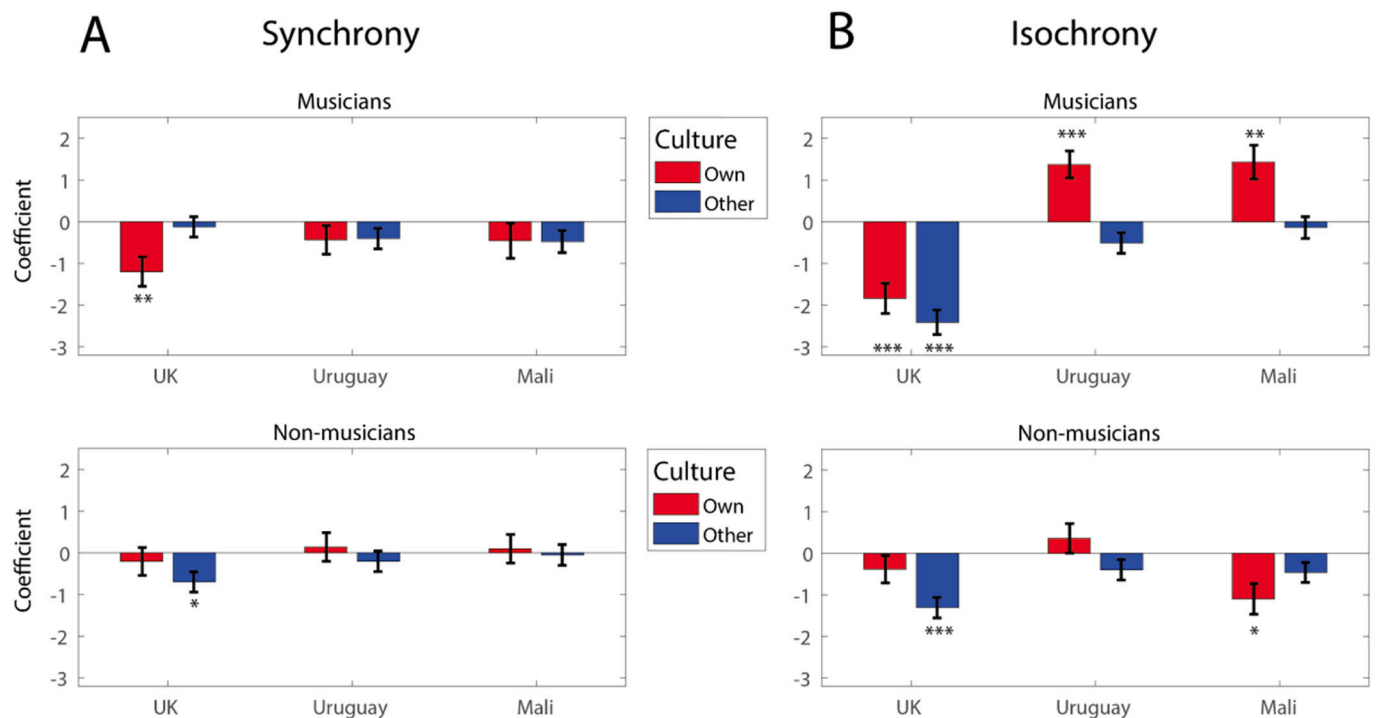


Fig. 5. Contrast analyses comparing preferences for fully synchronous/isochronous stimuli to original timings (1-ori), by country, expertise, and cultural familiarity. A. Bars display the coefficients from the pairwise contrasts comparing ratings for the original stimulus (1-ori) versus the quantized version (0-qua) for the synchrony manipulation. Positive coefficient values indicate a preference toward 1-ori and negative values indicate a preference toward 0-qua. Asterisks represent a significant difference between ratings of the two conditions (1-ori vs. 0-qua), with Bonferroni correction applied (* = $p < .05$, ** = $p < .01$, *** = $p < .001$). Error bars represent standard error of the mean. B. The analogous comparison for the isochrony manipulation (1-ori versus 0-iso). Positive coefficient values indicate a preference toward 1-ori and negative values indicate a preference toward 0-iso.

synchrony in the music styles used here were not perceived as expressive patterns of aesthetic relevance, even by those who were culturally familiar with, or expert performers in, a style. The most plausible explanation of this result is that the magnitude of the original asynchronies in our stimuli ($M = 13\text{--}19$ ms) is near or below the detectable threshold, as indicated by the finding that these asynchronies were barely noticeable in the discrimination task. This idea also aligns with previous findings that mean asynchronies of <30 ms in groove-based music were not perceptually salient (Butterfield, 2010), and when skilled drummers were asked to create a driving (“pushy”) or relaxed (“laid-back”) feeling in an experiment by playing before or after the beat, respectively, they used notably larger asynchronies, in the range of 25 to 40 ms (Cámara, Nymoen, Lartillot, & Danielsen, 2020).

Taken together, the relatively poor discrimination and aesthetic indifference toward the original asynchronies and the decreased preferences for larger asynchronies indicate an aesthetic ideal of perceptually perfect synchrony across all groups for all music styles tested here. This aligns with previous studies using Western music styles and participants, which have also found that performance-based asynchronies do not increase preferences or groove ratings over quantized versions (Datseris et al., 2019; Senn et al., 2016), and extends these findings to groove-based music styles and participants from other cultures. In our study, cultural familiarity with the music and musical expertise appeared to simply amplify the existing pattern of preferences for lower degrees of asynchrony (see Fig. 3C); this suggests the aesthetic ideal for synchrony in groove-based music is a relatively low-level, experience-independent construct that is strengthened by increased cultural familiarity and/or expertise with a style.

In the isochrony task, we found a more notable divergence in preferences as a function of culture and expertise. In general, preference ratings decreased with greater deviation from isochrony, however this

overall pattern of results varied depending on cultural familiarity and expertise. UK participants, in particular musicians, consistently preferred isochronous beat subdivision patterns for the music of their own culture, as well as music of other cultures. However, in Uruguay and Mali, expert musicians preferred the original, non-isochronous pattern for music from their own culture over an isochronous version. This preference pattern for non-isochrony did not extend to music from other cultures. Uruguayan and Malian non-musicians did not significantly prefer the original pattern over an isochronous version of their own music, despite the fact that they did have some exposure to this music (see Fig. S1) and were also able to recognize variants to the original pattern in the discrimination task. Furthermore, the participant subgroups in Uruguay and Mali generally did not show a significant preference for isochrony over the original pattern in any of the music styles (with one exception of Malian non-musicians listening to jembe music). These findings contrast the assumption that has been made in some previous literature that deviations from isochrony are perceptually and aesthetically irrelevant or disadvantageous (Merker, 2014), or that isochrony in music is a human “universal” (Ravignani & Madison, 2017; Savage et al., 2015). Rather, our findings support recent evidence that experience-based, listener-specific factors, such as familiarity and taste, shape our perception and appreciation of music (Madison & Schiöde, 2017), and rhythm and timing in particular (Danielsen et al., 2021; Senn et al., 2018; Senn, Bechtold, Hoesl, & Kilchenmann, 2021). This aligns closely with well-established findings on the role of exposure, familiarity, and prototypicality in shaping aesthetic preferences in general (Berlyne, 1970; Hekkert, Snelders, & van Wieringen, 2003; Hekkert & Wieringen, 1990; Zajonc, 1968).

The fact that this preference for non-isochrony in culturally familiar music was evidenced in both Uruguayan and Malian musicians leads to the question of whether similar preferences might be found in other

groups and musical styles. Comparative research at a more global level, using both stimuli and participants from other cultural contexts with different listening experiences, is needed to probe this question. Further research should also explore the reasons for the different pattern of results found in the UK in comparison to the Uruguayan and Malian groups. In the UK, both musicians and non-musicians tended to prefer isochrony in all styles (although in the non-musician group this result was only statistically significant for culturally unfamiliar music), suggesting that the interaction of cultural familiarity with expertise we found in both Uruguay and Mali does not seem to play a role here. It may be that the difference in musical exposure between university students and expert musicians in the UK is smaller than in Uruguay and Mali, where traditional musicians and university students (our non-musician groups) tend to form discrete social milieus and cultural sub-groups. Moreover, the overall preference for isochrony exhibited by the UK musicians (independent of the music's cultural familiarity) may be attributed to exposure/practice of other musical styles that favor isochrony, the use of recording and digital editing techniques that employ quantization, or other cultural influences such as the usage of music notation.

4.1. Limitations and future research

Several limitations should be noted for consideration in future research. Due to time constraints, our stimulus set for the preference ratings consisted of only a single, short, although representative, excerpt from each music style. We chose this approach because it allowed us to conduct an in-depth examination of the effects of both the magnitude and distribution of both asynchrony and non-isochrony on aesthetic preferences in a fully balanced, cross-cultural design. In addition, initial pilot testing of several of the manipulations using two excerpts to represent a musical style revealed very similar results across excerpts. However, it may be that different patterns of aesthetic preferences evolve over longer excerpts or whole performances, which would be difficult or impossible to test using the tasks and number of manipulations employed in the present study. A limitation more specific to the jazz style is the difficulty of defining a stylistic prototype to be represented by a single excerpt, due to the wide range of swing patterns and playing styles that can be adopted in this genre (Benadon, 2006; Dittmar et al., 2017; Friberg & Sundström, 2002). Future studies should aim to test a wider range of excerpts, as well as a more diverse sample of participants, as introduced here via the comparison of UK and US jazz musicians. In addition, though our stimuli were natural-sounding versions of real instrumental ensemble performances, we did not include other relevant ensemble parts such as singers or dancers, which might involve different types and degrees of asynchrony and non-isochrony. Finally, follow-up studies that extend our approach by also including a manipulation that combines deviations from both synchrony and isochrony would be particularly informative.

The degree of synchronization in our excerpts (mean asynchronies in the range of 13–19 ms), though typical of these music styles, can be described as relatively tight in comparison to other styles, such as art music traditions in both European and Asian countries (Clayton et al., 2020). It remains possible that other types of music with looser and more variable synchronization patterns (Clayton et al., 2020; Danielsen, 2010) may contain asynchronies that are not only more perceptually salient, but also more aesthetically relevant. For example, the interactions between the percussive accompaniment and singers in Japanese Noh songs show not only much rhythmic elasticity in general but also large-scale ensemble asynchronies, which are explicitly positively valued in corresponding aesthetic discourse (Fujita, 2019).

Finally, the present design does not enable us to fully disentangle the effects of expertise in performing a musical style from the effects of cultural familiarity with a musical style, as musicians also rated their familiarity with the music of their culture significantly higher than the non-musicians (see Fig. S1). These results therefore leave open the

question of whether the metrical patterns that are preferred by the musicians themselves (such as a non-isochronous beat subdivision) are actually conveyed to and preferred by their audiences. Subsequent research should use groups of avid listeners or participants with high levels of exposure that are not musicians (e.g., regular dancers, see Neuhoff et al., 2017), to further test whether familiarity in perceiving versus producing a particular rhythmic prototype leads to convergent results.

4.2. Methodological contributions and advantages of the present design

Previous cross-cultural studies of music cognition have often compared responses between two strongly contrasting cultural groups, typically one Western and one non-Western (e.g., Balkwill & Thompson, 1999; Eggermann, Fernando, Chuen, & McAdams, 2015; Fritz et al., 2009), or comprised large surveys of cultural materials and music recordings from a wide range of regions (e.g., Lomax, 1968; Mehr et al., 2019; Mehr, Singh, York, Glowacki, & Krasnow, 2018; Savage et al., 2015). Our approach falls somewhat in between these two approaches. We tested three musical cultures that are related, thus our design benefits from the comparability of both the musical materials as well as the familiarity of the participant groups with (one of) the styles comprising those materials. Since we chose three different musical styles with non-isochronous beat subdivisions, this allowed us to ask, for example, whether Malian musicians' preference for non-isochronous subdivisions in jembe music would generalize to their evaluations of other, relatively unfamiliar styles. This approach requires that the same feature is present in all the music styles under investigation (equivalence of stimuli and constructs, He & van de Vijver, 2012), and has the advantage that our results are likely to generalize to closely related musical styles, such as other African and African-diasporic traditions. However, this same design feature also limits the possibility of making broader claims about universality that extend to less related musical styles, which may exhibit entirely different rhythmic structures and performance patterns.

Another crucial consideration in designing this study was the choice between artificial and naturalistic stimuli. Artificial stimuli have the advantage that very specific perceptual features can be controlled with relative ease, whereas naturalistic stimuli introduce more complexity but are more ecologically valid. Here too, we took a middle-ground approach, which yields advantages from both sides, by creating stimuli that were based on information extracted from performed music, but were then re-synthesized from single sound samples, allowing for the isolation and manipulation of specific stimulus features. Our manipulations also went beyond earlier research by 1) making a clear differentiation between synchrony and isochrony and 2) varying not only the magnitude but also the temporal distribution of asynchrony/non-isochrony. These differentiations proved particularly informative, as greater preference variations were found for the isochrony and magnitude-related manipulations.

Two further methodological considerations are of note. First, we complemented our primary, subjective rating task (preferences) with perceptual measures (discrimination task), which generally supported the findings from the preference task, providing indications of the mechanisms underlying such preferences: stimuli that are more perceptually distinguishable elicit more defined preferences. Secondly, our recruitment of groups of expert musicians in each country, beyond the more common approach of recruiting university students, turned out to be particularly fruitful. In particular, the prominent cases of cultural variation we found concerned musicians in Uruguay and Mali, but not the non-musician (primarily university student) groups we tested in the same countries. This aligns with several previous studies showing that testing university student samples can underestimate the true cross-cultural variability of rhythm perception and production abilities (e.g., Jacoby, Polak, Grahn, et al., 2021; Yates et al., 2017).

5. Conclusion

In sum, we manipulated two types of rhythmic timing variations (synchrony and isochrony) in three musical styles, and presented these to six participant groups differentiated by country of residence and degree of musicianship, to test the effects of cultural familiarity and expertise on aesthetic preferences for such timing variations. Across all groups and styles, preferences increased for more synchronous stimuli, with greater cultural familiarity and expertise simply amplifying this pattern of results, suggesting a general preference for synchrony within groove-based music styles. On the other hand, a consistent preference for isochrony was found only in Western (UK) participants, and expert musicians in Uruguay and Mali preferred a non-isochronous subdivision pattern over an isochronous one in their own music, indicating that preferences for isochrony in music are more culturally contingent responses shaped by experience and exposure. These findings thereby resolve some previous conflicting views as to whether “microtiming” variations in music are aesthetically relevant, by demonstrating that the answer to this question is dependent on both the type of timing variation (synchrony/isochrony) and the experience of the listener.

More broadly, the divergent pattern of results that emerged for our synchrony and isochrony manipulations demonstrates that factors such as culture and expertise do not uniformly influence all aspects of the perception and aesthetic evaluation of expressive communication. That is, different components of the aesthetic experience may vary in their dependence on low-level perceptual versus experience-dependent factors. This highlights the potential of combining the systematic manipulation of constituent features of aesthetic objects with the recruitment of groups representing diverse levels of expertise and cultural background. As such, this study demonstrates the key role cross-cultural research can play in understanding aesthetic experiences.

Author note

This work was supported by the Arts and Humanities Research Council (grant no. AH/N00308X/1). Nori Jacoby was funded as a Presidential Scholar in Society and Neuroscience at Columbia University. Rainer Polak/data collection in Mali was funded by Max Planck Institute for Empirical Aesthetics. Luis Jure, Martín Rocamora, and data collection in Uruguay were partially funded by Universidad de la República, CSIC. Martín Rocamora was also partially funded by Agencia Nacional de Investigación (ANII), Uruguay. Thanks to Martin Clayton, Tuomas Eerola, and Simone Tarsitani (Durham University, UK), for assistance with the initial design and project funding. Thanks to Franco Toscano (Montevideo, Uruguay), Claudia Deluca (Montevideo, Uruguay), Sarah Nihad Nedjar (NYC, USA), Salabary Doumbia (Bamako, Mali), Noumouké Doumbia (Bamako, Mali), Madu Jakite (Bamako, Mali), Nikki Moran (University of Edinburgh, UK), and Matthias Lichtenfeld (Durham University, UK) for assistance with data collection. Thanks to Olivier Senn and Lorenz Kilchenmann (Lucerne University of Applied Sciences and Arts) for providing recordings and contextual materials from which we created the jazz stimuli. Associated stimuli and data for this article can be accessed at osf.io/uebdk. We have no conflicts of interest to disclose.

CRediT authorship contribution statement

Kelly Jakubowski: Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Project administration. **Rainer Polak:** Conceptualization, Methodology, Investigation, Resources, Writing – original draft, Visualization, Project administration. **Martín Rocamora:** Methodology, Investigation, Resources, Writing – review & editing. **Luis Jure:** Methodology, Investigation, Resources, Writing – review & editing. **Nori Jacoby:** Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Visualization, Supervision.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cognition.2022.105205>.

References

- Balkwill, L. L., & Thompson, W. F. (1999). A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music Perception*, 17(1), 43–64. <https://doi.org/10.2307/40285811>
- Benadon, F. (2006). Slicing the beat: Jazz eight-notes as expressive micro-timing. *Ethnomusicology*, 50, 73–98.
- Berlyne, D. E. (1970). Novelty, complexity, and hedonic value. *Perception & Psychophysics*, 8(5), 279–286. <https://doi.org/10.3758/BF03212593>
- Butterfield, M. W. (2010). Participatory discrepancies and the perception of beats in jazz. *Music Perception*, 27(3), 151–175.
- Cámara, G. S., & Danielsen, A. (2020). Groove (Hrsg.). In A. Rehding, & S. Rings (Eds.), *The Oxford handbook of critical concepts in music theory*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190454746.013.17>
- Cámara, G. S., Nymoen, K., Lartillot, O., & Danielsen, A. (2020). Timing is everything... or is it? Effects of instructed timing style, reference, and pattern on drum kit sound in groove-based performance. *Music Perception*, 38(1), 1–26. <https://doi.org/10.1525/mp.2020.38.1.1>
- Cameron, D. J., Zioga, I., Lindsen, J. P., Pearce, M. T., Wiggins, G. A., Potter, K., & Bhattacharya, J. (2019). Neural entrainment is associated with subjective groove and complexity for performed but not mechanical musical rhythms. *Experimental Brain Research*, 237(8), 1981–1991. <https://doi.org/10.1007/s00221-019-05557-4>
- Camurri, A., Mazzarino, B., Ricchetti, M., Timmers, R., & Volpe, G. (2003). Multimodal analysis of expressive gesture in music and dance performances. In A. Camurri, & G. Volpe (Eds.), *Gesture-based communication in human-computer interaction*. Springer. https://doi.org/10.1007/978-3-540-24598-8_3
- Chokron, S., & De Agostini, M. (2000). Reading habits influence aesthetic preference. *Cognitive Brain Research*, 10(1–2), 45–49. [https://doi.org/10.1016/S0926-6410\(00\)00021-5](https://doi.org/10.1016/S0926-6410(00)00021-5)
- Christensen, R. H. (2019). Ordinal: Regression models for ordinal data (R package version 2019.12-10). <https://CRAN.R-project.org/package=ordinal>
- Clarke, E. F. (1985). Structure and expression in rhythmic performance. In P. Howell, R. West, & I. Cross (Eds.), *Musical structure and cognition* (pp. 209–236). Academic Press.
- Clarke, E. F. (1993). Generativity, mimesis and the human body in music performance. *Contemporary Music Review*, 9(1–2), 207–219.
- Clayton, M., Jakubowski, K., Eerola, T., Keller, P. E., Camurri, A., Volpe, G., & Alborno, P. (2020). Interpersonal entrainment in music performance. *Music Perception*, 38(2), 136–194. <https://doi.org/10.1525/mp.2020.38.2.136>
- Danielsen, A. (2010). Here, there and everywhere: Three accounts of pulse in D'Angelo's 'Left and Right'. In A. Danielsen (Ed.), *Musical rhythm in the age of digital reproduction* (pp. 19–36). Ashgate.
- Danielsen, A., Nymoen, K., Langerød, M. T., Jacobsen, E., Johansson, M., & London, J. (2021). Sounds familiar(?): Expertise with specific musical genres modulates timing perception and micro-level synchronization to auditory stimuli. *Attention, Perception, & Psychophysics*. Advance online publication. <https://doi.org/10.3758/s13414-021-02393-z>
- Datseris, G., Ziereis, A., Albrecht, T., Hagmayer, Y., Priesemann, V., & Geisel, T. (2019). Microtiming deviations and swing feel in jazz. *Scientific Reports*, 9(1), 19824. <https://doi.org/10.1038/s41598-019-55981-3>
- Davies, M., Madison, G., Silva, P., & Gouyon, F. (2013). The effect of microtiming deviations on the perception of groove in short rhythms. *Music Perception*, 30(5), 497–510. <https://doi.org/10.1525/mp.2013.30.5.497>
- Dittmar, C., Pfeleiderer, M., Balke, S., & Müller, M. (2017). A swingogram representation for tracking micro-rhythmic variation in jazz performances. *Journal of New Music Research*, 47(2), 1–17. <https://doi.org/10.1080/09298215.2017.1367405>
- Egermann, H., Fernando, N., Chuen, L., & McAdams, S. (2015). Music induces universal emotion-related psychophysiological responses: Comparing Canadian listeners to Congolese pygmies. *Frontiers in Psychology*, 5, 1341. <https://doi.org/10.3389/fpsyg.2014.01341>
- Fabian, D., Timmers, R., Schubert, E., & (Eds.). (2014). *Expressiveness in music performance: Empirical approaches across styles and cultures (first edition.)*. Oxford University Press.
- Friberg, A., & Sundström, A. (2002). Swing ratios and ensemble timing in jazz performance: Evidence for a common rhythmic pattern. *Music Perception*, 19(3), 333–349. <https://doi.org/10.1525/mp.2002.19.3.333>
- Fritz, T. H., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., ... Koelsch, S. (2009). Universal recognition of three basic emotions in music. *Current Biology*, 19(7), 573–576. <https://doi.org/10.1016/j.cub.2009.02.058>
- Fujita, T. (2019). Layers and elasticity in the rhythm of Noh songs. In R. K. Wolf, S. Blum, & C. Hasty (Eds.), *Thought and play in musical rhythm: Asian, African, and Euro-American perspectives* (pp. 212–231). Oxford University Press. <https://doi.org/10.1093/oso/9780190841485.003.0009>
- Gabrielsson, A., & Juslin, P. N. (1996). Emotional expression in music performance: Between the performer's intention and the listener's experience. *Psychology of Music*, 24(1), 68–91.
- Hannon, E. E., Soley, G., & Ullal-Gupta, S. (2012). Familiarity overrides complexity in rhythm perception: A cross-cultural comparison of American and Turkish listeners.

- Journal of Experimental Psychology: Human Perception and Performance*, 38(3), 543–548. <https://doi.org/10.1037/a0027225>
- He, J., & van de Vijver, F. (2012). Bias and equivalence in cross-cultural research. *Online Readings in Psychology and Culture*, 2(2). <https://doi.org/10.9707/2307-0919.1111>
- Hekkert, P., & Leder, H. (2009). Product aesthetics. In P. Hekkert, & H. N. J. Schifferstein (Eds.), *Product experience* (pp. 259–285). Elsevier. <https://doi.org/10.1016/B978-008045089-6.50013-7>
- Hekkert, P., Snelders, D., & van Wieringen, P. C. W. (2003). “Most advanced, yet acceptable”: Typicality and novelty as joint predictors of aesthetic preference in industrial design. *British Journal of Psychology*, 94(1), 111–124. <https://doi.org/10.1348/000712603762842147>
- Hekkert, P., & Wieringen, P. C. W. (1990). Complexity and prototypicality as determinants of the appraisal of cubist paintings. *British Journal of Psychology*, 81(4), 483–495. <https://doi.org/10.1111/j.2044-8295.1990.tb02374.x>
- Hervé, M. (2022). RVAideMemoire-package: Testing and plotting procedures for biostatistics. R package version 0.9-81-2. <https://CRAN.R-project.org/package=RVAideMemoire>
- Hofmann, A., Wesolowski, B. C., & Goebel, W. (2017). The tight-interlocked rhythm section: Production and perception of synchronisation in jazz trio performance. *Journal of New Music Research*, 12(4), 1–13. <https://doi.org/10.1080/09298215.2017.1355394>
- Jacobsen, T. (2006). Bridging the arts and sciences: A framework for the psychology of aesthetics. *Leonardo*, 39(2), 155–162. <https://doi.org/10.1162/leon.2006.39.2.155>
- Jacoby, N., & McDermott, J. H. (2017). Integer ratio priors on musical rhythm revealed cross-culturally by iterated reproduction. *Current Biology*, 27(3), 359–370. <https://doi.org/10.1016/j.cub.2016.12.031>
- Jacoby, N., Polak, R., Grahn, J., Cameron, D. J., Lee, K. M., Godoy, R., ... McDermott, J. (2021). Universality and cross-cultural variation in mental representations of music revealed by global comparison of rhythm priors. *PsyArXiv*. <https://doi.org/10.31234/osf.io/b879v>
- Jacoby, N., Polak, R., & London, J. (2021). Extreme precision in rhythmic interaction is enabled by role-optimized sensorimotor coupling: Analysis and modelling of West African drum ensemble music. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences*, 376(1835), 20200331. <https://doi.org/10.1098/rstb.2020.0331>
- 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. <https://doi.org/10.1037/a0024208>
- Jure, L., & Rocamora, M. (2016, May). Microtiming in the rhythmic structure of Candombe drumming patterns. In *Fourth international conference on Analytical Approaches to World Music (AAWM 2016)*, New York, NY, United States.
- Jure, L., Rocamora, M., Tarsitani, S., & Clayton, M. (2020). IEMP Uruguayan Candombe. *Open Science Framework*. <https://doi.org/10.17605/OSF.IO/WFX7K>
- Keil, C. (1987). Participatory discrepancies and the power of music. *Cultural Anthropology*, 2(3), 275–283. <https://doi.org/10.1525/can.1987.2.3.02a00010>
- Kilchenmann, L., & Senn, O. (2015). Microtiming in swing and funk affects the body movement behavior of music expert listeners. *Frontiers in Psychology*, 6(903), 55. <https://doi.org/10.3389/fpsyg.2015.01232>
- Laukka, P., Eröla, T., Thingujam, N. S., Yamasaki, T., & Beller, G. (2013). Universal and culture-specific factors in the recognition and performance of musical affect expressions. *Emotion*, 13(3), 434–449. <https://doi.org/10.1037/a0031388>
- Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, 95(4), 489–508.
- Leder, H., Tinio, P. P. L., Brieber, D., Kröner, T., Jacobsen, T., & Rosenberg, R. (2018). Symmetry is not a universal law of beauty. *Empirical Studies of the Arts*, 37(1), 104–114. <https://doi.org/10.1177/0276237418777941>
- Lenth, R. V. (2022). *Emmeans: Estimated marginal means, aka least-squares means*. R package version 1.7.2. <https://CRAN.R-project.org/package=emmeans>
- Lisboa, T., Williamson, A., Zicari, M., & Eiholzer, H. (2005). Mastery through imitation: A preliminary study. *Musicae Scientiae*, 9(1), 75–110.
- Lomax, A. (1968). *Folk song style and culture*. American Association for the Advancement of Science.
- Madison, G., & Schiöde, G. (2017). Repeated listening increases the liking for music regardless of its complexity: Implications for the appreciation and aesthetics of music. *Frontiers in Neuroscience*, 11, 147. <https://doi.org/10.3389/fnins.2017.00147>
- Martindale, C. (1990). *The clockwork muse: The predictability of artistic change*. Basic Books.
- Masuda, T., Gonzalez, R., Kwan, L., & Nisbett, R. E. (2008). Culture and aesthetic preference: Comparing the attention to context of East Asians and Americans. *Personality and Social Psychology Bulletin*, 34(9), 1260–1275. <https://doi.org/10.1177/0146167208320555>
- Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, 44(2), 314–324. <https://doi.org/10.3758/s13428-011-0168-7>
- Matthews, T. E., Thibodeau, J. N. L., Gunther, B. P., & Penhune, V. B. (2016). The impact of instrument-specific musical training on rhythm perception and production. *Frontiers in Psychology*, 7, 69. <https://doi.org/10.3389/fpsyg.2016.00069>
- Mehr, S. A., Singh, M., Knox, D., Ketter, D. M., Pickens-Jones, D., Atwood, S., ... Glowacki, L. (2019). Universality and diversity in human song. *Science*, 366(6468). <https://doi.org/10.1126/science.aax0868>
- Mehr, S. A., Singh, M., York, H., Glowacki, L., & Krasnow, M. M. (2018). Form and function in human song. *Current Biology*, 28(3), 356–368. <https://doi.org/10.1016/j.cub.2017.12.042>
- Merker, B. (2014). Groove or swing as distributed rhythmic consonance: Introducing the groove matrix. *Frontiers in Human Neuroscience*, 8, 195. <https://doi.org/10.3389/fnhum.2014.00454>
- Neuhoff, H., Polak, R., & Fischinger, T. (2017). Perception and evaluation of timing patterns in drum ensemble music from Mali. *Music Perception*, 34(4), 438–451. <https://doi.org/10.1525/mp.2017.34.4.438>
- Polak, R., Jacoby, N., Fischinger, T., Goldberg, D., Holzapfel, A., & London, J. (2018). Rhythmic prototypes across cultures: A comparative study of tapping synchronization. *Music Perception*, 36(1), 1–23. <https://doi.org/10.1525/mp.2018.36.1.1>
- Polak, R., Jacoby, N., & London, J. (2016). Both isochronous and non-isochronous metrical subdivision afford precise and stable ensemble entrainment: A corpus study of Malian jembe drumming. *Frontiers in Neuroscience*, 10, 285. <https://doi.org/10.3389/fnins.2016.00285>
- Polak, R., & London, J. (2014). Timing and meter in Mande drumming from Mali. *Music Theory Online*, 20(1). <http://www.mtosmt.org/issues/mto.14.20.1/mto.14.20.1.polo-london.html>
- Polak, R., Tarsitani, S., & Clayton, M. (2018). IEMP Malian jembe. *Open Science Framework*. <https://doi.org/10.17605/OSF.IO/M652X>
- Prögler, J. A. (1995). Searching for swing: Participatory discrepancies in the jazz rhythm section. *Ethnomusicology*, 39(1), 21–54. <https://doi.org/10.2307/852199>
- Rasch, R. A. (1988). Timing and synchronization in ensemble performance. In J. A. Sloboda (Ed.), *Generative processes in music* (pp. 70–90). Clarendon Press.
- Ravignani, A., & Madison, G. (2017). The paradox of isochrony in the evolution of human rhythm. *Frontiers in Psychology*, 8, 1820. <https://doi.org/10.3389/fpsyg.2017.01820>
- Repp, B. H. (1990). Patterns of expressive timing in performances of a Beethoven minuet by nineteen famous pianists. *The Journal of the Acoustical Society of America*, 88, 622–641.
- Repp, B. H. (1997a). Expressive timing in a Debussy prelude: A comparison of student and expert pianists. *Musicae Scientiae*, 1(2), 257–268.
- Repp, B. H. (1997b). The aesthetic quality of a quantitatively average music performance: Two preliminary experiments. *Music Perception*, 14(4), 419–444. <https://doi.org/10.2307/40285732>
- Repp, B. H. (1998). Obligatory ‘expectations’ of expressive timing induced by perception of musical structure. *Psychological Research*, 61(1), 33–43.
- Repp, B. H. (2010). Sensorimotor synchronization and perception of timing: Effects of music training and task experience. *Human Movement Science*, 29(2), 200–213. <https://doi.org/10.1016/j.humov.2009.08.002>
- Rocamora, M. (2018). *Computational methods for percussion music analysis: The Afro-Uruguayan Candombe drumming as a case study* [Doctoral dissertation, Universidad de la República (Uruguay)]. <https://iie.fing.edu.uy/publicaciones/2018/Roc18/>
- Rocamora, M., Jure, L., Marengo, B., Fuentes, M., Lanzaro, F., & Gómez, A. (2015, November 17). An audio-visual database of Candombe performances for computational musical studies. In *Segundo Congreso Internacional de Ciencia y Tecnología Musical (CICTeM 2015)*, Buenos Aires, Argentina. <https://doi.org/10.5281/zenodo.3647666>
- Savage, P. E., Brown, S., Sakai, E., & Currie, T. E. (2015). Statistical universals reveal the structures and functions of human music. *Proceedings of the National Academy of Sciences*, 112(29), 8987–8992. <https://doi.org/10.1073/pnas.1414495112>
- Senn, O., Bechtold, T. A., Hoesl, F., & Kilchenmann, L. (2021). Taste and familiarity affect the experience of groove in popular music. *Musicae Scientiae*, 25(1), 45–66. <https://doi.org/10.1177/1029864919839172>
- Senn, O., Kilchenmann, L., Bechtold, T., & Hoesl, F. (2018). Groove in drum patterns as a function of both rhythmic properties and listeners’ attitudes. *PLoS One*, 13(6), Article e0199604. <https://doi.org/10.1371/journal.pone.0199604>
- Senn, O., Kilchenmann, L., von Georgi, R., & Bullerjahn, C. (2016). The effect of expert performance microtiming on listeners’ experience of groove in swing or funk music. *Frontiers in Psychology*, 7, 1487. <https://doi.org/10.3389/fpsyg.2016.01487>
- Senn, O., Rose, D., Bechtold, T., Kilchenmann, L., Hoesl, F., Jerjen, R., ... Alessandri, E. (2019). Preliminaries to a psychological model of musical groove. *Frontiers in Psychology*, 10, 1228. <https://doi.org/10.3389/fpsyg.2019.01228>
- Silvia, P. J., & Barona, C. M. (2009). Do people prefer curved objects? Angularity, expertise, and aesthetic preference. *Empirical Studies of the Arts*, 27(1), 25–42. <https://doi.org/10.2190/EM.27.1.b>
- Spring, H. (2014, January 31). *Swing*. Grove Music Online. <https://doi.org/10.1093/gmo/9781561592630.article.A2258226>
- Stamkou, E., van Kleef, G. A., & Homan, A. C. (2018). The art of influence: When and why deviant artists gain impact. *Journal of Personality and Social Psychology*, 115(2), 276–303. <https://doi.org/10.1037/pspi0000131>
- Thompson, W. F., & Balkwill, L.-L. (2012). Cross-cultural similarities and differences. In J. A. Sloboda (Ed.), *Handbook of music and emotion: Theory, research, applications* (pp. 755–788). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780192301433.003.0027>
- Van de Cruys, S., & Wagemans, J. (2011). Putting reward in art: A tentative prediction error account of visual art. *i-Perception*, 2(9), 1035–1062.
- Widmer, G., & Goebel, W. (2004). Computational models of expressive music performance: The state of the art. *Journal of New Music Research*, 33(3), 203–216.
- Wing, A. M. (1993). The uncertain motor system: Perspectives on the variability of movement. In D. E. Meyer, & S. Kornblum (Eds.), *Attention and performance 14: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 709–744). The MIT Press. <https://doi.org/10.7551/mitpress/1477.003.0043>
- Witek, M. A. G., Liu, J., Kuubertzie, J., Yankvera, A. P., Adzei, S., & Vuust, P. (2020). A critical cross-cultural study of sensorimotor and groove responses to syncope among Ghanaian and American university students and staff. *Music Perception*, 37(4), 278–297. <https://doi.org/10.1525/mp.2020.37.4.278>
- Wolf, A., Kopiez, R., Platz, F., Lin, H.-R., & Mütze, H. (2018). Tendency towards the average? The aesthetic evaluation of a quantitatively average music performance: A

- successful replication of Repp's (1997) study. *Music Perception*, 36(1), 98–108. <https://doi.org/10.1525/mp.2018.36.1.98>
- Yates, C. M., Justus, T., Atalay, N. B., Mert, N., & Trehub, S. E. (2017). Effects of musical training and culture on meter perception. *Psychology of Music*, 45(2), 231–245. <https://doi.org/10.1177/0305735616657407>
- Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, 9(2), 1–27. <https://doi.org/10.1037/h0025848>