

MICROTIMING AND MENTAL EFFORT: ONSET ASYNCHRONIES IN MUSICAL RHYTHM MODULATE PUPIL SIZE

JO FUGNER SKAANSAR, BRUNO LAENG, &
ANNE DANIELSEN
University of Oslo, Oslo, Norway

THE PRESENT STUDY TESTED TWO ASSUMPTIONS concerning the auditory processing of microtiming in musical grooves (i.e., repeating, movement-inducing rhythmic patterns): 1) Microtiming challenges the listener's internal framework of timing regularities, or meter, and demands cognitive effort. 2) Microtiming promotes a "groove" experience—a pleasant sense of wanting to move along with the music. Using professional jazz musicians and nonmusicians as participants, we hypothesized that microtiming asynchronies between bass and drums (varying from -80 to 80 ms) were related to a) an increase in "mental effort" (as indexed by pupillometry), and b) a decrease in the quality of sensorimotor synchronization (as indexed by reduced finger tapping stability). We found bass/drums-microtiming asynchronies to be positively related to pupil dilation and negatively related to tapping stability. In contrast, we found that steady time-keeping (presence of eighth note hi-hat in the grooves) decreased pupil size and increased tapping performance, though there were no conclusive differences in pupil response between musicians and nonmusicians. However, jazz musicians consistently tapped with higher stability than nonmusicians, reflecting an effect of rhythmic expertise. Except for the condition most closely resembling real music, participants preferred the on-the-grid grooves to displacements in microtiming and bass-succeeding-drums-conditions were preferred over the reverse.

Received: January 11, 2019, accepted August 22, 2019.

Key words: meter, microtiming, groove, pupillometry, effort

IN MUSIC, THE CONCEPT OF GROOVE CAPTURES three fundamental aspects of sound: rhythmic properties, embodiment, and pleasure (Cámara & Danielsen, 2018; Witek, 2017). First, grooves are musical

patterns that have a certain rhythmic, symmetrical, continuously repeating and danceable quality to them (Pressing, 2002). Groove-based genres include African-American and derived music heritages, such as jazz, soul, reggae, hip-hop, and funk (Pressing, 2002), as well as contemporary computer-programmed styles (e.g., electronic dance music, EDM; Butler, 2006). Second, groove is also a psychological construct, being a subjective sensorimotor response to the above types of music. Third, groove can be defined as "that aspect of the music that induces a pleasant sense of wanting to move along with the music" (Janata, Tomic, & Haberman, 2012, p. 56), which underlies a general association of groove to positive affect. The phenomenological state of feeling "in the groove" is assumed to be closely related to smooth and effortless musical attendance for the above-mentioned musical styles (Danielsen, 2006; Roholt, 2014).

A groove-based musical texture usually contains features that afford successful entrainment, including temporal information that increases rhythmic predictability. These features may include a "locomotion-friendly" tempo (Etani, Marui, Kawase, & Keller, 2018; Janata et al., 2012), dynamic repetition (Danielsen, 2006, 2019), and structural "low-level" features like event density and beat salience (Madison, Gouyon, Ullén, & Hörnström, 2011). However, it seems essential that certain forms of *complexity* are present to a certain extent so as to challenge the listener's perception and expectations, e.g., of a sensed rhythmic virtual reference structure (meter), as well as to create musical "tension" (Huron, 2006, p. 305). For example, researchers have debated the role of syncopation density (Sioros, Miron, Davies, Gouyon, & Madison, 2014), polyrhythm (Danielsen, 2006; Vuust, Gebauer, & Witek, 2014), as well as microtiming, which is the focus of the present study (Butterfield, 2010; Danielsen, Haugen, & Jensenius, 2015; Iyer, 2002).

Specifically, microtiming refers to subtle timing asynchronies. These are typically applied systematically and intentionally (although not always consciously) for expressive purposes, throughout a variety of musical genres and contexts (Clarke, 1989; Collier & Collier, 1996; Danielsen, Haugen, et al., 2015; Friberg & Sundström, 2002; Iyer, 2002; Keil, 1987; Palmer, 1996; Pressing, 2002; Rasch, 1988). As examples, the musician or

ensemble may play at times rhythmically “outside” a presumed norm (most commonly, nearly isochronous beat series/metric grid at different levels), fine-grained asynchronies between the onsets of the players’ tones may occur, or there may be subtle tempo changes. Notes containing microtiming will likely stand out perceptually from notes placed on the rhythmic grid and will tend to attract the listener’s attention, so that the perceptual system treats them as something worthy of closer analysis (Iyer, 2002). Musically, then, the presence of microtiming may help achieve a greater perceptual salience for some structural aspects of the music, like melodic material or even single tones (Iyer, 2002; Palmer, 1996; Repp, 1996). Groove players’ discourse involves discussing microrhythmic nuances like “pushing,” “pulling,” or playing “laid back” or “on the beat,” often with specific reference to asynchronies between the bassist and the drummer (see, for example, Collier & Collier, 1996). Indeed, the more skilled groove musicians are able of playing with a stable microtiming profile over time (Danielsen, Waadeland, Sundt, & Witek, 2015). Microtiming configurations in the rhythm section are also well documented from ethnographic studies (Monson, 1996) and analyses of musical recordings. For example, in several songs by the R&B artist D’Angelo, systematic timing asynchronies between the drums and electric bass have been found to reach up to 90 ms (Danielsen, 2010; Danielsen, Haugen, et al., 2015).

In groove contexts, a proposed function of rhythmic complexity generally—and of microtiming specifically—is to create patterns of tension and equilibrium in relation to the listeners’ anticipations of a *metrical* pulse (Roholt, 2014). The influential Dynamic Attending Theory (DAT; Jones, 2016; Large & Jones, 1999; Large & Snyder, 2009) provides one plausible theoretical background for such a proposal (see also Danielsen, 2019; London, 2012). With support from recent neuroscientific evidence, DAT addresses the relations between rhythmic input and temporal cognitive, that is, metric, structures. The general notion is that when external rhythmic events are attended, neural population *oscillations* are set into action, forming rhythmic or repetitive neural activity (Fujioka, Trainor, Large, & Ross, 2009; Large, Herrera, & Velasco, 2015; Lehmann, Arias, & Schönwiesner, 2016; Nozaradan, Peretz, Missal, & Mouraux, 2011; Zoefel, ten Oever, & Sack, 2018). Importantly, these neural oscillations, dubbed the *attending rhythm*, correspond to multiple time levels of the rhythm heard—or what is imagined within a metric framework (Iversen, Repp, & Patel, 2009). Furthermore, being self-persistent, these oscillations “expect” stability and engender a pattern of temporal expectations. Thus,

their cyclic nature entails anticipations concerning the placement of the next pulse beat and, hence, cues to the temporal allocation of attentional energy (Jones & Boltz, 1989) and coordination of overt movement with the music (Danielsen, Haugen, et al., 2015).

According to DAT, expectancy violations, as those resulting from microtiming asynchronies between bass and drums, are continuously taken into account by the neural networks. Phase perturbations caused by microtiming lead to a widening of the attentional focus for each perceived beat in order to encompass deviations or multiple onsets, sometimes with the consequence that the phase of the oscillation is adjusted (Danielsen, Haugen, et al., 2015). Importantly, although not stated explicitly by the above theorists, we may also assume that these adjustments demand considerable attentional and cognitive resources in the listener’s brain.

Moreover, it has recently been theorized (Keil & Feld, 1994; Witek, 2017), with some supportive empirical evidence (Witek, Clarke, Wallentin, Kringelbach, & Vuust, 2014), that rhythmic complexities (like microtiming, syncope, and polyrhythm) may contribute to groove by virtue of inviting the participant to “participate” with bodily movements. Note that such rhythmic complexities generate “tension” between the rhythm and the underlying meter. Such an experience could even allow the listeners’ filling in of the open spaces, or metrical ambiguities in the groove, with motion of their own bodies (e.g., tapping a foot, head nodding; Witek, 2017). It is paramount, however, that for a rhythm to be groove-promoting, its rhythmic complexity challenges but not disrupts the listeners’ metrical model (Vuust, Dietz, Witek, & Kringelbach, 2018). Interestingly, Witek and colleagues (2014) found an inverted U-shaped curve between syncopation density in drum kit grooves and groove experience. That is, either too little or too much syncopation (complexity) yielded lower groove experience, while medium syncopation positively related to the highest groove rating. These findings suggest a generally nonlinear relation between structural complexity and aesthetic pleasure in artistic objects (as also originally pointed out by Berlyne, 1971). However, other systematic experimental investigations attempting to establish a causal link between presence of the rhythmic complexity of *microtiming* and groove experience have yielded inconsistent results. Some studies have found that listeners generally rate rhythm patterns as groovier when the patterns are on-the-grid (quantized) than when they contain microtiming configurations (Butterfield, 2010; Davies, Madison, Silva, & Gouyon, 2013; Madison et al., 2011). Others have found that in some musical instances, microtiming may

achieve as high (but not higher) groove rating as full synchronization (Matsushita & Nomura, 2016; Senn, Kilchenmann, von Georgi, & Bullerjahn, 2016).

Despite the important role of rhythmic complexity in groove-based music, its relation to the psychological variable of *mental effort* has not yet been investigated. Mental effort refers to the intensity of processing in the brain or cognitive system (Just, Carpenter, & Miyake, 2003) and, as originally suggested by Kahneman (1973), it can be measured reliably by phasic pupil size diameter changes by use of the psychophysiological method of *pupillometry*. Currently, measuring the pupil and therefore indexing physiologically the moment-by-moment mental effort can be easily accomplished on the basis of computerized eye-tracking technology (Alnæs et al., 2014; Laeng, Sirois, & Gredeback, 2012). When luminance is kept constant during the experiment, increases in pupil size, or dilation over a baseline, can be taken as a gauge of an increase in mental (cognitive) workload (Kahneman & Beatty, 1966, 1967). Note that mental effort is not another term synonymous to performance level but is concerned with the underlying difference in allocation of cognitive capacity (Walle, Nordvik, Espe-seth, Becker, & Laeng, 2019). Indeed, the same task may result in different levels of effort from different individuals, in turn reflecting individual differences in ability (Ahern & Beatty, 1979), cognitive resources (Alnæs et al., 2014), affective and hedonic states (Bradley, Miccoli, Escrig, & Lang, 2008; Libby, Lacey, & Lacey, 1973), or other temporary psychological and brain states (Lean & Shan, 2012; McGinley, David, & McCormick, 2015), including actual physical effort (e.g., Zénon, Sidibé & Olivier, 2014). The pupil size has also been shown to index interest and affective processing during music listening (Hammerschmidt & Wöllner, 2018; Laeng, Eidet, Sulutvedt, & Panksepp, 2016; Partala & Surakka, 2003), as well as for quantifying listening effort in speech comprehension (Winn, Wendt, Koelewijn, & Kuchinsky, 2018). The eye pupils dilate to several cognitive mechanisms like surprise and violation of expectation (Friedman, Hakerem, Sutton, & Fleiss, 1973; Preuschoff, Hart, & Einhäuser, 2011; Quirins et al., 2018; this also in the domain of music: Damsma & van Rijn, 2017; Liao, Yoneya, Kashino, & Furukawa, 2018), as well as cognitive conflict or interference (Laeng, Ørbo, Holmlund, & Miozzo, 2011), working memory load (Granholm, Asarnow, Sarkin, & Dykes, 1996; Kahneman & Beatty, 1966; Schwalm, Keinath, & Zimmer, 2008), and perceptual/attentional shifts (Einhäuser, Stout, Koch, & Carter, 2008).

In the present study, our goal was to investigate the effects on listeners' cognitive processing of one type of

rhythmic complexity in musical groove: microtiming asynchronies. This appears relevant for at least three reasons: First and foremost, we are able, in a novel fashion, to put directly to test a prediction derived from Dynamic Attending Theory, namely that a) microtiming in musical grooves contributes to increased rhythmic complexity in a fashion that challenges listeners' fundamental cognitive structures for musical timing (i.e., the listeners' internal metric framework and associated moment-to-moment anticipations), and that b) this added complexity in the auditory input requires mental effort.

Second, by combining psychophysiological effort measurement with behavioral data, such as tapping variability, and ratings of subjective experience of groove, we are able to explore possible systematic relations between effort and groove that, in turn, may be moderated by rhythmic structural features. This seems particularly relevant since groove-experiences appear to stem from predictability (effortless attending) of the sound stream, but also from its complexity (effortful attending).

Third, by comparing groups of musicians with non-musicians, the study may also document effects of musical expertise. Individual differences in music training may result in noticeable differences in the processing efficiency of microtiming events, since some of these are posited to be cognitively demanding. Although some previous experimental findings suggested that elementary (Western) meter perception is independent of musical sophistication (Bouwer, Van Zuijlen, & Honing, 2014; Damsma & van Rijn, 2017), others suggest that musicianship and expertise develop cognitive metrical frameworks that can enhance the perception of rhythm (Chen, Penhune, & Zatorre, 2008; Drake, Penel, & Bigand, 2000; Geiser, Sandmann, Jäncke, & Meyer, 2010; Matthews, Thibodeau, Gunther, & Penhune, 2016; Stupacher et al., 2013; Stupacher, Wood, & Witte, 2017; Vuust et al., 2005). When it comes to more sophisticated rhythmic structure (at least according to Western standards), musicians have indeed been found to be more sensitive than nonmusicians in perceiving and/or superior in synchronizing to complex time signatures (Snyder, Hannon, Large, & Christiansen, 2006), musical syncopations (Ladinig, Honing, Haden, & Winkler, 2009), polyrhythmic musical texture (Jones, Jagacinski, Yee, Floyd, & Klapp, 1995), and onset (microtiming) asynchronies (Hove, Keller, & Krumhansl, 2007).

Previous psychophysiological studies—related to rhythm and meter perception generally—have often used electroencephalography (EEG; Näätänen, Paavilainen, Rinne, & Alho, 2007) and, specifically, the

mismatch-negativity (MMN) component of event related potentials (ERPs). Magnitude and latency of early EEG-signals appear to mirror the mismatch between expectation and what is heard (e.g., Honing, 2012; Vuust et al., 2005), so that this response has been interpreted as reflecting the degree of expectation violation between rhythm and underlying meter. Interestingly, similar to the MMN component, pupil dilations can also signal metric violation/“surprise” effects to single rhythmic deviants (Damsma & van Rijn, 2017; Fink, Hurley, Geng, & Janata, 2018). However, an alternative approach is to monitor time-averaged pupillary diameters while participants listen continuously to running rhythms with varying degrees of complexity, but without time-locked, single deviants. Moreover, monitoring pupillary responses can provide a more direct measure of attentional demands (e.g., mental effort) than electroencephalography (e.g., Rozado & Dunser, 2015) or other psychophysiological measures, like changes in heart rate or skin resistance (e.g., Kahneman, Tursky, Shapiro, & Crider, 1969; Libby et al., 1973). Thus, the present pupillometry paradigm, where participants can adjust their internal metrical model in accordance with the incoming rhythmic texture, may be able to capture a different process than the MMN-studies mentioned above. Consistent with the DAT account, the pupil could elucidate dynamic attention as it encompasses both the initial response (e.g., prediction error) *and* the following adjustment of the metric framework to “fit” the incoming rhythm.

In line with assumptions from DAT, rhythmic complexity in the form of microtiming would be expected to be reflected in behavioral measures, like the quality of sensorimotor synchronization (e.g., “the coordination of rhythmic movement with an external rhythm”; Repp & Su, 2013, p. 403). Specifically, some studies have shown that fluctuations in intra-individual tapping accuracy can reflect the variation of external rhythmic complexity (e.g., Chen et al., 2008), while inter-individual tapping accuracy differences may reflect variation in musical/rhythmic expertise (e.g., Hove et al., 2007). Indeed, finger-tapping paradigms remain popular in investigating the above mechanisms in both music psychology and the neuropsychology of motor control (e.g., Dhamala et al., 2003; Jäncke et al., 2000; Repp, 2005; Repp & Su, 2013); consequently, we included a measure of tapping accuracy in the present study.

In our experimental paradigm, we exposed participants to the auditory stimuli of short (30 s) groove patterns played by double bass and drum kit and systematically changed the asynchrony magnitude between the double bass and the drums into five distinct

microtiming conditions. We included three different groove patterns with increasing levels (low, medium, high) of syncopation density and note onsets per bar. Within the low and medium levels, we compared the same grooves with and without hi-hat eighth notes, to directly measure an effect of “timekeeping,” that is, a layer of events with a faster “density referent” supporting the metric reference (i.e., hi-hat eighth notes; see Nketia, 1974, p. 127). Importantly, we presented the auditory stimuli in both a passive condition (“listening only”) and an active tapping condition (“synchronizing with the beat”). The motivation for this experimental manipulation was to make sure that pupil responses do not simply reflect the added effort of performing an action (tapping) or motor recruitment but that they are also driven by attentional processes *per se* (see Laeng et al., 2016).

The dependent variables in the present study consisted of three complementary data sources (Jack & Roepstorff, 2002), covering several aspects of pre-attentive and conscious processing: 1) levels of mental effort or cognitive workload, as measured by baseline-corrected pupil diameter changes; 2) quality of sensorimotor synchronization (i.e., tapping accuracy or variability), operationalized as the standard deviation (in ms) of tapping offset from rhythmic reference points; and 3) subjective ratings of groove on a Likert scale (i.e., degree of “wanting to move the body to the music” and “musical well-formedness”).

The main hypothesis was that the cognitive processing of microtiming in these groove contexts would increase pupil dilation while decreasing tapping accuracy, compared to when such rhythmic challenges are not present. In accordance with extensive empirical research documenting stronger and better-refined metric models for musicians than nonmusicians, we expected consistently higher tapping accuracy in the former than in the latter group. Regarding possible pupillary effects of rhythmic expertise, on the basis of the MMN-studies referred above, we also expected *stronger* responses to rhythmic deviants (e.g., microtiming) in musicians than nonmusicians, which would be accounted for by the enhanced sensitivity of musicians to musical incongruity (Brattico et al., 2008; van Zuijlen, Sussman, Winkler, Näätänen, & Tervaniemi, 2005; Vuust, Ostergaard, Pallesen, Bailey, & Roepstorff, 2009). Hence, one could expect that professional musicians would pay more attention to the pulse, which in turn could be mirrored in greater pupillary responses in musicians. However, expertise has also been *negatively* related to mental effort both in musical tasks (Bianchi, Santurette, Wendt, & Dau, 2016) and other domains, like mathematical reasoning (Ahern & Beatty, 1979) or face recognition memory (Wu, Laeng, &

Magnussen, 2012), indicating more efficient processing in experts. This aspect of efficiency may supersede or counteract an increase in attentional allocation. Thus, we may also observe a greater pupillary response in the nonmusician group compared to professional musicians. A third possibility is that these opposite processes are at work simultaneously and neutralize each other, leading to inconclusive results in relation to expertise-based differences. Hence, it is not straightforward to predict pupillary changes based on differences in expertise.

Finally, in line with earlier research, we expected that on-the-grid playing, or the very moderate microtiming versions, would be rated by all participants as highest on subjective measures of groove and musical well-formedness. Nevertheless, we expected a group difference also on these ratings, as musicians' more robust metric models might be more sensitive and perhaps tolerant to subtle onset asynchronies in timing (Hove et al., 2007).

Method

PARTICIPANTS

We recruited 63 volunteers (i.e., unpaid participants) by personal invitation or word-of-mouth. Twenty of the participants (eight females) were professional jazz musicians (mean age = 29.5, range = 20–44), and another 43 (twenty females) were nonmusicians/amateur musicians (mean age = 30.0, range = 20–51). The two groups' age distributions did not differ (as revealed by a two-sample Kolmogorov-Smirnov test, $p = .99$), nor did their gender distribution (as revealed by a chi-square test of independence; $X^2(1, N = 63) = .23, p = .63$). The average years of musical instrument experience for the musicians were 19.60 ($SD = 6.01$), and for the nonmusicians 3.98 ($SD = 7.81$). All the professional musicians had obtained a university degree in music and they verbally described themselves as professionals in groove-based genres (e.g., jazz, pop, soul). The group comprised four pianists, four guitarists, one bass player, two drummers, two vocalists, three sax players, three trumpeters, and one trombonist. Among the nonmusicians, three participants played the piano at amateur level, three the guitar, one the bass guitar, two the trumpet, and one played the flute. No other demographic data were registered. All participants had (by self-report) normal or corrected-to-normal (with contact lenses) vision as well as hearing ability within the normal range. Participants signed a written informed consent before participation and were treated in accordance with the Declaration of Helsinki. The study was approved by the Internal Research Board or Ethics

Committee of the Department of Psychology at the University of Oslo (No. 1417777). All 63 participants were able to complete the experiment and were included in the final analyses of the pupil and subjective rating data. However, the entire *tapping* data sets of ten participants (all nonmusicians) did not meet the inclusion criteria for tapping data (as clarified below) and these were excluded prior to performing the statistical analyses. Following these participants' exclusions, there were still no significant age (as revealed by a two-sample Kolmogorov-Smirnov test, $p = .96$), nor gender differences (as revealed by a chi-square test of independence; $X^2(1, N = 53) = .70, p = .79$) between the remaining nonmusicians ($N = 33$) and the group of musicians ($N = 20$).

MUSICAL EAR TEST (MET)

To measure objectively the musical expertise of the two groups (musicians and nonmusicians), all participants went through the Musical Ear Test (MET; Wallentin, Nielsen, Friis-Olivarius, Vuust, & Vuust, 2010). MET has a melodic and a rhythmic part (but only the rhythmic part was used in the present study). Despite MET not addressing microtiming specifically, it measures general musical/rhythmic competence through the proxy of musical working memory. The test appears to be able to discriminate individuals' sensitivity to auditory fine-scaled rhythms (Wallentin et al., 2010), which fits the goal of the present study. MET has shown good psychometric attributes of validity and reliability, with practically neither ceiling- nor floor-effects. Moreover, MET can successfully distinguish between groups of professional musicians from amateur musicians and nonmusicians (e.g., Wallentin et al., 2010).

Each participant's percentage accuracy score in the MET was entered as the dependent variable in a one-way ANOVA using JASP (v.0.9) software. This analysis revealed a significant effect of Musicianship (musicians vs. nonmusicians), $F(1, 61) = 23.22, p < .001$. As expected, musicians ($M = 87.21; SD = 5.04$) outperformed the nonmusicians ($M = 77.64; SD = 8.17$) on this test.

AUDITORY STIMULI

We generated 25 novel groove excerpts lasting 30 s each (Figure 1). Professional jazz musicians on a standard drum kit and double (upright) bass recorded all excerpts acoustically. Drums and bass constitute the typical rhythm section in many contemporary musical genres, hence, most people should be used to pay "rhythmic attention" to these instruments. This is also a typical constellation where microtiming asynchronies

The figure displays five musical conditions (A-E) for a groove, each consisting of an Upright Bass line and a Drums line. The tempo is marked as $\text{♩} = 80$. Condition A is labeled 'Uniformly shifted' and shows a simple bass line with quarter notes and a drum pattern with a metronome click and a kick drum. Conditions B, C, D, and E show increasingly complex bass lines and drum patterns, including hi-hats and snare (rim) drums. Condition E is the most complex, featuring a dense bass line and a complex drum pattern with hi-hats and snare (rim) drums. The microtiming conditions are indicated by a vertical axis on the left of each condition: -80 ms, -40 ms, 0 ms, +40 ms, and +80 ms.

FIGURE 1. A) Low structural complexity groove without hi-hat (“low”); B) Low structural complexity groove with hi-hat (“lowHH”); C) Medium structural complexity groove without hi-hat (“medium”); D) Medium structural complexity groove with hi-hat (“mediumHH”). E) High structural complexity groove (“highHH”). Nota bene: there was no “high” condition, since the high complexity groove was presented with hi-hat only. All five groove-excerpts were presented in five microtiming conditions (relative placement of bass compared to drums): -80 ms, -40 ms, 0 ms (or synchronous), $+40$ ms, $+80$ ms. Figure setup adapted from Matsushita and Nomura (2016).

TABLE 1. Overview of Experimental Stimuli

| | | Timekeeping (eighth note hi-hat): | |
|--|--------|-----------------------------------|-------------------------------|
| | | No | Yes |
| Complexity (number of syncopations/ note events per bar ex. hi-hat): | Low | <i>Figure 1A</i> (“low”) | <i>Figure 1B</i> (“lowHH”) |
| | Medium | <i>Figure 1C</i> (“medium”) | <i>Figure 1D</i> (“mediumHH”) |
| | High | * | <i>Figure 1E</i> (“highHH”) |

Note: *There was no “high” condition, since the high complexity groove was presented with hi-hat only. All five groove-excerpts were presented in five microtiming conditions (relative placement of bass compared to drums): -80 ms, -40 ms, 0 ms (or synchronous), $+40$ ms, $+80$ ms.

often occur in groove contexts (Keil, 1987). By only presenting bass and drum, both playing a key role in timing perception, we simultaneously removed possible confounders of attending to other aspect of the musical piece.

The recorded groove excerpts were categorized into three levels (low, medium, high) of structural Complexity, which in this context refers to syncopation density as well as the number of note onsets per bar (i.e., these two variables were confounded in the present stimuli since number of note onsets per bar varied independently with number of eighth and sixteenth note syncopations per bar in the “low,” “medium,” and “high” complexity conditions). Syncopation is a common measure of structural complexity in a groove (Witek et al., 2014). Specifically, the “low” complexity excerpts (Figure 1A & B) featured two notes per bar in the bass and no syncopations; the “medium” level (Figures 1C & D) featured three bass notes per bar, in which two of these were syncopated, and no drum syncopations. To be able to directly examine the processing effect of a steady timekeeper, the “low” and “medium” levels also featured versions with (Figures 1B & D) and without (Figures 1A & C) eighth notes hi-hat (“HH”). Generally, there is no obvious connection between an increase in event density caused by steady timekeeping and structural complexity. Rather, as long the acoustical articulations of the time-keeping instrument support rather than challenge the listener’s metric reference, timekeeping will likely simplify the processing of the rhythmic material.

The “high” Complexity condition (Figure 1E) was recorded in a single version only, which included both hi-hat and snare drum and featured a more intricate drum kit pattern with two sixteenth note syncopations per bar in the kick drum. The bass line was melodic and non-syncopated, “walking bass like,” with eight eighth notes per bar. The structure was eight bars long, in contrast to the low and medium complexity-grooves that had a repeated two-bar structure. The reason for not including a version of the high complexity

condition without hi-hat was that this groove was inspired by a real musical example, namely the R&B/soul-tune *Really Love* (Mayfield, D’Angelo, Figueroa & Foster, 2014) by the artist D’Angelo, where microtiming is a part of the groove matrix between bass and drums. An analysis (performed with Amadeus Pro, v.2.3.1) scrutinising the microtiming asynchrony between bass and drums in the *original* track of *Really Love*, showed that the timing of the electric bass is approximately 40-60 ms *behind* the kick drum timing in most of the tune. This feature of the song also enabled us to compare the effects from an “ecological” groove with the two “custom made” grooves. Indeed, ecological validity represents a challenge in the field of music cognition generally (Demorest, 1995). We surmise that the effect microtiming has on its listeners, depends on the musical context into which it is experienced. Hence, we found it pertinent to supplement the present study with adding an “ecological” stimulus.

To produce the stimuli (see Table 1), we proceeded as follows: Bass and drum tracks were recorded separately. Drums were recorded using one microphone on each drum (kick, hi-hat, snare drum). For the double bass recording, we used two microphones placed close to the bridge/F hole, as well as a Realist contact microphone. All excerpts were 16 bars (in 4/4 tactus) long. The bassist and drummer were instructed to play “on the grid” of a metronome click track set to 80 bpm. Eighty bpm is an adequate tempo for groove-based music (especially in the R&B/soul-genre; Danielsen, Haugen, et al., 2015), and it is the actual tempo of *Really Love* that served as the model for our high complexity groove. The notes played by the bass player and drummer varied in duration and loudness in a natural manner throughout the recordings. The drums (i.e., kick, hi-hat, and snare drum) were aligned to the metronomic grid’s eighth notes (and to the swung sixteenth notes in the “highHH”-excerpt) post-recording using the quantization tool in Cubase DAW (v.8.5). To keep the stimuli as musical and natural as possible, the bass part was not quantized post-recording; this yielded minute non-

TABLE 2. Onset Asynchronies of Recorded Instrument vs. Metronome Reference in On-the-grid Versions of Experimental Stimuli

| Complexity | Onsets | Mean (ms) | SD (ms) | Min (ms) | Max (ms) |
|------------------------------------|--------|-----------|---------|----------|----------|
| Drums (quantized) | | | | | |
| low & medium | 17 | 2 | 0 | 2 | 2 |
| lowHH & mediumHH | 83 | 2 | 0 | 2 | 2 |
| highHH | 127 | 2 | 0 | 2 | 2 |
| Double bass (non-quantized) | | | | | |
| low & lowHH | 17 | 12.9 | 2.6 | 5 | 17 |
| medium & mediumHH | 25 | 3.6 | 9.2 | -17 | 21 |
| highHH | 64 | 10.2 | 10.2 | -7 | 41 |

Note: Negative numbers indicate that the recorded instrument is ahead of the metronomic reference.

systematic microtiming of the bass in relation to drums. The onset of each bass event in relation to the grid was identified using LARA computer software (v.2.6.3). First, we verified that there was adequate correspondence between onsets in the waveform (first-zero-crossing) and LARA's "perceptual onset" measurement. Second, a manual inspection of all LARA results was performed and those bass notes that were not identified by LARA or seemed incorrect according to the waveform's shape, were manually measured (Amadeus Pro, v.2.3.1) and included in the analysis. The timing analysis of the experiment sound stimuli is summarized in Table 2.

Next, corresponding bass and drums tracks were combined in AVID Pro Tools (v.10) (with help from technician MTL). Five microtiming levels (-80, -40, 0, +40, +80) of relative timing (in ms) of bass in relation to the original bass recordings were produced for each of the five complexity levels (low, lowHH, medium, mediumHH, and highHH). The microtiming conditions were produced by placing the click metronomes of the bass and drum track in the correct relationship for each of the five microtiming conditions in the digital audio workstation. Consequently, the entire bass track was before (-80, -40), aligned (0), or after (+40, +80) the drum track/metronome reference, which means that all levels differed by 40 ms. For simplicity, we use the term "0 ms" for the originally recorded condition, even though the non-quantized bass was not timed exactly on the metronomic grid. Each trial began with three metronome ticks aligned with the metronome of the drum set recording. The last/fourth tick was removed to obscure which of the instruments deviated from the initial metronome reference.

In summary, the experiment included three grooves of different structural complexity (low, medium, and high degree of syncopations and note onsets per bar), of which the low- and medium complexity groove were presented both with ("HH") and without

a meter-supporting time-keeping hi-hat. Each of the grooves ("low," "lowHH," "medium," "mediumHH," "highHH") were presented in five microtiming levels (-80, -40, 0, +40, +80), making a total number of 25 excerpts (see Table 1).

SETUP AND PROCEDURE

Testing took place at the Cognitive Laboratory in the Department of Psychology, University of Oslo, in a windowless and quiet room with constant illumination of about 170 Lux and constant temperature and humidity. We used a Remote Eye Tracking Device (RED) and I-View software commercially available from SensoMotoric Instruments (SMI), Berlin (Germany) to record oculomotor data (pupil diameters, eye movements, gaze fixations, and blinks). The RED system records continuous binocular pupil diameter with high precision (detecting changes as small as .0004 mm, according to SMI specs) using infrared light lamps and an infrared light sensitive video camera. The sampling frequency was set to 60 Hz, which is appropriate for pupilometry. Instructions in English were given on a flat DELL LCD monitor with a screen resolution of 1680 X 1050, installed above the RED video camera. The distance from the participant's eyes to the monitor and eye tracking device was set to 60 cm by use of a chin/head stabilizer. Participants gave tapping responses during the active condition by pressing a key with their chosen hand on the PC's keyboard (Dell L30U) placed in front of the participant. The keyboard was linked to the PC via a STLAB USB port hub. All participants listened to the stimuli via a set of stereo headphones (Philips SBC HP840).

All participants were individually tested by the first author (JFS) or by another experimenter (LKG) who were present in the room at all times during the session. They were requested to keep their eyes open at all times (in order to obtain continuous pupil recordings), except for short eye blinks, and maintain gaze on a black

fixation cross (+) displayed in the middle of the monitor. Furthermore, they were informed that it was possible to rest their eyes when the fixation cross was not present. The fixation cross was used to facilitate anchoring gaze and avoid that participants would look away from the screen, causing a loss of oculomotor data.

The screen had a neutral grey color as background (pixels' RGB coordinates: 163,163,163). All sessions began with a standard 4-point eye calibration procedure, then a general instruction slide appeared on screen, followed by the individual adjusting of the sound volume to a comfortable and clearly audible level. Most participants kept the default volume that was preset by the experimenter, or made only minute adjustments. Each trial was self-paced by the participant, who pressed the spacebar to start each trial. The fixation cross was visible in the middle of the screen for three seconds before (as a baseline measurement) as well as during the playback of each of the music clips. First, the 25 excerpts in the "listening only" condition were played. The order of stimuli was pseudorandomized within each block, with the constraint that none of the same groove types (e.g., "lowHH") were played back-to-back. The test blocks were counterbalanced, so that half of the participants in each group (musicians/nonmusicians) listened to the stimuli in the opposite order. After listening to each excerpt, the participants were asked two questions (on a 5-step Likert scale: 1 = *not at all*, 5 = *to a great extent*): 1) *To what extent does listening to this rhythm, make you want to move your body?* and 2) *To what extent do you find that this rhythm sounds musically well-formed?* Question 1 probed the concept of groove experience, treating groove exclusively as a sensorimotor phenomenon, according to Madison's (2006) definition of groove. We did not enquire about "positive affect," since microtiming stimuli were brief bass-and-drums tracks and thus contained a reduced musical context. Question 2 enquired about rating of musical "well-formedness," similar to Davies and colleagues' (2013) "musical naturalness," i.e., to what extent the musical excerpt "sounds like a typical musical performance" (p. 502). After a break, there was an introductory test trial where participants were asked to tap with their index finger, synchronizing to a 30-s isochronous metronome beat. Participants did not receive feedback concerning their performance on this task, and the data were not analyzed. Next, participants were asked to tap (using their index finger of choice) to the three metronome ticks and continue tapping to the basic pulse of the music, as steady and regularly as possible, until the music stopped. For all participants the stimuli were presented in reverse order in the active tapping

condition. After completing the tapping condition, participants were given the Musical Ear Test ("rhythm part" only). The test began with instructions and two practice trials. Participants' task was to decide whether two succeeding rhythms, presented in pairs, were identical or not. Every new trial was initiated by pressing the spacebar. All participants were exposed to the MET stimuli in the same order and no feedback was given during testing. Immediately following the experiment session, each participant filled out an online questionnaire probing their demographics and musical experience and taste, based on that previously used by Laeng and colleagues (2016).

DATA PROCESSING

Pupillary diameters, tapping time events, and subjective rating data were exported with use of BeGaze software (SMI, v.3.4) and imported to Microsoft® Excel where we computed descriptive statistics for each participant. As pupil sizes of left and right eye were nearly identical, we used data only from the left eye. We applied subtractive baseline correction which is common in pupillometry research (Mathôt, Fabius, Van Heusden, & Van der Stigchel, 2018). In the present study, average individual left eye's pupil size in each trial (lasting 30,000 ms) were baseline corrected by subtracting the average pupil size recorded the last 3,000 ms before stimulus onset for each participant per trial. This provides a measure of "pupil size change" that expresses dilations from baseline as positive numbers and relative constrictions as negative numbers.

As for the tapping data, each tap's timing was compared to a corresponding *reference point* of the musical pulse. The interpulse interval for musical reference points was 750 ms (80 bpm). The timing of the drums/metronome (not the double bass) served as the basis for the reference points. To identify and exclude outliers, a pre-processing of tapping data was done by custom-scripting in MATLAB according to rules specified in Figure 2.

In tapping studies, the typical dependent variables can be either a) the mean "absolute" offsets between reference points (e.g., a metronome tick) and the tap, and b) their standard deviation (SD); additionally, one can compute the mean and standard deviation of inter-tap intervals (ITI; see Repp, 2005; Repp & Su, 2013). From preliminary tapping analyses, we discovered the presence of an artifactual "time lag" or "instability" in the software system (most likely generated from the use of a standard keyboard and USB port in collecting taps as key presses; as confirmed by SensoMotoric Instruments' (personal communication), which made the

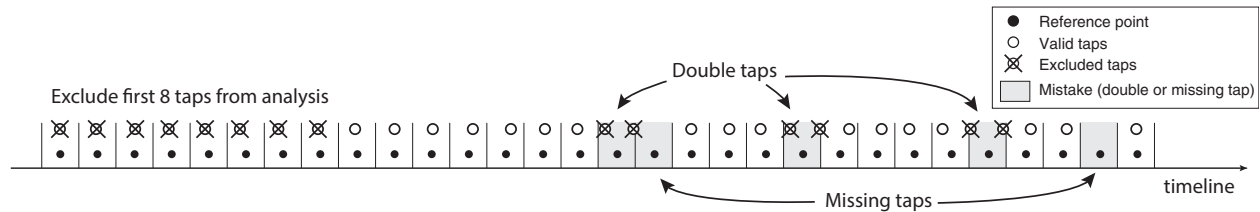


FIGURE 2. Exclusion criteria for finger tapping: A) exclude first eight taps in each series, B) identify and exclude mistakes (i.e., double or missing taps). Series with more than six mistakes were excluded prior to the analyses.

absolute offsets (typically a measure of “tapping accuracy”) unreliable. Nevertheless, the intra-participant tapping offset distributions (typically a measure of “tapping stability”) appeared quite reliable and within the normal range. Hence, we selected “SD offset” (i.e., the standard deviation of offset distribution per trial per participant) as the main dependent variable for tapping in our analyses, as a reversed measure of “tapping stability.” There was still, however, a small element of uncertainty relating to exclusions of occasional taps; as shown in Figure 2, taps were included when being $\pm 50\%$ around each reference points (± 250 ms). Since taps generally could be placed ± 100 ms too late due to technical issues, ± 150 ms were considered outliers and excluded from the statistical analyses. We decided to ignore the intertap interval (ITI), since missing and excluded double taps tended to generate double ITIs (for example 1,000 ms instead of 500 ms). Single trials that had more than six mistakes (i.e., reference points containing double or no taps) were excluded. Furthermore, ten of the nonmusicians’ tapping data sets contained more than 20% (i.e., more than five out of 25) excluded trials. Hence, they were excluded prior to performing the statistical analyses.

Missing pupil data were rare (20 out of totally 3,150 trials, or 0.6%). The total number of excluded and missing tapping trials in included participants was 60 out of 1,325 trials, or 4.5%. There was no link between missing/excluded trials in the pupil and the tapping data sets, respectively. Standard mixed-between repeated-measures ANOVAs were applied for all measures, since these are robust to non-normality in real data found in social sciences, also when the group sizes are unequal (Blanca, Alarcón, Arnau, & Bendayan, 2017). Missing/excluded trials in both pupil and tapping data were estimated using that participant’s group (musician/nonmusician) mean pupil change/SD offset in the respective condition for mean substitution, to fill in the missing cells in the statistical spreadsheet. Such a between-subject mean substitution method is one of the most commonly practiced approaches (Rubin,

Witkiewitz, Andre, & Reilly, 2007) since it preserves the mean of a variable’s distribution, though it reduces the condition’s variance in proportion to the number of missing values. Statistical analyses were performed with either Statview or JASP (v.0.9) statistical software. Greenhouse-Geisser corrections were applied when appropriate.

Results

PUPIL RESPONSES

Figure 3 shows participants’ mean pupil diameter size (in mm) as a function of time (0–30 s) for all trials combined. By visual inspection we can observe a response peak 5–7 s after stimulus onset, and later a negative “drifting” effect with a more or less continuous decay—the pupil generally tended to decrease in size as the trial progressed. Hence, it was reasonable to expect the mean baseline-corrected pupil sizes in both experiments to be either close to or below the zero level, as the baseline-condition was only three seconds in duration and measured right before each stimulus, whereas the experimental condition took 30 s.

A $5 \times 5 \times 2 \times 2$ mixed between-within repeated-measures ANOVA of mean baseline-corrected pupil size change was performed. Within-subjects factors were Microtiming asynchrony (-80 ms, -40 ms, 0, $+40$ ms, $+80$ ms), Complexity (“low,” “lowHH,” “medium,” “mediumHH,” “highHH”) and Activity (“passive” vs. “active” conditions); the between-subjects factor was Musicianship (musicians vs. nonmusicians). The analysis yielded significant main effects of Microtiming asynchrony, $F(4, 244) = 3.55, p = .008, \eta_p^2 = .06$; Complexity (i.e., different levels of syncopations/note onsets per bar), $F(3.39, 206.46) = 3.22, p = .019, \eta_p^2 = .05$; and Activity, $F(1, 61) = 75.22, p < .001, \eta_p^2 = .55$. The last effect was evident by significantly larger pupil dilation changes when participants tapped to the beat ($M = -.01$ mm; $SD = .04$) compared to when they listened only ($M = -.18$ mm; $SD = .05$). The analysis did not reveal any significant difference in pupillary response

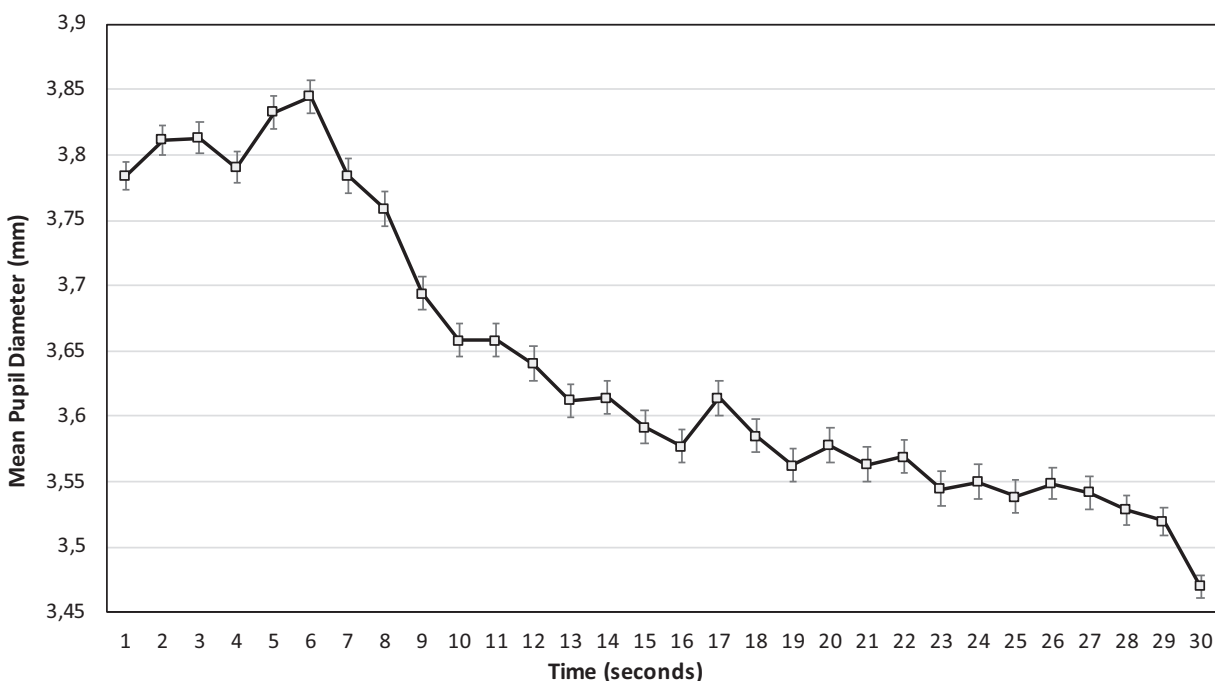


FIGURE 3. Mean pupil size as function of time (in seconds) for all experimental trials combined. Error bars: ± 1 standard errors.

due to Musicianship, $p = .30$. No interaction effects involving any of the factors were found.

The effect of Microtiming on pupil change is illustrated in Figure 4 and was statistically examined further by four planned contrasts, inspired by Hove and colleagues (2007). The first contrast compared all the microtiming asynchrony conditions (averaged across the -80 ms, -40 ms, $+40$ ms, and 80 ms levels) with the averaged on-the-grid (0 ms) condition. The analysis showed that microtiming asynchrony did increase pupil dilation, compared to when microtiming was not present, $F(1, 61) = 4.10$, $p = .047$. In the second contrast, we investigated the effect of increasing microtiming magnitude from 40 ms to 80 ms in either direction; the ± 80 ms conditions yielded larger pupil diameters compared to the ± 40 ms conditions, $F(1, 61) = 7.42$, $p = .008$. The third contrast looked at effects of direction of asymmetry (i.e., we compared the averaged -40 ms and -80 ms conditions with the averaged $+40$ ms and $+80$ ms conditions, respectively). No significant difference between the (-) and (+) conditions, $p = .437$, was found. The fourth and final contrast investigated whether an interaction effect between the ± 40 ms vs. ± 80 ms conditions and the (-) vs. (+) conditions was present; that is, whether the increase in microtiming magnitude from 40 ms to 80 ms affected pupil response differently depending on bass being timed ahead of

versus after the drums. This analysis revealed a nonsignificant result, $p = .421$.

Within the significant main effect of the factor Complexity, we were specifically interested in how the participants' pupil response was influenced by 1) timekeeping (i.e., presence vs. absence of a steady eighth note hi-hat), and 2) syncopation density as well as number of note onsets per bar, when timekeeping was controlled for. Thus, with regards to timekeeping, we performed a planned contrast that compared the pupil response from the "low" and "medium" complexity grooves that included hi-hat ("lowHH" and "mediumHH"), with the "low" and "medium" complexity grooves without hi-hat ("low" and "medium"). This analysis revealed a significant Timekeeper effect, $F(1, 61) = 6.67$, $p = .012$, where the low and medium grooves without hi-hat eighth notes generated a stronger pupil response ($M = -.10$; $SD = .13$) in the participants, than did the corresponding grooves with hi-hat ($M = -.12$; $SD = .14$). With regards to syncopation density as well as number of note onsets per bar, a separate $5 \times 3 \times 2 \times 2$ (Microtiming x ComplexityHH x Activity x Musicianship) mixed between-within repeated-measures ANOVA was performed, that compared pupil change from the three complexity levels only that included hi-hat ("lowHH," "mediumHH," "highHH"), in addition to the original Microtiming, Musicianship, and Activity factors. The

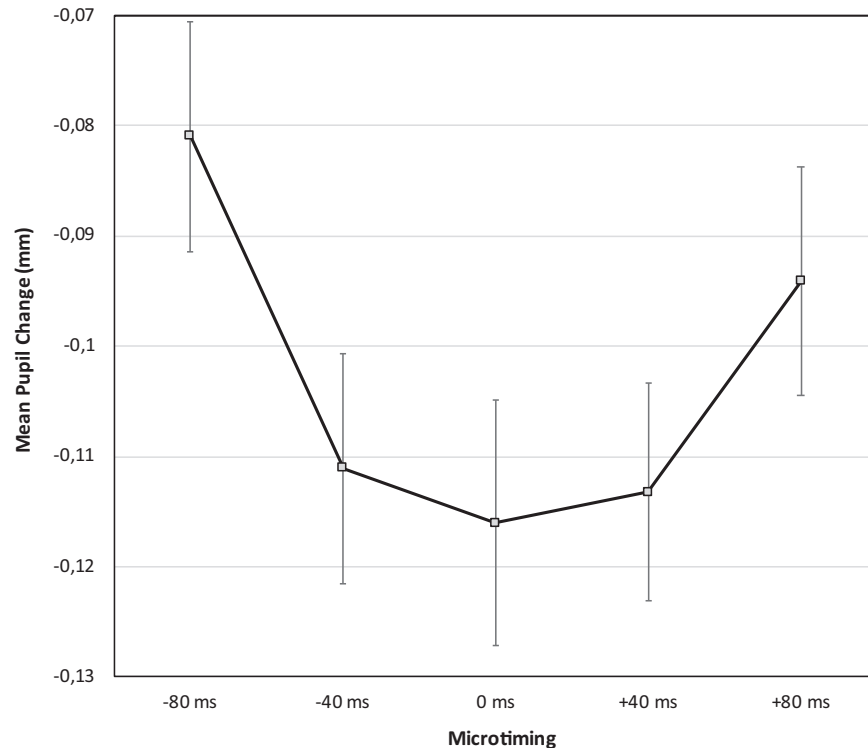


FIGURE 4. Pupil change as a function of Microtiming asynchrony between bass and drums. Error bars: ± 1 standard errors.

effect of this modified Complexity variable was not significant, $p = .108$.

TAPPING STABILITY

We performed a $5 \times 5 \times 2$ (Microtiming \times Complexity \times Musicianship) mixed between-within repeated-measures ANOVA with *SDs* of tapping offsets as the dependent variable and with the same independent variables as in the pupil data analysis above (except Activity, since there was by definition no tapping in the passive condition). This ANOVA revealed significant main effects of Microtiming, $F(2.83, 144.41) = 6.48$, $p < .001$, $\eta_p^2 = .11$; Complexity, $F(2.74, 139.60) = 8.23$, $p < .001$, $\eta_p^2 = .14$; and Musicianship $F(1, 51) = 28.61 < .001$, $\eta_p^2 = .36$. Regarding the latter, musicians ($M = 23.16$ ms; $SD = 11.51$) outperformed (i.e., had lower *SDs* of offsets than) nonmusicians ($M = 74.45$ ms; $SD = 57.64$) in tapping stability across all levels of microtiming asynchrony (See Figure 5). In addition to the main effects, we found a significant Microtiming by Musicianship interaction, $F(2.83, 144.41) = 2.73$, $p = .05$, $\eta_p^2 = .05$; that is, presence of microtiming influenced musicians' and nonmusicians' tapping performance differently, as addressed more specifically below. There were no other significant interaction effects. Visual

inspection of the tapping *SDs* histograms (shown separately for musicians and nonmusicians) revealed two distributions that are both skewed towards the left. Consequently, we verified our results by the use of non-parametric analyses (Independent Mann-Whitney *U* test and Friedman's two-way ANOVA by Ranks) of the same tapping data, which showed corresponding results to the *F* test. Hence, we decided to keep the original ANOVAs of Tapping *SDs*.

The Microtiming factor was investigated similar to the pupil response data above, with the same four planned contrasts. Given the significant main and interactive effects involving Musicianship, analyses for musicians and nonmusicians were performed separately. The first planned contrast revealed that presence of microtiming (i.e., the averaged -80 ms, -40 ms, $+40$ ms, and 80 ms conditions) decreased tapping stability for musicians, $F(1, 19) = 10.51$, $p = .004$, compared to the averaged on-the-grid versions. However, for the nonmusicians, this analysis did not yield a significant result ($p = .16$); that is, their tapping was equally inaccurate across microtiming. In the second planned contrast, an effect of microtiming magnitude was found, since the ± 80 ms conditions yielded lower tapping stability compared to the ± 40 ms conditions, both for musicians

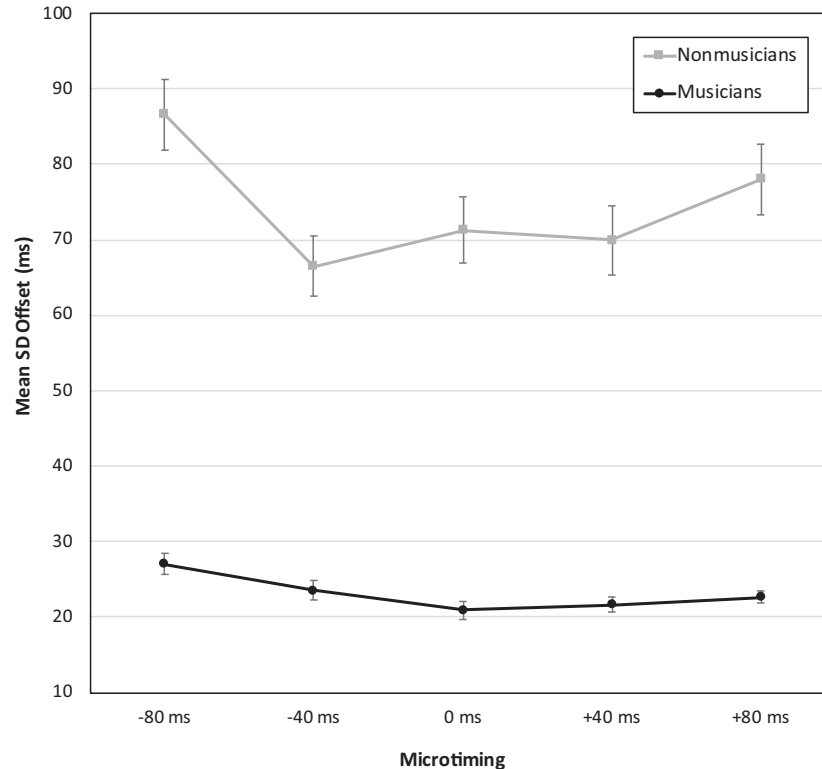


FIGURE 5. Tapping stability (SD offset in ms) as a function of Microtiming asynchrony between bass and drums split by Musicianship. Error bars: ± 1 standard errors.

$F(1, 19) = 6.39, p = .021$, and for nonmusicians $F(1, 32) = 13.29, p < .001$. According to the third planned contrast, the averaged -40 ms and -80 ms conditions gave larger tapping SDs of offsets (e.g., lower tapping stability) than the averaged $+40$ ms and $+80$ ms conditions in the musicians, $F(1, 19) = 17.65, p < .001$, indicating an asymmetry between the effects of “bass ahead of the drums,” versus “bass behind the drums.” However, this asymmetry did not appear for the nonmusicians ($p = .37$). Fourth, the interaction effect between the ± 40 ms vs. ± 80 ms conditions and the (-) vs. (+) conditions, approached significance for the musicians, $F(1, 19) = 4.17, p = .055$, and was significant for the nonmusicians, $F(1, 32) = 5.65, p = .024$. This indicates that tapping stability was more negatively influenced by an increase in microtiming asynchrony magnitude (from 40 ms to 80 ms), when bass timing preceded drum timing, than when bass timing succeeded drum timing.

The main effect of Complexity was also scrutinized similarly to the pupil response analyses above (with the exception that musicians and nonmusicians were analyzed separately). First, we investigated the potential Timekeeper effect (significant for the pupil data).

A planned contrast compared the averaged tapping SDs from the “low” and “medium” complexity grooves with (“lowHH” and “mediumHH”), and without (“low” and “medium”) hi-hat. Indeed, presence of eighth notes hi-hat increased tapping stability (i.e., decreased tapping SDs of offsets), both for musicians, $F(1, 19) = 108.11, p < .001$; with hi-hat: $M = 19.09, SD = 9.15$; without hi-hat: $M = 28.65, SD = 12.49$, and for nonmusicians, $F(1, 32) = 29.06, p < .001$; with hi-hat: $M = 65.01, SD = 54.98$; without hi-hat: $M = 82.57, SD = 56.76$.

Further, we calculated a $5 \times 3 \times 2$ (Microtiming x ComplexityHH x Musicianship) mixed between-within repeated-measures ANOVA of tapping SDs from the three complexity levels that included hi-hat (“lowHH,” “mediumHH,” “highHH”), as well as the original levels of the two other factors. According to this analysis, complexity did not significantly affect tapping stability when timekeeping (i.e., eighth notes hi-hat) was controlled for ($p = .12$).

RATING SCALES

The two rating scales were based on the following questions: 1) *To what extent does listening to this rhythm,*

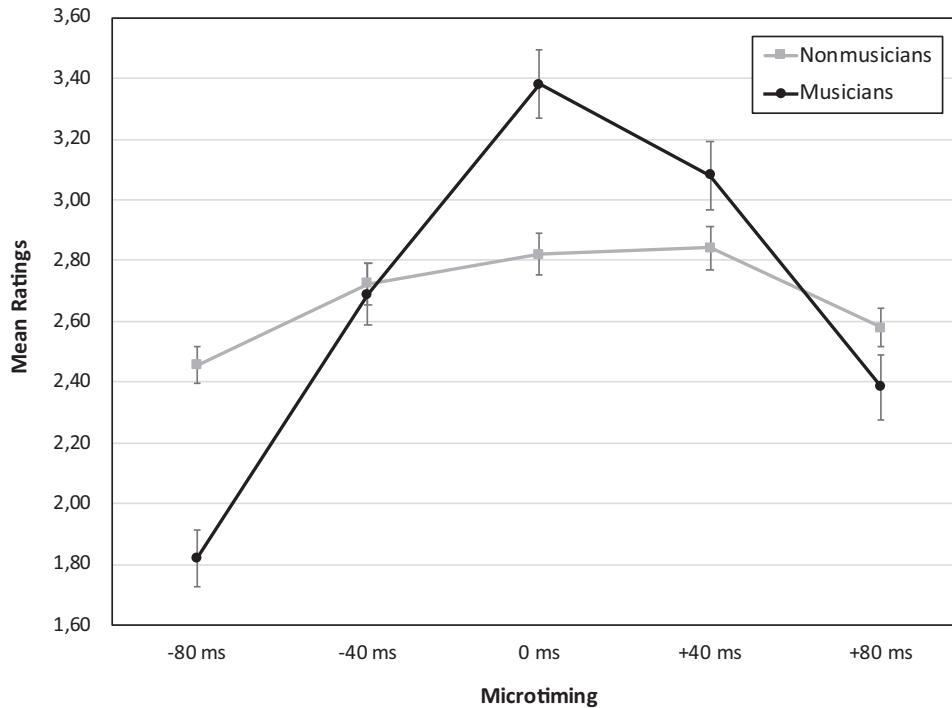


FIGURE 6. Microtiming asynchrony on subjective ratings (based on two rating scales, ranging 1-5) split by Musicianship. Error bars: ± 1 standard errors.

make you want to move your body? (“move body”), and 2) *To what extent do you find that this rhythm sounds musically well-formed?* (“well-formed”). Participants’ responses on the two items were moderately to highly correlated (musicians: $r = .78$, $p < .001$; nonmusicians: $r = .58$, $p < .001$). Thus, for simplicity, we collapsed data across the two questions (“move body” and “well-formed”). A $5 \times 5 \times 2$ (Microtiming \times Complexity \times Musicianship) mixed between-within repeated-measures ANOVA of subjective ratings was performed, with the same independent variables as in the pupil and tapping analyses (the Activity factor was not included, since subjective ratings were done in the passive condition only). The results show significant main effects of Microtiming, $F(3.11, 189.77) = 65.88$, $p < .001$, $\eta_p^2 = .52$; and Complexity, $F(2.77, 168.78) = 42.66$, $p < .001$, $\eta_p^2 = .41$. The main effect of Musicianship was not significant ($p = .93$). However, there were significant interaction effects of Microtiming by Musicianship, $F(3.11, 189.77) = 22.67$, $p < .001$, $\eta_p^2 = .27$ (as shown in Figure 6) and Complexity by Musicianship, $F(2.77, 168.78) = 5.83$, $p = .001$, $\eta_p^2 = .09$. These results are further addressed below and imply that musicians’ and nonmusicians’ ratings were differently affected by variations in both Microtiming and Complexity, respectively. In

addition, the Microtiming by Complexity interaction, $F(10.86, 662.71) = 7.82$, $p < .001$, $\eta_p^2 = .11$ is shown in Figure 7 and suggests that variations in structural Complexity affected subjective ratings differently, depending on the microtiming configuration (bass preceding drums or bass succeeding drums). There was no significant three-way interaction.

Similar to the procedure for the pupil and tapping data, the Microtiming factor was further investigated through four planned contrasts. As with the tapping data, we performed analyses for musicians and nonmusicians separately. First, a planned comparison of the averaged -80 ms, -40 ms, $+40$ ms, and $+80$ ms conditions, with the 0 ms condition, showed that presence of microtiming significantly reduced subjective ratings of groove in both musicians $F(1, 19) = 50.22$, $p < .001$, and nonmusicians $F(1, 42) = 14.43$, $p < .001$. Second, musicians, $F(1, 19) = 30.62$, $p < .001$, and nonmusicians, $F(1, 42) = 29.2$, $p < .001$, rated the ± 80 ms conditions lower than the ± 40 ms conditions. Thus, the more closely aligned in time the bass and drums, the higher ratings. As confirmed by the significant Microtiming by Musicianship interaction (and clearly illustrated in Figure 6) the musicians were to a greater degree responsible for this effect; musicians were more

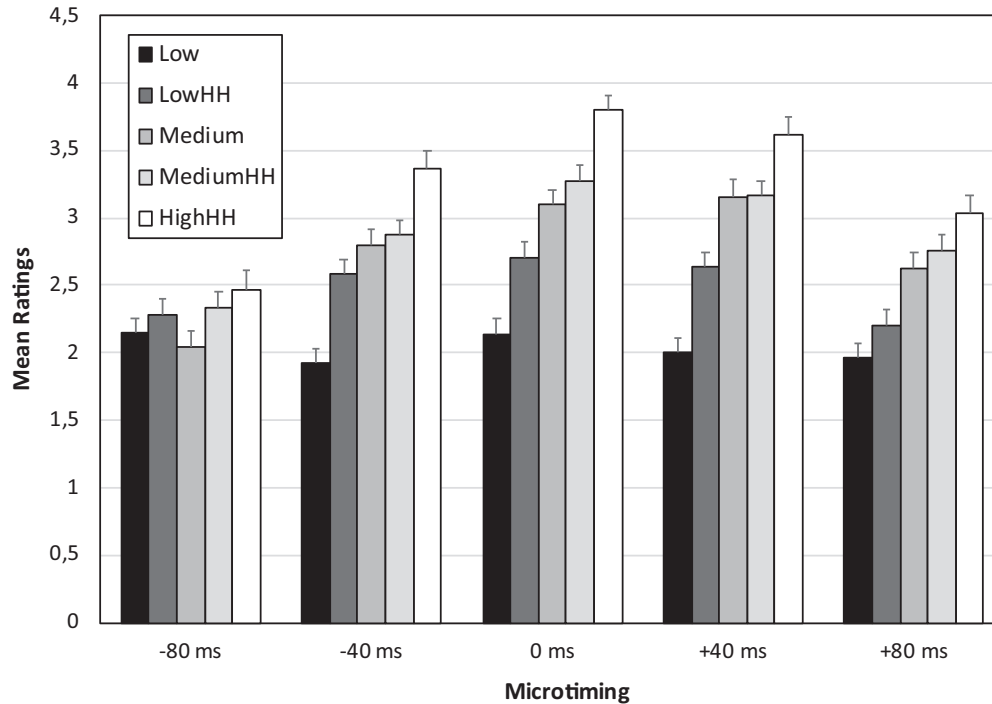


FIGURE 7. Interaction effect of groove Complexity and Microtiming asynchrony on subjective ratings (based on two rating scales, ranging 1-5). Error bars: ± 1 standard errors.

responsive to microtiming than nonmusicians and tended to use a wider range of ratings. The third contrast confirmed that musicians, $F(1, 19) = 23.26$, $p < .001$, and nonmusicians, $F(1, 42) = 12.36$, $p = .001$, rated the excerpts higher when bass timing succeeded (+), rather than preceded (-) drum timing. From the graph in Figure 6 it is evident that the -80 ms has a specially detrimental effect of groove among the musicians. The fourth planned contrast looked for possible interaction effects between the ± 40 ms vs. ± 80 ms conditions and the (-) vs. (+) conditions, leaving non-significant results for both groups (musicians: $p = .39$; nonmusicians: $p = .96$). In other words, the increase in microtiming magnitude from 40 ms to 80 ms did not affect ratings differently depending on bass being timed ahead of versus after the drums.

As to the Complexity factor, we first looked for a possible Timekeeper-effect. This was significant for the nonmusicians only, where presence of eighth note hi-hat increased ratings, $F(1, 42) = 75.75$ (with hi-hat: $M = 2.76$, $SD = .88$; without hi-hat: $M = 2.32$, $SD = .92$). For the musicians, timekeeping did not affect subjective rating, $p = .72$. Second, a $5 \times 3 \times 2$ (Microtiming \times ComplexityHH \times Musicianship) mixed between-within repeated-measures ANOVA, analyzed the subjective

ratings from the three complexity levels that included hi-hat (“lowHH,” “mediumHH,” “highHH”) and the original levels of the two other factors. According to this analysis, Complexity* positively affected subjective ratings when timekeeping (i.e., eighth notes hi-hat) was controlled for, $F(1.68, 102.20) = 30.31$, $p < .001$, $\eta_p^2 = .33$. There was no interaction effect between ComplexityHH and Musicianship, $p = .19$. Post hoc t -tests with Bonferroni correction revealed that high complexity grooves with hi-hat ($M = 3.26$; $SD = 1.11$) was rated higher than both low ($M = 2.48$; $SD = 0.93$; $t = 7.21$, $p_{\text{bonf}} < .001$) and medium ($M = 2.88$; $SD = 0.96$), $t = 3.71$, $p_{\text{bonf}} < .001$, grooves with hi-hat. Additionally, the mediumHH grooves was rated higher than lowHH grooves, $t = 5.49$, $p_{\text{bonf}} < .001$.

Finally, we performed a multiple regression analysis with pupil change as the dependent variable and Tapping stability and Rating scales as the independent variables. Tapping stability was a significant predictor of pupil change (Regression coefficient = $-.001$, $t = -3.15$, $p = .002$), indicating that the more attentional resources or effort were allocated to the task (i.e., larger pupils), the more stable was the tapping. In contrast, there was no linear relationship between Rating scales and pupils (Regression coefficient = $-.008$, $t = 0.67$, $p = .50$).

Discussion

The present results support our hypothesis that the brain allocates increased effort to process musical stimuli whenever the rhythmic relations challenge fundamental metrical models. When microtiming asynchronies between bass and drums in the groove excerpts increased, we observed corresponding increases in cognitive workload as reflected in pupil size. Furthermore, microtiming asynchrony also negatively influenced tapping stability. However, the nonmusicians had highly variable and instable tapping rates, so that only the more “extreme” microtiming conditions of ± 80 ms caused further deterioration of their tapping performance.

Based on Dynamic Attending Theory (Large & Jones, 1999), increased attentional processing demands would be a result of participants experiencing prediction errors (also as manifested by lower tapping stability). Thus, accordingly, listeners would need to adjust the neural oscillations’ phase and/or widen the temporal attentional focus, to account for the incoming rhythmic events across the 30-s music duration; that is, larger asynchronies between the timing of the bass and drums increased the need for adjustments in the locus and shape of the attentional focus.

One should note that both the pupil diameter and performance (tapping) may be indicative of response conflicts (Kamp & Donchin, 2015). The inherent temporal ambiguity of exact beat placement, especially in the 80 ms microtiming asynchrony conditions, may generate conflicts (either at the attentional or motoric levels) about the timing of the taps. Overall, the pupil dilation and SDs of tapping offsets increased when bidirectional microtiming asynchrony increased from 40 ms to 80 ms. Research has shown that when asynchronies exceed 100 ms, sound onsets—even when in a similar frequency range—tend to be experienced as separate, rather than integrated events (see Repp, 2005; Warren, 1993). It is thus possible that the 80 ms approaches the threshold for *temporal integration* whereas a 40 ms asynchrony may be well within it.

Further supporting DAT was the finding that time-keeping (adding hi-hat eighth notes) was negatively related to effort while positively related to tapping stability. Furthermore, the overall “Complexity” effect of tapping was driven exclusively by the lower tapping stability of the non-HH (i.e., “low” and “medium”) stimuli. We surmise that timekeeping reduced metric ambiguity and increased predictability, independent of whether microtiming or other complexity features was present or absent. Likely, the hi-hat eighth-notes provided listeners with more temporal information and

added extra prominence to the drum kit timing, making it appear more like the “main” timing or a salient temporal anchor, in favor of the timing of the bass or a combined temporal event that was “somewhere between the timing of the bass and drum kit.”

Regarding musical expertise or musicianship, a number of previous studies have documented enhanced perceptual skills of musicians for rhythm and meter (e.g., Chen et al., 2008; Drake et al., 2000; Geiser et al., 2010; Matthews et al., 2016; Stupacher et al., 2017). We reasoned that expertise could result in decreases in effort but also, conversely, in increases in attentional allocation for salient events during music making or listening. Thus, we expected that musicianship would have measurable effects on the pupillary response to microtiming changes, although being unable to make specific predictions about their *direction* (decrease or increase of response). However, our analyses failed to reveal any conclusive evidence one way or the other. One possibility is that there might be actual but small differences between the groups that could surface with larger samples or higher statistical power. Another possibility is that the hypothesized processes cancelled each other out. Nevertheless, we did observe enhanced processing of microtiming features in the tapping data of the musicians. We surmise that differences in tapping stability may reflect differences in neural functionality, engendered by levels of rhythmic expertise (e.g., Stupacher et al., 2017). Due to space limitations and small *N*, we did not investigate tapping stability as a function of musical instrument (e.g., drummers vs. pianists), which has been found to play a role in earlier studies (e.g., Krause, Pollok, & Schnitzler, 2010).

As mentioned earlier, although microtiming is assumed to possess a vital function in much groove-based music, systematic experimental investigations have given inconsistent evidence as to whether microtiming asynchronies in grooves actually promote groove experience (Davies et al., 2013; Kilchenmann & Senn, 2015). The present findings seem consistent with several previous studies (Davies et al., 2013; Senn et al., 2016), where musicians showed an increased responsivity to different microtiming conditions, compared to nonmusicians, and used a wider range of ratings. Moreover, with one exception, we also found that the larger the microtiming asynchrony magnitudes, the lower the ratings of the clips. Interestingly, however, the exception to this general pattern was the groove resembling the D’Angelo tune *Really Love*. The +40 ms version of this high complexity groove was the second highest rated clip in the experiment (as shown in Figure 7; the only clip that was rated slightly higher, was the on-the-grid

version of this high complexity groove). This suggests that the musical context probably is crucial to the groove effect of microtiming asynchronies. In the model for this high complexity groove, the bass is consistently timed 40–60 milliseconds after the kick drum. The high rating of this substantial microtiming asynchrony might be explained by the ways in which it interacts with the other rhythmic events in the groove pattern as well as preferences associated with the musical style of this clip (i.e., R&B/soul).

It is plausible to think that musical competence might increase not only responsivity, but also tolerance (as for example suggested by Hove et al., 2007), or even preference for microtiming differences; however, except for the D'Angelo example discussed above, such tolerance was not reflected in subjective rating data in the present study. Instead, it appeared that when rhythmic information (i.e., microtiming) challenged the metrical model, the musicians' stronger metrical model was more "negatively" responsive, as seen in their ratings, than for the nonmusicians. One important issue, however, is that ratings were done in the passive non-movement condition only. Vuust and colleagues (2018) speculate that metric models are strengthened by bodily movement synchronized with the music; hence, a movement condition could have increased participants' tolerance for microtiming.

Janata and colleagues (2012) suggest that the affectively-positive psychological state of "feeling in the groove" is closely linked to *perceptual fluency*; that is, a state of being where individuals are able to anticipate meter and musical onsets, and actualize or imagine them in relation to body movements. According to such a view, high predictability and "effortless" feeling are two related key aspects of groove experience. However, research has also emphasized how groove-based music clearly contains structural complexities that may challenge or even violate listeners' sense of meter (Huron, 2006; Vuust et al., 2014; Witek, 2017). Given this somewhat paradoxical "effortless-but-complex (i.e., attention-demanding)"-dialectic of groove and groove experience, it seems an intriguing endeavor for future studies to unravel possible systematic relations between the objective measurement of effort (e.g., via pupillometry) and the subjective rating of groove. In the present study, microtiming as a rhythmic complexity seemed to *increase* prediction error in a way that concomitantly increased effort and decreased feelings of groove. In other words, under influence of microtiming, effort was negatively related to groove. Remarkably, the microtiming asynchronies' effect on pupil size yielded a U-shaped curve (Figure 4), while the microtiming effect on groove

rating gave an *inverted* U-shaped curve (Figure 6). The ± 80 ms (and in particular the -80 ms) conditions were probably experienced more like metric violations than groove promoters, and had detrimental effects on groove rating. The on-the-grid excerpts on the other hand, being rhythmically most predictable, were the highest rated, had the highest stability of tapping, and demanded minimum mental effort from the participants.

Further supporting the *effortless*-side of groove (i.e., a negative correlation between groove and effort), was the finding that presence of hi-hat eighth notes decreased effort and tapping SDs, while it—albeit only for nonmusicians—increased subjective groove ratings. Interestingly, musicians were indifferent to presence of a steady hi-hat in their ratings. Given the assumption that the musicians' metrical model is stronger than the nonmusicians', it is plausible to speculate that the temporal information from the eighth note subdivisions, is to a greater degree *already present* in the expert group's internal reference structure. Hence, although the eighth note hi-hat *did* decrease musicians' effort and increase their tapping stability, it did not influence how musically well-formed or body move-inducing they rated the clips to be.

The present results indicate that groove does not only emanate from predictability but also from *complexity* since structural complexity (i.e., increasing numbers of syncopations and note onsets per bar) promoted the participants' feeling of groove. Interestingly, respondents were also *more* responsive to (i.e., tolerated less) microtiming when structural complexity increased (this is seen in Figure 7 as larger fluctuations in ratings/steeper profiles on the graph from on-the-grid to microtiming conditions). Nevertheless, we note that our structural complexity variable does "contain" both increases in "real" rhythmic complexity (syncopations) as well as increments in note onsets; the latter may in fact support the listeners' metric models, as suggested by the pupil and tapping effect of adding hi-hat eighth notes to the grooves. The increase in note onsets might well be a reason for why the structural complexity variable was unrelated to mental effort and tapping stability when controlling timekeeping. Furthermore, we did not include musical clips with an "exaggerated" syncopation density, as done by for example Witek and colleagues (2014). Therefore, we did not have the chance to investigate the effect on effort, tapping, and rating of temporal relations exceeding the number of syncopations beyond a degree that is indeed groove-promoting.

An intriguing dissociation in the present study was that participants seemed to both prefer (as evidenced by subjective ratings) and tolerate (as evidenced by tapping

data) grooves where the double bass was timed *after* (+) the drums compared to *ahead* (-) of the drums. This can be accounted for by the sound of the bass having a different musical function than the drums in musical grooves. Drums often work as the primary timing reference and if the sound of the bass is placed after the drums, the longer notes of the bass may work to widen the beat, adding weight to the pulse after the attack. This is a common way to shape the beats of the pulse in many African-American musical styles (see, for example, Danielsen, 2006; Iyer, 2002). Bass succeeding drums may thus be more in line with the way Western participants often hear the combination of these two instruments in genres that are related to the present stimuli; that is, as an accent followed by a widening of the sound (but see Matsushita & Nomura, 2016, for an opposite conclusion using Japanese participants).

Finally, we generally found indications of higher effort in the active “tap with the beat” condition, compared to the passive “listen only” condition. This is not surprising, as music production is clearly a more complex activity than music perception (Zatorre, Chen, & Penhune, 2007) and the act of synchronizing or entraining two events—an internally generated (the rhythm of taps) and an external one (the rhythm heard)—is likely to require a great deal of attentional resources. Indeed, tapping involves an active participation in the groove, demanding a continuously sustained and coordinated motoric “realization” of the internal metric model. Even making a simple motor response like an occasional single key press may enhance the intensity of attentional focusing (Moresi et al., 2008; Van der Molen, Boomsma, Jennings, & Nieuwboer, 1989) over passively listening, thus yielding larger pupil sizes in the former than the latter (Laeng et al., 2016).

Conclusion and Limitations

This is, to our knowledge, the first time an index of mental effort—measured by pupillary response—was employed during exposure to instances of microtiming in a groove context. Magnitudes of microtiming asynchronies between bass and drums were found to be positively related to mental effort and negatively related to tapping stability. Timekeeping was negatively related to mental effort and positively related to stability of tapping. These results are consistent with the Dynamic Attending Theory (DAT), which predicts that rhythmic complexity lowers prediction (as well as tapping stability) and thus should increase effort. We also note that recent accounts based on Friston’s Predictive Coding Theory (Clark, 2013; Friston, 2005; Koelsch, Vuust, &

Friston, 2019; Vuust et al., 2018, 2009) seem consistent with the present findings and with DAT. On-the-grid grooves were generally preferred to (i.e., rated as higher in groove than) microtiming grooves, consistent with previous research. Timing conditions where the bass was timed ahead of the drums were less preferred than the reverse and yielded lower tapping stability. When moderated by microtiming, groove ratings were negatively related to effort, possibly because microtiming was experienced more like a metric violation than a groove promoter. The groove that was modelled on a real musical example was an exception to this pattern. The on-the-grid and +40 versions of this groove were the highest rated clips in the experiment. Better perceptual skills related to rhythm perception in professional jazz musicians appeared to be mirrored in tapping performance, and increased responsivity in subjective ratings.

As in most scientific investigations, the present study has several limitations. One is that the asynchronies were artificially constructed. This manipulation made experimental control possible, but has drawbacks related to the question of microtiming’s relation to groove. A rhythmic event played by drums or bass is usually not only timed but also sonically shaped in accordance with its microtemporal position (Cámara, Nymoen, Lartillot & Danielsen, 2019; Danielsen, Waa-deland, et al., 2015). We decided to present the ecological stimuli (“high”) in a version *with* hi-hat only, while the “low” and “medium” complexity stimuli were presented both with and without hi-hat. However, this prevented us from investigating the particular effect of the hi-hat and resulted in a not completely balanced experimental design. Nevertheless, it shortened the experimental session and possibly reduced fatigue. There were also technical problems related to the collection of the tapping key presses. The absolute offset tapping data seemed unreliable, which made it difficult to investigate how microtiming influenced absolute offsets, including negative asynchronies. Clearly, it was non-optimal to tap on a standard PC-keyboard; a device specially made for tapping would perhaps have improved the tapping performance. A final and potentially important caveat is that each participant’s head was kept in a stable position by a chinrest while listening to music, which might have improved pupil data but constrained spontaneous movements to the music. In future research, we wish to combine pupillometry with motion capture to allow for spontaneous movements and also be able to measure participants’ actual movement rates under music listening (cf. Kilchenmann & Senn, 2015). Such movements might in fact be crucial to both the processing and appreciation of groove and the

microtiming asynchronies of the magnitudes tested in the present study.

Author Note

We are grateful to Jakop Filip Janssønn Hauan for playing and recording the drums on the auditory stimuli; Martin Torvik Langerød (MTL) for engineering the double bass recordings and preparing the experimental sound clips; Kristian Nymo for preparation of tapping data with custom MATLAB analyses and design of Figure 2; Peter Vuust, Dag Erik Eilertsen and Rebekah Oomen for advice at earlier stages of the manuscript; and Lara Katrina Garvija (LKG) for testing some of the

participants. This work was partially supported by the Research Council of Norway through its Centres of Excellence scheme (Project 262762) and the TIME project (Grant 249817).

The experimental raw data from the eye-tracker is publicly available in four folders on Figshare. Web addresses (DOI) are: 10.6084/m9.figshare.8246189; 10.6084/m9.figshare.8246201; 10.6084/m9.figshare.8246213; 10.6084/m9.figshare.8246273

Correspondence concerning this article should be addressed to Professor Anne Danielsen, RITMO Centre for Interdisciplinary Studies in Rhythm, Time and Motion, University of Oslo, Box 1133 Blindern, NO-0318 Oslo, Norway. E-mail: anne.danielsen@imv.uio.no

References

- AHERN, S., & BEATTY, J. (1979). Pupillary responses during information processing vary with scholastic aptitude test scores. *Science*, 205, 1289–1292. <https://doi.org/10.1126/science.472746>
- ALNÆS, D., SNEVE, M. H., ESPESETH, T., ENDESTAD, T., VAN DE PAVERT, S. H. P., & LAENG, B. (2014). Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the locus coeruleus. *Journal of Vision*, 14(4), 1–20. <https://doi.org/10.1167/14.4.1>
- BERLYNE, D. E. (1971). *Aesthetics and psychobiology*. East Norwalk, CT: Appleton-Century-Crofts.
- BIANCHI, F., SANTURETTE, S., WENDT, D., & DAU, T. (2016). Pitch discrimination in musicians and non-musicians: Effects of harmonic resolvability and processing effort. *Journal of the Association for Research in Otolaryngology*, 17(1), 69–79. <https://doi.org/10.1007/s10162-015-0548-2>
- BLANCA, M. J., ALARCÓN, R., ARNAU, J., BONO, R., & BENDAYAN, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema*, 29, 552–557. <https://doi.org/10.7334/psicothema2016.383>
- BOUWER, F. L., VAN ZUIJEN, T. L., & HONING, H. (2014). Beat processing is pre-attentive for metrically simple rhythms with clear accents: An ERP study. *PLoS ONE*, 9(5), 1–9. <https://doi.org/10.1371/journal.pone.0097467>
- BRADLEY, M. M., MICCOLI, L., ESCRIG, M. A., & LANG, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, 45, 602–607.
- BRATTICO, E., PALLESEN, K. J., VARYAGINA, O., BAILEY, C., ANOUROVA, I., JÄRVENPÄÄ, M., ET AL. (2008). Neural discrimination of nonprototypical chords in music experts and laymen: An MEG study. *Journal of Cognitive Neuroscience*, 21, 2230–2244.
- BUTLER, M. J. (2006). *Unlocking the groove: Rhythm, meter, and musical design in electronic dance music*. Bloomington, IN: Indiana University Press.
- BUTTERFIELD, M. (2010). Participatory discrepancies and the perception of beats in jazz. *Music Perception*, 27, 157–176. <https://doi.org/10.1525/mp.2010.27.3.157>
- CÂMARA, G. S., & DANIELSEN, A. (2018). Groove. In A. Rehding & S. Rings (Eds.), *The Oxford handbook of critical concepts in music theory*. Oxford, UK: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190454746.013.17>
- CÂMARA, G. S., NYMOEN, K., LARTILLOT, O., & DANIELSEN, A. (2019). *Effect of instructed timing on electric guitar and bass sound in groove performance*. Manuscript submitted for publication.
- CHEN, J. L., PENHUNE, V. B., & ZATORRE, R. J. (2008). Moving on time: Brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *Journal of Cognitive Neuroscience*, 20, 226–239. <https://doi.org/10.1162/jocn.2008.20018>
- CLARK, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36, 181–204. <https://doi.org/10.1017/S0140525X12000477>
- CLARKE, E. F. (1989). The perception of expressive timing in music. *Psychological Research*, 51, 2–9. <https://doi.org/10.1007/BF00309269>
- COLLIER, G. L., & COLLIER, J. L. (1996). Microrhythms in jazz: A review of papers. *Annual Review of Jazz Studies*, 8, 117–139.
- DAMSMA, A., & VAN RIJN, H. (2017). Pupillary response indexes the metrical hierarchy of unattended rhythmic violations. *Brain and Cognition*, 111, 95–103. <https://doi.org/10.1016/j.bandc.2016.10.004>
- DANIELSEN, A. (2006). *Presence and pleasure: The funk grooves of James Brown and Parliament*. Middletown, CT: Wesleyan University Press.

- DANIELSEN, A. (2010). Here, there and everywhere: Three accounts of pulse in D'Angelo's 'Left & Right.' In A. Danielsen (Ed.), *Musical rhythm in the age of digital reproduction* (pp. 19–36). Farnham, UK: Ashgate Publishing, Ltd.
- DANIELSEN, A. (2019). Pulse as dynamic attending: Analysing beat bin metre in neo soul grooves. In C. Scotto, K. Smith, & J. Brackett (Eds.), *The Routledge companion to popular music analysis: Expanding approaches* (pp. 179–189). New York: Routledge.
- DANIELSEN, A., HAUGEN, M. R., & JENSENIUS, A. R. (2015). Moving to the beat: Studying entrainment to micro-rhythmic changes in pulse by motion capture. *Timing and Time Perception*, 3, 133–154.
- DANIELSEN, A., WAADELAND, C. H., SUNDT, H. G., & WITEK, M. A. G. (2015). Effects of instructed timing and tempo on snare drum sound in drum kit performance. *Journal of the Acoustical Society of America*, 138, 2301–2316. <https://doi.org/10.1121/1.4930950>
- 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, 497–510. <https://doi.org/10.1525/mp.2013.30.5.497>
- DEMAREST, S. M. (1995). Issues of ecological validity for perceptual research in music. *Psychomusicology*, 14, 173–181. <https://doi.org/10.1037/h0094085>
- DHAMALA, M., PAGNONI, G., WIESENFELD, K., ZINK, C. F., MARTIN, M., & BERNS, G. S. (2003). Neural correlates of the complexity of rhythmic finger tapping. *NeuroImage*, 20, 918–926. [https://doi.org/10.1016/S1053-8119\(03\)00304-5](https://doi.org/10.1016/S1053-8119(03)00304-5)
- DRAKE, C., PENEL, A., & BIGAND, E. (2000). Tapping in time with mechanically and expressively performed music. *Music Perception*, 18, 1–23. <https://doi.org/10.2307/40285899>
- EINHÄUSER, W., STOUT, J., KOCH, C., & CARTER, O. (2008). Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 1704–1709. <https://doi.org/10.1073/pnas.0707727105>
- ETANI, T., MARUI, A., KAWASE, S., & KELLER, P. E. (2018). Optimal tempo for groove: Its relation to directions of body movement and Japanese nori. *Frontiers in Psychology*, 9, 1–13. <https://doi.org/10.3389/fpsyg.2018.00462>
- FINK, L. K., HURLEY, B. K., GENG, J. J., & JANATA, P. (2018). A linear oscillator model predicts dynamic temporal attention and pupillary entrainment to rhythmic patterns. *Journal of Eye Movement Research*, 11(2), 1–24. <https://doi.org/10.16910/jemr.11.2.12>
- FRIBERG, A., & SUNDSTRÖM, A. (2002). Swing ratios and ensemble timing in jazz performance: Evidence for a common rhythmic pattern. *Music Perception*, 19, 333–349. <https://doi.org/10.1525/mp.2002.19.3.333>
- FRIEDMAN, D., HAKEREM, G., SUTTON, S., & FLEISS, J. L. (1973). Effect of stimulus uncertainty on the pupillary dilation response and the vertex evoked potential. *Electroencephalography and Clinical Neurophysiology*, 34, 475–484. [https://doi.org/10.1016/0013-4694\(73\)90065-5](https://doi.org/10.1016/0013-4694(73)90065-5)
- FRISTON, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B*, 360, 815–836. <https://doi.org/10.1098/rstb.2005.1622>
- FUJIOKA, T., TRAINOR, L. J., LARGE, E. W., & ROSS, B. (2009). Beta and gamma rhythms in human auditory cortex during musical beat processing. *Annals of the New York Academy of Sciences*, 1169, 89–92. <https://doi.org/10.1111/j.1749-6632.2009.04779.x>
- GEISER, E., SANDMANN, P., JÄNCKE, L., & MEYER, M. (2010). Refinement of metre perception - Training increases hierarchical metre processing. *European Journal of Neuroscience*, 32, 1979–1985. <https://doi.org/10.1111/j.1460-9568.2010.07462.x>
- GRANHOLM, E., ASARNOW, R. F., SARKIN, A. J., & DYKES, K. L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology*, 33, 457–461.
- HAMMERSCHMIDT, D., & WÖLLNER, C. (2018). The impact of music and stretched time on pupillary responses and eye movements in slow-motion film scenes. *Journal of Eye Movement Research*, 11(2), 1–16. <https://doi.org/10.16910/jemr.11.2.10>
- HONING, H. (2012). Without it no music: Beat induction as a fundamental musical trait. *Annals of the New York Academy of Sciences*, 1252, 85–91. <https://doi.org/10.1111/j.1749-6632.2011.06402.x>
- HOVE, M. J., KELLER, P. E., & KRUMHANSL, C. L. (2007). Sensorimotor synchronization with chords containing tone-onset asynchronies. *Perception and Psychophysics*, 69, 699–708. <https://doi.org/10.3758/BF03193772>
- HURON, D. B. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: MIT Press.
- IVERSEN, J. R., REPP, B. H., & PATEL, A. D. (2009). Top-down control of rhythm perception modulates early auditory responses. *Annals of the New York Academy of Sciences*, 1169, 58–73. <https://doi.org/10.1111/j.1749-6632.2009.04579.x>
- IYER, V. (2002). Embodied mind, situated cognition, and expressive microtiming in African-American music. *Music Perception*, 19, 387–414. <https://doi.org/10.1525/mp.2002.19.3.387>
- JACK, A. I., & ROEPSTORFF, A. (2002). Introspection and cognitive brain mapping: From stimulus–response to script–report. *Trends in Cognitive Sciences*, 6, 333–339. [https://doi.org/10.1016/S1364-6613\(02\)01941-1](https://doi.org/10.1016/S1364-6613(02)01941-1)
- 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, 54–75. <https://doi.org/10.1037/a0024208>

- JÄNCKE, L., PETERS, M., HIMMELBACH, M., NÖSSELT, T., SHAH, J., & STEINMETZ, H. (2000). fMRI study of bimanual coordination. *Neuropsychologia*, 38, 164–174. [https://doi.org/10.1016/S0028-3932\(99\)00062-7](https://doi.org/10.1016/S0028-3932(99)00062-7)
- JONES, M. R. (2016). Musical time. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (2nd ed., pp. 1–23). Oxford, UK: Oxford University Press.
- JONES, M. R., & BOLTZ, M. (1989). Dynamic attending and responses to time. *Psychological Review*, 96, 459–491. <https://doi.org/10.1037/0033-295X.96.3.459>
- JONES, M. R., JAGACINSKI, R. J., YEE, W., FLOYD, R. L., & KLAPP, S. T. (1995). Tests of attentional flexibility in listening to polyrhythmic patterns. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 293–307. <https://doi.org/10.1037/0096-1523.21.2.293>
- JUST, M. A., CARPENTER, P. A., & MIYAKE, A. (2003). Neuroindices of cognitive workload: Neuroimaging, pupillometric and event-related potential studies of brain work. *Theoretical Issues in Ergonomics Science*, 4, 56–88. <https://doi.org/10.1080/14639220210159735>
- KAHNEMAN, D. (1973). *Attention and effort*. Engelwood Cliffs, NJ: Prentice Hall.
- KAHNEMAN, D. (2011). *Thinking, fast and slow*. New York: Farrar, Strauss, Giroux.
- KAHNEMAN, D., & BEATTY, J. (1966). Pupil diameter and load on memory. *Science*, 154, 1583–1585.
- KAHNEMAN, D., & BEATTY, J. (1967). Pupillary responses in a pitch-discrimination task. *Perception and Psychophysics*, 2, 101–105. <https://doi.org/10.3758/BF03210302>
- KAHNEMAN, D., TURSKY, B., SHAPIRO, D., & CRIDER, A. (1969). Pupillary, heart rate, and skin resistance changes during a mental task. *Journal of Experimental Psychology*, 79, 164–167.
- KAMP, S. M., & DONCHIN, E. (2015). ERP and pupil responses to deviance in an oddball paradigm. *Psychophysiology*, 52, 460–471. <https://doi.org/10.1111/psyp.12378>
- KEIL, C. (1987). Participatory discrepancies and the power of music. *Cultural Anthropology*, 2, 275–283. <https://doi.org/10.1525/can.1987.2.3.02a00010>
- KEIL, C., & FELD, S. (1994). *Music grooves*. Chicago, IL: University of Chicago Press.
- KILCHENMANN, L., & SENN, O. (2015). Microtiming in swing and funk affects the body movement behavior of music expert listeners. *Frontiers in Psychology*, 6, 1–14. <https://doi.org/10.3389/fpsyg.2015.01232>
- KOELSCH, S., VUUST, P., & FRISTON, K. (2019). Predictive processes and the peculiar case of music. *Trends in Cognitive Sciences*, 23, 63–77. <https://doi.org/10.1016/j.tics.2018.10.006>
- KRAUSE, V., POLLOK, B., & SCHNITZLER, A. (2010). Perception in action: The impact of sensory information on sensorimotor synchronization in musicians and non-musicians. *Acta Psychologica*, 133, 28–37. <https://doi.org/10.1016/j.actpsy.2009.08.003>
- LADINIG, O., HONING, H., HÄDEN, G., & WINKLER, I. (2009). Probing attentive and preattentive emergent meter in adult listeners without extensive music training. *Music Perception*, 26, 377–386. <https://doi.org/10.1525/mp.2009.26.4.377>
- LAENG, B., EIDET, L. M., SULUTVEDT, U., & PANKSEPP, J. (2016). Music chills: The eye pupil as a mirror to music's soul. *Consciousness and Cognition*, 44, 161–178. <https://doi.org/10.1016/j.concog.2016.07.009>
- LAENG, B., ØRBO, M., HOLMLUND, T., & MIOZZO, M. (2011). Pupillary stroop effects. *Cognitive Processing*, 12, 13–21. <https://doi.org/10.1007/s10339-010-0370-z>
- LAENG, B., SIROIS, S., & GREDEBÄCK, G. (2012). Pupillometry: A window to the preconscious? *Perspectives on Psychological Science*, 7, 18–27. <https://doi.org/10.1177/1745691611427305>
- LARGE, E. W., HERRERA, J. A., & VELASCO, M. J. (2015). Neural networks for beat perception in musical rhythm. *Frontiers in Systems Neuroscience*, 9, 1–14. <https://doi.org/10.3389/fnsys.2015.00159>
- LARGE, E. W., & JONES, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, 106, 119–159. <https://doi.org/10.1037/0033-295X.106.1.119>
- LARGE, E. W., & SNYDER, J. S. (2009). Pulse and meter as neural resonance. *Annals of the New York Academy of Sciences*, 1169, 46–57. <https://doi.org/10.1111/j.1749-6632.2009.04550.x>
- LEAN, Y., & SHAN, F. (2012). Brief review on physiological and biochemical evaluation of human mental workload. *Human Factors and Ergonomics in Manufacturing and Service Industries*, 22, 177–187.
- LEHMANN, A., ARIAS, D. J., & SCHÖNWIESNER, M. (2016). Tracing the neural basis of auditory entrainment. *Neuroscience*, 337, 306–314. <https://doi.org/10.1016/j.neuroscience.2016.09.011>
- LIAO, H.-I., YONEYA, M., KASHINO, M., & FURUKAWA, S. (2018). Pupillary dilation response reflects surprising moments in music. *Journal of Eye Movement Research*, 11(2), 1–13. <https://doi.org/10.16910/jemr.11.2.13>
- LIBBY, W. L., LACEY, B. C., & LACEY, J. I. (1973). Pupillary and cardiac activity during visual attention. *Psychophysiology*, 10, 270–294.
- LONDON, J. (2012). *Hearing in time: Psychological aspects of musical meter* (2nd ed.). Oxford, UK: Oxford University Press.
- MADISON, G. (2006). Experiencing groove induced by music: Consistency and phenomenology. *Music Perception*, 24, 201–208. <https://doi.org/10.1525/mp.2006.24.2.201>
- MADISON, G., GOUYON, F., ULLÉN, F., & HÖRNSTRÖM, K. (2011). Modeling the tendency for music to induce movement in humans: First correlations with low-level audio descriptors across music genres. *Journal of Experimental Psychology: Human Perception and Performance*, 37, 1578–1594. <https://doi.org/10.1037/a0024323>

- MATHÔT, S., FABIUS, J., VAN HEUSDEN, E., & VAN DER STIGCHEL, S. (2018). Safe and sensible preprocessing and baseline correction of pupil-size data. *Behavior Research Methods*, *50*, 94–106. <https://doi.org/10.3758/s13428-017-1007-2>
- MATSUSHITA, S., & NOMURA, S. (2016). The asymmetrical influence of timing asynchrony of bass guitar and drum sounds on groove. *Music Perception*, *34*, 123–131. <https://doi.org/10.1525/mp.2016.34.2.123>
- 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*, 1–16. <https://doi.org/10.3389/fpsyg.2016.00069>
- MAYFIELD, C., D'ANGELO, FIGUEROA, G., & FOSTER, K. (2014). Really love [Recorded by D'Angelo]. On *Black Messiah* [CD]. New York: RCA Records.
- MCGINLEY, M. J., DAVID, S. V., & MCCORMICK, D. A. (2015). Cortical membrane potential signature of optimal states for sensory signal detection. *Neuron*, *87*, 179–192. <https://doi.org/10.1016/j.neuron.2015.05.038>
- MONSON, I. (1996). *Saying something: Jazz improvisation and interaction*. Chicago, IL: University of Chicago Press.
- MORESI, S., ADAM, J. J., RIJCKEN, J., VAN GERVEN, P. W. M., KUIPERS, H., & JOLLES, J. (2008). Pupil dilation in response preparation. *International Journal of Psychophysiology*, *67*, 124–130. <https://doi.org/10.1016/j.ijpsycho.2007.10.011>
- NÄÄTÄNEN, R., PAAVILAINEN, P., RINNE, T., & ALHO, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, *118*, 2544–2590. <https://doi.org/10.1016/j.clinph.2007.04.026>
- NKETIA, J. H. K. (1974). *The music of Africa*. New York: Norton.
- NOZARADAN, S., PERETZ, I., MISSAL, M., & MOURAUX, A. (2011). Tagging the neuronal entrainment to beat and meter. *Journal of Neuroscience*, *31*, 10234–10240. <https://doi.org/10.1523/JNEUROSCI.0411-11.2011>
- PALMER, C. (1996). On the assignment of structure in music performance. *Music Perception*, *14*, 23–56. <https://doi.org/10.2307/40285708>
- PARTALA, T., & SURAKKA, V. (2003). Pupil size variation as an indication of affective processing. *International Journal of Human-Computer Studies*, *59*, 185–198. [https://doi.org/10.1016/S1071-5819\(03\)00017-X](https://doi.org/10.1016/S1071-5819(03)00017-X)
- PRESSING, J. (2002). Black atlantic rhythm: Its computational and transcultural foundations. *Music Perception*, *19*, 285–310. <https://doi.org/10.1525/mp.2002.19.3.285>
- PREUSCHOFF, K., HART, B. M. 'T, & EINHÄUSER, W. (2011). Pupil dilation signals surprise: Evidence for noradrenaline's role in decision making. *Frontiers in Neuroscience*, *5*, 1–12. <https://doi.org/10.3389/fnins.2011.00115>
- QUIRINS, M., MAROIS, C., VALENTE, M., SEASSAU, M., WEISS, N., EL KAROUI, I., ET AL. (2018). Conscious processing of auditory regularities induces a pupil dilation. *Scientific Reports*, *8*, 1–11. <https://doi.org/10.1038/s41598-018-33202-7>
- RASCH, R. A. (1988). Timing and synchronization in ensemble performance. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition* (pp. 70–90). Oxford, UK: Oxford University Press/Clarendon Press.
- REPP, B. H. (1996). Patterns of note onset asynchronies in expressive piano performance. *Journal of the Acoustical Society of America*, *100*, 3917–3932. <https://doi.org/10.1121/1.417245>
- REPP, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin and Review*, *12*, 969–992. <https://doi.org/10.3758/BF03206433>
- REPP, B. H., & SU, Y.-H. (2013). Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonomic Bulletin and Review*, *20*, 403–452. <https://doi.org/10.3758/s13423-012-0371-2>
- ROHOLT, T. C. (2014). *Groove: A phenomenology of rhythmic nuance*. New York: Bloomsbury Publishing.
- ROZADO, D., & DUNSER, A. (2015). Combining EEG with pupillometry to improve cognitive workload detection. *Computer*, *48*(10), 18–25.
- RUBIN, L. H., WITKIEWITZ, K., ANDRE, J. S., & REILLY, S. (2007). Methods for handling missing data in the behavioral neurosciences: Don't throw the baby rat out with the bath water. *Journal of Undergraduate Neuroscience Education*, *5*, A71–A77.
- SCHWALM, M., KEINATH, A., & ZIMMER, H. D. (2008). Pupillometry as a method for measuring mental workload within a simulated driving task. In D. de Waard, F. O. Flemisch, B. Lorenz, H. Oberheid, & K. A. Brookhuis (Eds.), *Human Factors for Assistance and Automation* (pp. 75–87). Maastricht, NL: Shaker Publishing.
- 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*, 1–16. <https://doi.org/10.3389/fpsyg.2016.01487>
- SIOROS, G., MIRON, M., DAVIES, M., GOUYON, F., & MADISON, G. (2014). Syncopation creates the sensation of groove in synthesized music examples. *Frontiers in Psychology*, *5*, 1–10. <https://doi.org/10.3389/fpsyg.2014.01036>
- SNYDER, J. S., HANNON, E. E., LARGE, E. W., & CHRISTIANSEN, M. H. (2006). Synchronization and continuation tapping to complex meters. *Music Perception*, *24*, 135–146. <https://doi.org/10.1525/mp.2006.24.2.135>
- STUPACHER, J., HOVE, M. J., NOVEMBRE, G., SCHÜTZ-BOSBACH, S., & KELLER, P. E. (2013). Musical groove modulates motor cortex excitability: A TMS investigation. *Brain and Cognition*, *82*, 127–136. <https://doi.org/10.1016/j.bandc.2013.03.003>

- STUPACHER, J., WOOD, G., & WITTE, M. (2017). Neural entrainment to polyrhythms: A comparison of musicians and non-musicians. *Frontiers in Neuroscience*, *11*, 1–17. <https://doi.org/10.3389/fnins.2017.00208>
- VAN DER MOLEN, M. W., BOOMSMA, D. I., JENNINGS, J. R., & NIEUWBOER, R. T. (1989). Does the heart know what the eye sees? A cardiac/pupillometric analysis of motor preparation and response execution. *Psychophysiology*, *26*, 70–80. <https://doi.org/10.1111/j.1469-8986.1989.tb03134.x>
- VAN ZUIJEN, T. L., SUSSMAN, E., WINKLER, I., NÄÄTÄNEN, R., & TERVANIEMI, M. (2005). Auditory organization of sound sequences by a temporal or numerical regularity—A mismatch negativity study comparing musicians and non-musicians. *Cognitive Brain Research*, *23*, 270–276. <https://doi.org/10.1016/j.cogbrainres.2004.10.007>
- VUUST, P., DIETZ, M. J., WITEK, M. A. G., & KRINGELBACH, M. L. (2018). Now you hear it: A predictive coding model for understanding rhythmic incongruity. *Annals of the New York Academy of Sciences*, *1423*, 19–29. <https://doi.org/10.1111/nyas.13622>
- VUUST, P., GEBAUER, L. K., & WITEK, M. A. G. (2014). Neural underpinnings of music: The polyrhythmic brain. In H. Merchant & V. De Lafuente (Eds.), *Neurobiology of Interval Timing* (pp. 339–356). New York: Springer.
- VUUST, P., OSTERGAARD, L., PALLESEN, K. J., BAILEY, C., & ROEPSTORFF, A. (2009). Predictive coding of music - Brain responses to rhythmic incongruity. *Cortex*, *45*, 80–92. <https://doi.org/10.1016/j.cortex.2008.05.014>
- VUUST, P., PALLESEN, K. J., BAILEY, C., VAN ZUIJEN, T. L., GJEDDE, A., ROEPSTORFF, A., & ØSTERGAARD, L. (2005). To musicians, the message is in the meter: Pre-attentive neuronal responses to incongruent rhythm are left-lateralized in musicians. *NeuroImage*, *24*, 560–564. <https://doi.org/10.1016/j.neuroimage.2004.08.039>
- WALLE, K. M., NORDVIK, J. E., ESPESETH, T., BECKER, F., & LAENG, B. (2019). Multiple object tracking and pupillometry reveal deficits in both selective and intensive attention in unilateral spatial neglect. *Journal of Clinical and Experimental Neuropsychology*, *41*, 270–289. <https://doi.org/10.1080/13803395.2018.1536735>
- WALLENTIN, M., NIELSEN, A. H., FRIIS-OLIVARIUS, M., VUUST, C., & VUUST, P. (2010). The Musical Ear Test, a new reliable test for measuring musical competence. *Learning and Individual Differences*, *20*, 188–196. <https://doi.org/10.1016/j.lindif.2010.02.004>
- WARREN, R. M. (1993). Perception of acoustic sequences: Global integration versus temporal resolution. In S. E. McAdams & E. Bigand (Eds.), *Thinking in sound* (pp. 37–68). Oxford, UK: Clarendon Press/Oxford University Press.
- WINN, M. B., WENDT, D., KOELEWIJN, T., & KUCHINSKY, S. E. (2018). Best practices and advice for using pupillometry to measure listening effort: An introduction for those who want to get started. *Trends in Hearing*, *22*, 1–32. <https://doi.org/10.1177/2331216518800869>
- WITEK, M. A. G. (2017). Filling in: Syncopation, pleasure and distributed embodiment in groove. *Music Analysis*, *36*, 138–160. <https://doi.org/10.1111/musa.12082>
- WITEK, M. A. G., CLARKE, E. F., WALLENTIN, M., KRINGELBACH, M. L., & VUUST, P. (2014). Syncopation, body-movement and pleasure in groove music. *PLoS ONE*, *9*(4), 1–12. <https://doi.org/10.1371/journal.pone.0094446>
- WU, E. X. W., LAENG, B., & MAGNUSSEN, S. (2012). Through the eyes of the own-race bias: Eye-tracking and pupillometry during face recognition. *Social Neuroscience*, *7*, 202–216. <https://doi.org/10.1080/17470919.2011.596946>
- ZATORRE, R. J., CHEN, J. L., & PENHUNE, V. B. (2007). When the brain plays music: Auditory-motor interactions in music perception and production. *Nature Reviews Neuroscience*, *8*, 547–558. <https://doi.org/10.1038/nrn2152>
- ZÉNON, A., SIDIBÉ, M., & OLIVIER, E. (2014). Pupil size variations correlate with physical effort perception. *Frontiers in Behavioral Neuroscience*, *8*, 1–8. <https://doi.org/10.3389/fnbeh.2014.00286>
- ZOEFEL, B., TEN OEVER, S., & SACK, A. T. (2018). The involvement of endogenous neural oscillations in the processing of rhythmic input: More than a regular repetition of evoked neural responses. *Frontiers in Neuroscience*, *12*, 1–13. <https://doi.org/10.3389/fnins.2018.00095>