

THERE'S MORE TO TIMING THAN TIME: INVESTIGATING MUSICAL MICRORHYTHM ACROSS DISCIPLINES AND CULTURES

ANNE DANIELSEN & RAGNHILD BRØVIG
RITMO, University of Oslo, Oslo, Norway

KJETIL KLETTE BØHLER
University of Southeastern Norway, Borre, Norway

GUILHERME SCHMIDT CÂMARA,
MARI ROMARHEIM HAUGEN, & EIRIK JACOBSEN
RITMO, University of Oslo, Oslo, Norway

MATS S. JOHANSSON
University of Southeastern Norway, Rauland, Norway

OLIVIER LARTILLOT, KRISTIAN NYMOEN,
KJELL ANDREAS ODDEKALV, BJØRNAR SANDVIK,
& GEORGE SOROS
RITMO, University of Oslo, Oslo, Norway

JUSTIN LONDON
Carleton College & RITMO, University of Oslo, Oslo, Norway

THE TIME PROJECT: TIMING AND SOUND IN Musical Microrhythm (2017–2022) studied micro-rhythm; that is, how dynamic envelope, timbre, and center frequency, as well as the microtiming of a variety of sounds, affect their perceived rhythmic properties. The project involved theoretical work regarding the basic aspects of microrhythm; experimental studies of microrhythm perception, exploring both stimulus features and the participants' enculturated expertise; observational studies of how musicians produce particular microrhythms; and ethnographic studies of musicians' descriptions of microrhythm. Collectively, we show that: (a) altering the microstructure of a sound ("what" the sound is) changes its perceived temporal location ("when" it occurs), (b) there are systematic effects of core acoustic factors (duration, attack) on microrhythmic perception, (c) microrhythmic features in longer and more complex sounds can give rise to different perceptions of the same sound, and (d) musicians are highly aware of microrhythms and have developed vocabularies for describing them. In addition, our

results shed light on conflicting results regarding the effect of microtiming on the "grooviness" of a rhythm. Our use of multiple, interdisciplinary methodologies enabled us to uncover the complexity of microrhythm perception and production in both laboratory and real-world musical contexts.

Received: April 4, 2023, accepted November 4, 2023.

Key words: microrhythm, groove, beat perception, synchronization, interdisciplinarity

WHEN JAZZ GUITARISTS WANT TO PRODUCE more "laidback" sounds, they play with their fingers rather than with a pick. When a producer wants a rhythmically floating feel, s/he chooses a bass sound with a muffled onset over one with a clearly articulated attack. Musicians in all styles and genres make these performance decisions all the time, as they intuitively know that changing the nature of the attack of a sound changes its rhythmic properties, as well as how the sound will blend/align with other sounds in the musical texture. For the examples just given, these choices involve sounds whose moment of occurrence will be later in time than an alternative guitar note played with a pick or a bass sound with a percussive attack. That is, although both the fingered and picked guitar sounds might *physically* start at the same moment in time (as would be seen in an audio file or DAW), they are not *perceived* as starting at the same time.

This review paper summarizes the results and lessons learned from a collaborative, interdisciplinary, and systematically comparative research project (the *TIME* project) on how the microstructure of sounds affects their perceived rhythmic properties on higher levels. This inquiry emerged from an earlier collaborative research project on rhythm and groove in the context of digital music production (Danielsen, 2010), which made it clear that, while onset timing is important regarding the feel and the shaping of groove, it is but one factor. In addition, another motivation was our awareness of musicians' strong interest in as well as their highly developed

language for microrhythmic nuance, which invited more genre-specific investigations of ways of shaping feel and groove beyond onset timing. We assumed that research into such genre-specific interactions of different microrhythmic aspects would also benefit our understanding of what is *not* genre-specific and would illuminate generic aspects/constraints on micro-level auditory perception.

In particular, we became interested in the perceptual interaction between the sonic features or “what” aspects of a sound, such as, attack shape, frequency content, intensity, etc., and the basic perception of the same sound’s “when,” that is, its location and sense of alignment with other sounds. We were also curious about effects of listener’s particular musical enculturation in this regard. Accordingly, the core questions of the *TIME* project were: (a) how do sonic parameters influence a sound’s perceived temporal position, (b) how do sonic parameters influence the tolerance (Johansson, 2010a) for the temporal location(s) of rhythmic events in a beat-based musical context; that is, our sense that two or more sounds are simultaneous, and (c) in metered music, how is the beat shaped and perceived in different music/cultural contexts? We hypothesized that sonic parameters would influence the listener’s perception of temporal relationships at the micro level of rhythm; in short, that the “what” would influence the perception of the “when”—and that it would do so differently in different musical genres.

The project team investigated these questions using a combined multidisciplinary and cross-cultural approach that is unique in research into rhythm and timing. Through perception and performance experiments, qualitative interviews with musicians and producers, and analyses of their music, we compared five musical genres and their corresponding communities of practice for which rhythm is a key aesthetic marker: jazz, samba, electronic dance music (EDM), contemporary R&B/hip-hop, and traditional Scandinavian folk music. Two aspects make these genres particularly suited to systematic, comparative investigation of how sonic parameters influence beat perception. First, a regularly recurring matrix of beats is a basic structure in all of them. Second, although groove is to varying degrees part of the discourse of the different genres, they are all “groove directed” in the sense that their musical patterns and ways of performance are or have been associated with dance and a “pleasurable urge to move” (Câmara & Danielsen, 2018; Janata et al., 2012).

This paper aggregates the principal results and methodological underpinnings of an otherwise dispersed array of published research, allowing for the derivation

of certain higher-level implications.¹ We begin by reviewing and clarifying the core concepts of perceptual centers (P-centers), beat bins, microrhythm, and microtiming (Section 1). Section 2 summarizes the experimental methods used in our perceptual experiments and shows how various acoustic factors give rise to both different perceptions of a sound’s temporal location as well as varying beat bins; that is, degrees of precision when expecting a sound in beat-based contexts. Section 3 summarizes our findings regarding the effect of musical expertise on the perception of microrhythm, and section 4 reviews our findings regarding what sonic aspects performers use when asked to produce different microrhythms on demand, as well as the bodily actions they employ in doing so. Section 5 gives excerpts from our ethnographic research, showing the commonalities and differences amongst the ways different groups of expert musicians describe microrhythms, shedding light on their cognitive representations of both microrhythm and its higher-level musical effects. In our discussion (Section 6) we discuss our findings in light of an embodied perspective on perception and cognition of rhythm, review the implications of the *TIME* project for our understanding of the relationship between the “what” and “when” of rhythmic sounds as well as the relation between microrhythm and groove. We conclude (Section 7) with some reflections on the advantages and the challenges of the project’s collaborative, interdisciplinary, and cross-cultural approach to rhythm research, and we outline some potential paths for future research.

P-Centers, Beat Bins, Microrhythm and Microtiming

The distinction between the acoustic, physical onset, and the perceived timing of a sound has been well studied in the perception of speech, where these locations are known as perceptual centers or “P-centers” (Morton et al., 1976; for a review of subsequent literature see Villing, 2010). P-centers were first noticed when listeners were presented with a series of counting syllables: “one, two, three, four” (etc.) whose acoustic onsets were perfectly isochronous, but they were not perceived as such (Morton et al., 1976), as the different vowel sounds have different rise times and spectral properties. In music the differences in P-centers amongst instruments have also long been known, along with their implications for coordination in ensemble performance (Rasch,

¹ The project *TIME: Timing and Sound in Musical Microrhythm* was funded by the Research Council of Norway and the University of Oslo and ran from 2017 through 2022.

1979; Vos & Rasch, 1981). However, not only is the temporal location of a sound separable from its acoustic or perceptual onset; those locations are also variable in their “width,” as there is a range in which one sound is heard as occurring in synchrony with another sound. Gordon (1987) and Wright (2008) have thus characterized P-centers not as points in time, but as probability distributions that have both a mean/peak as well as a temporal spread (for more recent applications of the same approach, see Danielsen et al., 2019, and Hosken, 2021).

The experience of musical rhythm, which most typically involves repeated patterns of sound, is characterized by an interaction between acoustically sounding events and endogenous reference structures that are activated in the listener (see Bengtsson & Gabrielsson, 1980; Clarke, 1985; Danielsen, 2006; Honing, 2013; Johansson, 2017; Kivite, 2007; London, 2012). Repeated sounds/sound patterns give rise to a basic pulse or beat, the grouping and hierarchically organization of those beats (i.e., meter), segmenting and grouping figural units (style- and song-specific rhythmic figures), and then coordinating the formation of larger sonic and metric structures. The independence of beats and meters from sounding events is manifest in phenomena such as subjective pulse (also called internal beat), subjective accentuation, subjective rhythmization, and the perception of “loud rests” (see London, 2012, for a summary of rhythm-meter interactions). Danielsen (2010, 2018) developed the “beat bin” hypothesis to account for our perceptual response to sounds with different P-center widths; that is, how sounds with different shapes can give rise to different senses of “beat”: Sharp, percussive sounds lead to narrow bins for the perception of beats, whereas indistinct, “muddy” or compound sounds induce considerably wider bins.

The complex interplay between periodic sounds and our endogenous sense of beat is illustrated in Figure 1, which also summarizes how the interpretation of experimental results has evolved in recent years. The uppermost panel (Figure 1a) illustrates an experimental context where the stimulus is a series of very brief sounds (e.g., metronome clicks), and the response is measured as a time point relative to the stimulus (e.g., taps on a drum pad). Given the brevity of the stimulus (a click) and the acoustic profile of the response (impact sounds), both are represented here as points in time. In experiments with this type of stimuli and task, the dependent measure may be the asynchrony between taps and clicks, or the variability of interonset interval between successive taps (see Repp, 2005, Repp & Su, 2013). Variability in responses is regarded as noise, and

differences amongst participants in variability is a measure of differences in their temporal acuity and/or motor control.

A similar result (and approach) is evident in studies that use sounds with longer durations but relatively sharp onsets as stimuli, such as drum strokes or fast-ramped sine tones (Figure 1b). Here the tap placement (or click alignment, in studies using the method of adjustment) is close to the beginning of the sound (though see the discussion of “negative mean asynchrony” in Repp, 2005, and Repp & Su, 2013), such that the acoustic onset is often regarded as the perceptual center and location of the sound. With musical sounds produced by bowing or breathing, the situation is more akin to the determination of P-centers in speech. Violins and voices, for example, involve “softer” attacks (that is, with a longer rise time of their amplitude envelope) and also take some time for the stabilization of pitch, timbre (e.g., vocal formants) and other features such as vibrato. Figure 1c illustrates how responses from a tapping or a click alignment task may be related to such sounds. Responses occur many milliseconds after the initial onset of the sound (and well past the perceptual threshold for the sound). Moreover, if testing the same sounds again, the mean P-center location will probably not be identical to the first trial. And if repeated a third time, yet another location might be the result. At this point we face an epistemic problem. For the data may indicate:

- a) The extent to which a given sound affords a precise temporal location;
- b) The degree of endogenous beat precision of a participant or group of participants;
- c) Or *both* the degree of temporal precision that the sound affords *and* the endogenous precision of the participants’ sense of beat.

Our results, which involve multiple experimental methods as well as a range of stimuli, show that the answer is (c). But it is not that the listeners’ responses are simply “fuzzy” in the context of sounds with slow attacks. Rather, we found that listeners’ endogenous (internal) sense of beat is matched to the temporal affordances of the sounds with which they are engaged, as well as their musical/aesthetic goals in listening to, moving with, and/or performing such sounds. This is shown in Figure 1d, which illustrates the linkage between sounds with slow attack and more complex shapes and the listener’s correspondingly larger/more complex “beat bins.”

Thus, *what* a sound is and *when* it appears to happen cannot be wholly separated, as they are interdependent. This has implications for studies of rhythm and timing

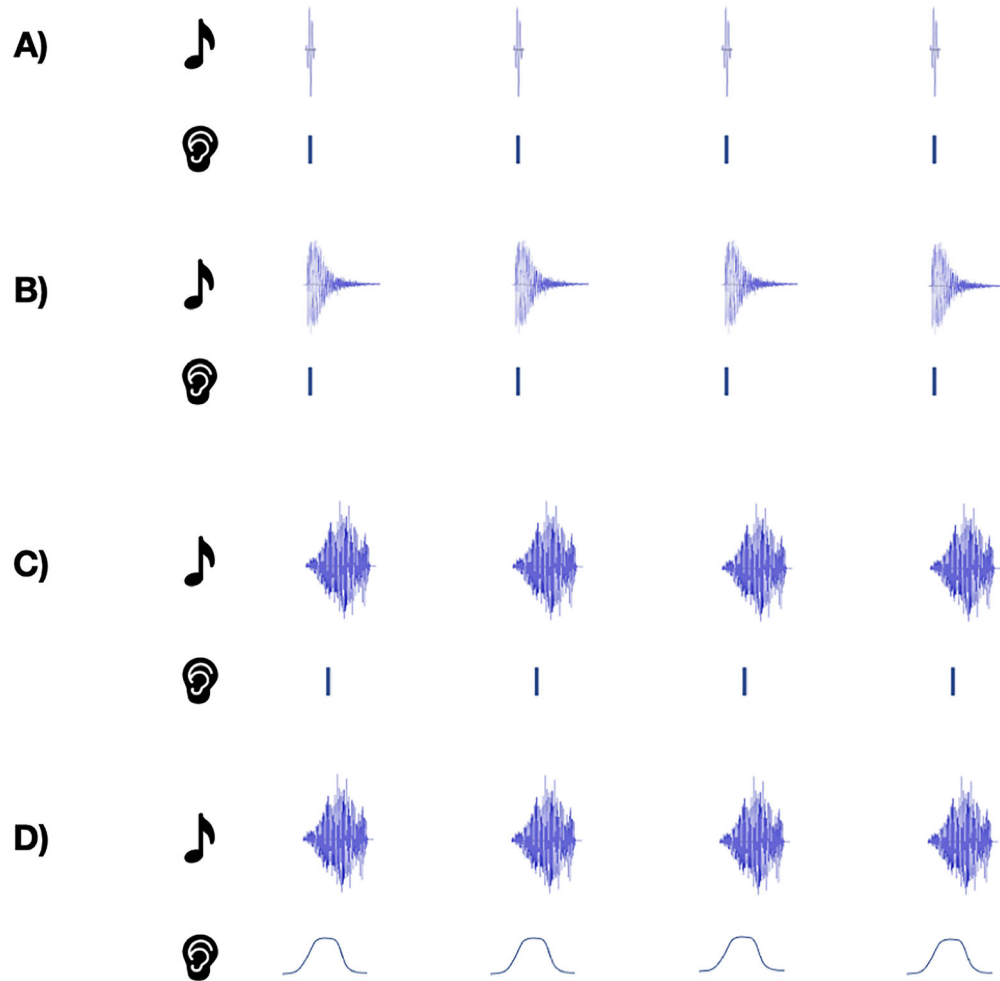


FIGURE 1. Stimuli and corresponding endogenous responses in different experimental conditions. The note indicates the kind of sound used as a stimulus/target, while the ear indicates what is (presumably) perceived: a) clicks and taps, b) sounds with sharp onsets and taps, c) sounds with slow attack/complex shape and taps, d) cumulative mapping of the tapping (or other) responses to sounds with slow attack/complex shape.

beyond that of single notes in (nominally) isochronous experimental contexts. As noted by Bengtsson and Gabrielsson (1983), patterns at the micro level of rhythm can be either idiomatic and systematic; that is, a structural feature, or expressive and varied. While terminology is not consistent, the former has often been denoted *microtiming* or *swing* (Butterfield, 2011; Iyer, 2002; Madison et al., 2011) and the latter *expressive timing* (Clarke, 1985, 1989).² When the *TIME* project started in 2016, research addressing microtiming in African American groove-based musics and expressive timing

² Madison et al. (2011, p. 1579) make a similar distinction between systematic (repeating) and unsystematic (non-repeating) varieties of microtiming. However, the latter category can in principle include both intentional (that is, expressive) and random microtiming (that is, noise).

in European art music had begun to increase in number and scope (see, for example, Bengtsson & Gabrielsson, 1983; Butterfield, 2010, 2011; Clarke, 1985, 1989; Desain & Honing, 1989; Friberg & Sundström, 2002; Iyer, 2002; Keller, 2014). Numerous studies have since examined the nature and role of microtiming in more diverse musical cultures.³ Still, only a few of these specifically

³ Some examples are Alén (1995) on Tumba Francesca; Clayton (2000) on North Indian Raga; Gerischer (2006) and Haugen & Danielsen (2020) on samba; Jankowsky (2013) on Tunisian Stambeli; Johansson (2010a, 2010b) and Kvifte (2004, 2007) on Scandinavian folk music; Polak (2010) and Polak & London (2014) on Malian Jembe music; Berliner (2009), Doffman (2009), Hodson (2007), Monson (1996), and Prögler (1995) on jazz; Danielsen (2006) on funk; Stover (2009) on salsa; and Bjerke (2010), Danielsen (2010, 2012), and Zeiner-Henriksen (2010) on neo-soul, disco, and electronic dance music.

address the relationship between dimensions such as the shape, timbre, and intensity of the individual elements within a rhythmic pattern and the perception and production of those elements' precise temporal locations in music (see Butterfield, 2011; Danielsen et al., 2015; Hofmann et al., 2017).

Comparing these different ways of thinking about the endogenous response to musical beats can help clarifying the relationship between microtiming and microrhythm: *Microtiming* refers to systematic patterns of onset timing, often with reference to an expected beat position (that is, early, late, etc.). *Microrhythm* is a more encompassing term that refers a range of sub-tactus musical features, as well as their interactions, and is paralleled by an endogenous reference structure that has both width and shape (cf. panel d in Figure 1). Put differently, in addition to microtiming's focus on the sound's "when," microrhythm takes into consideration a variety of additional features related to the sound's "what": attack (sharp or gradual?), duration (short or long?), decay (rapid or gradual?), pitch (high or low?), timbre (bright or dark?), and relative intensity. Attack, duration, and decay are aspects of the *shape* of a sound; that is, the distribution of energy over time or the sound's amplitude envelope. Relative intensity, on the other hand, refers to the overall energy of an event—how loud it is in relation to other rhythmic events. Aspects related to the spectral envelope of the sound, such as spectral centroid, pitch, and *timbre*⁴ can also play a role (Danielsen et al., 2019; Hove et al., 2007; Seton, 1989). Focusing on microrhythm instead of microtiming means to widen the focus and study how all these aspects (including timing) in various combinations may produce a wide variety of rhythmic feels (i.e., laid-back, pushed, tight, loose, and so on). Even though its different aspects might be difficult to distinguish at a perceptual level, their physical constituents can still be measured in the signal and analyzed after the fact.

The fact that the location and precision of the endogenous pulse response varies with the incoming sounds has implications at two levels. First, identifying patterns of physical onset timing seems to be but a first step towards identifying microtiming patterns *as perceived*. (The exception might be soundscapes made up of click-

like or sharp-attack sounds, as discussed above.) Second, as the conflicting results from groove research show (for reviews, see Câmara 2021, Chapter 2; Câmara & Danielsen, 2018; Etani et al., 2023; Malone, 2022), it is an open question whether timing patterns alone can account for a groove's characteristic microrhythmic feel and produce the related "pleasurable urge to move" (Janata et al., 2012; Madison, 2006). The fact that producers and musicians invest a lot of energy in shaping and talking about how they articulate the micro level of their music may point to microrhythmic dimensions beyond onset timing also being important for why and how certain groove feels come across as so irresistible. We will revisit both in the discussion of our results below.

Finally, it is well documented that the endogenous reference structures being activated in listeners depend on their musical enculturation—that is, the extent of their exposure to and experience with certain styles of music and their characteristic rhythmic organizations (see, for example, Hannon, 2010; Hannon et al., 2012; Polak et al., 2018). A starting point for the present project was the assumption that different musical genres are characterized—alongside differences at the level of macro-rhythmic and metrical structure—by characteristic and enculturated differences at the micro or beat level of rhythm.

Mapping the Shape of the Beat Bin

Even a fairly simple sound (i.e., one without many noise components or vibrato, a stable F0, etc.) provides many potential cues for its P-center, as shown in Figure 2. As Nymoen et al. (2017) noted, these descriptors encompass both physical/acoustic attributes of the sound (here mainly relative to its RMS envelope) as well as perceptual attributes (e.g., perceptual onset and perceptual attack, analogous to the P-center location). Most acoustic analyses for onset detection and (at least implicitly) the temporal location for sounds (i.e., in Music Information Retrieval [MIR] and signal processing contexts) have focused on the attack portion of the sound, attempting to locate the perceptual attack/P-center somewhere between the physical onset of the sound and its energy peak. Nymoen et al. (2017) compares the MIR toolbox and Timbre Toolbox onset functions with "ground truth" for P-center location obtained via a perceptual experiment for a range of sounds. As Nymoen et al. (2017) note, P-centers cannot be measured directly, but rather must be estimated by comparing the alignment of a target sound with another sound with a short duration (a "probe"). Moreover, as noted above,

⁴The physical correlate to perceived timbre is not straightforward, and specific dimensions of the timbre space may depend on the sound in question (Grey, 1977; Lakatos, 2000; McAdams & Giordano, 2008). In the context of this article, we refer to timbre as all the components of timbre that are not directly related to attack / rise time, separating the purely temporal and the mainly spectral aspects of the timbre space.

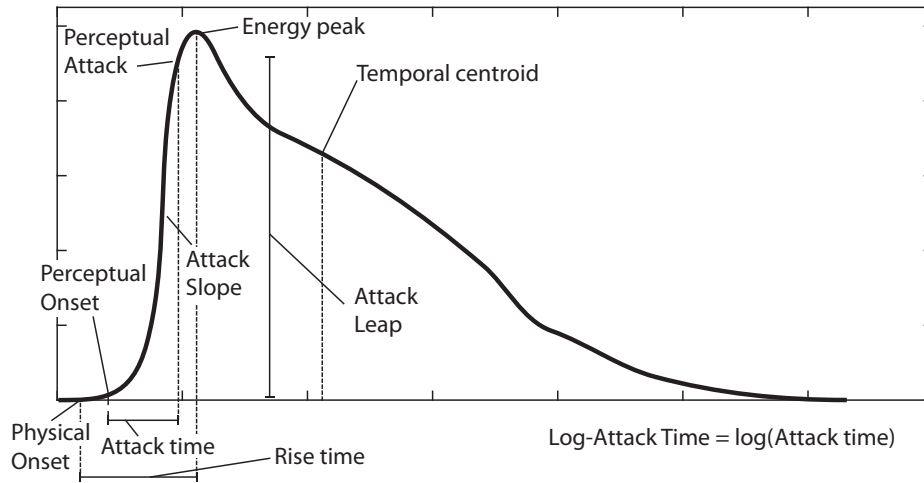


FIGURE 2. Terminology/Descriptors for various portions of the amplitude envelope of a sound, from Nymoer et al. (2017).

TABLE 1. Stimuli Used in Danielsen et al. (2019) and London et al. (2019)

Stimulus Parameters	Click	Noise*	Fast Short Low	Fast Short High	Fast Long Low	Fast Long High	Slow Short Low	Slow Short High	Slow Long Low	Slow Long High
Instrument			Kick Drum	Snare Drum	Dark Piano	Light Piano	Arco Bass	Cabasa	Synth Bass	Fiddle
Attack	0 ms	Slow	Fast	Fast	Fast	Fast	Slow	Slow	Slow	Slow
Duration (ms)	1	100	80-130	25	487	318	66	49	220	105
Frequency range	High	High	Low	High	Low	High	Low	High	Low	High
Pitch where relevant (Hz)	3000	Bandpass filter centered at 3000			65,4	659,3	65,4		32,7	479
Spectral Centroid (Hz)	3720	4809	780	2831	623	893	538	8199	781	1581

*Noise was not used as probe in Danielsen et al. 2019.

P-centers are not precise time points, but rather a probability distribution that occupies some region of the attack portion of the sound.

London et al. (2019) summarizes a series of experiments that explored various methods for investigating the P-centers for a set of musical sounds that were systematically varied in their attack (slow versus fast attack time), duration (long versus short), and center frequency; see Table 1. The target sounds were presented in “looped” fashion (600 ms ISI), and thus the context evokes both the sense of beat in the listener/participant as well as an isochronous interval from P-center to P-center in the stimulus.

As in prior P-center experiments, we used the method of adjustment, in which participants aligned a probe sound (either a click or a short noise burst) with the target sound, and both in-phase and antiphase alignments were used. In addition, we also used a tapping task in which participants tapped a set of clave sticks in synchrony with the looped target sound. For each method, the dependent variables were: (a) the mean P-center location found for each stimulus type, and (b) the variability of the mean P-center location found for each stimulus type. Figure 3 summarizes the results of the in-phase click, anti-phase click, and tapping tasks. Our main take-aways from these methodological studies were:

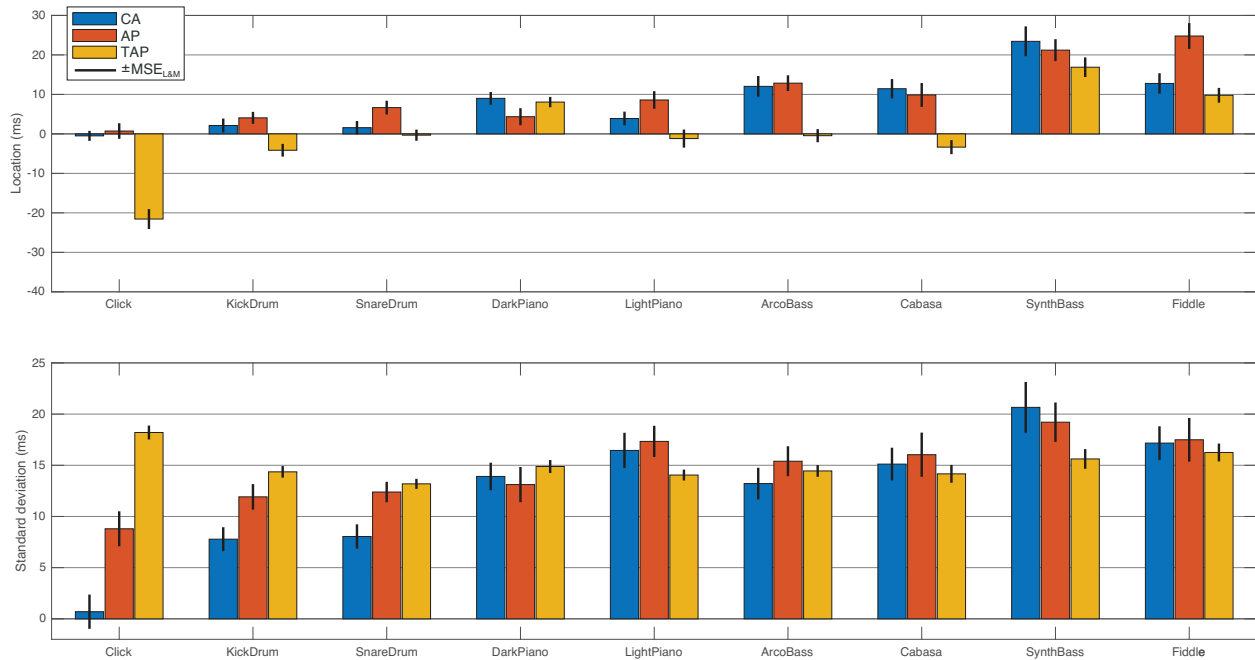


FIGURE 3. Click alignment (CA), anti-phase click alignment (AP) and tapping (TAP) results from London et al. (2019); upper panel shows the mean location of participant responses for each target sound; lower panel shows the standard deviation of those responses, a measure of beat-bin width.

TABLE 2. Artificial Stimuli Used in Danielsen et al. (2019), Experiment 2

	Click	Fast Short High	Fast Short Low	Fast Long High	Fast Long Low	Slow Short High	Slow Short Low	Slow Long High	Slow Long Low
Attack	0 ms	Fast	Fast	Fast	Fast	Slow	Slow	Slow	Slow
Rise Time (ms)		3	3	3	3	50	50	50	50
Duration (ms)	1	100	100	400	400	100	100	400	400
Center frequency (Hz)	3000	700	100	700	100	700	100	700	100

Note: Fast = fast attack, Slow = slow attack, Short = short duration, Long = long duration, Low = low center frequency of the passband, and High = high center frequency of the passband.

- In-phase and anti-phase methods of adjustment using clicks produce nearly identical results, and hence in-phase alignment used in subsequent experiments.
- Tapping vs. click alignment can provide different yet useful information regarding P-center locations:
 - The method of adjustment is sensitive to different sounds in terms of variability while tapping is not.
 - The tapping task involves perception-action synchronization, and thus may involve different mechanisms.
- Using filtered noise as an alignment probe yields consistently earlier probe-onset locations in comparison to using a click as a probe, which means that alignment tasks inherently involve the alignment of P-centers, not onsets.

Danielsen et al. (2019) presents further analysis of the data from London et al. (2019), along with results of a companion experiment that replicated the London et al. (2019) results using a set of wholly artificial stimuli generated from bandpass filtered noise (see Table 2). The experiment with artificial stimuli used the same 2 x 2 x 2 factorial design (fast vs. slow attack/rise time;

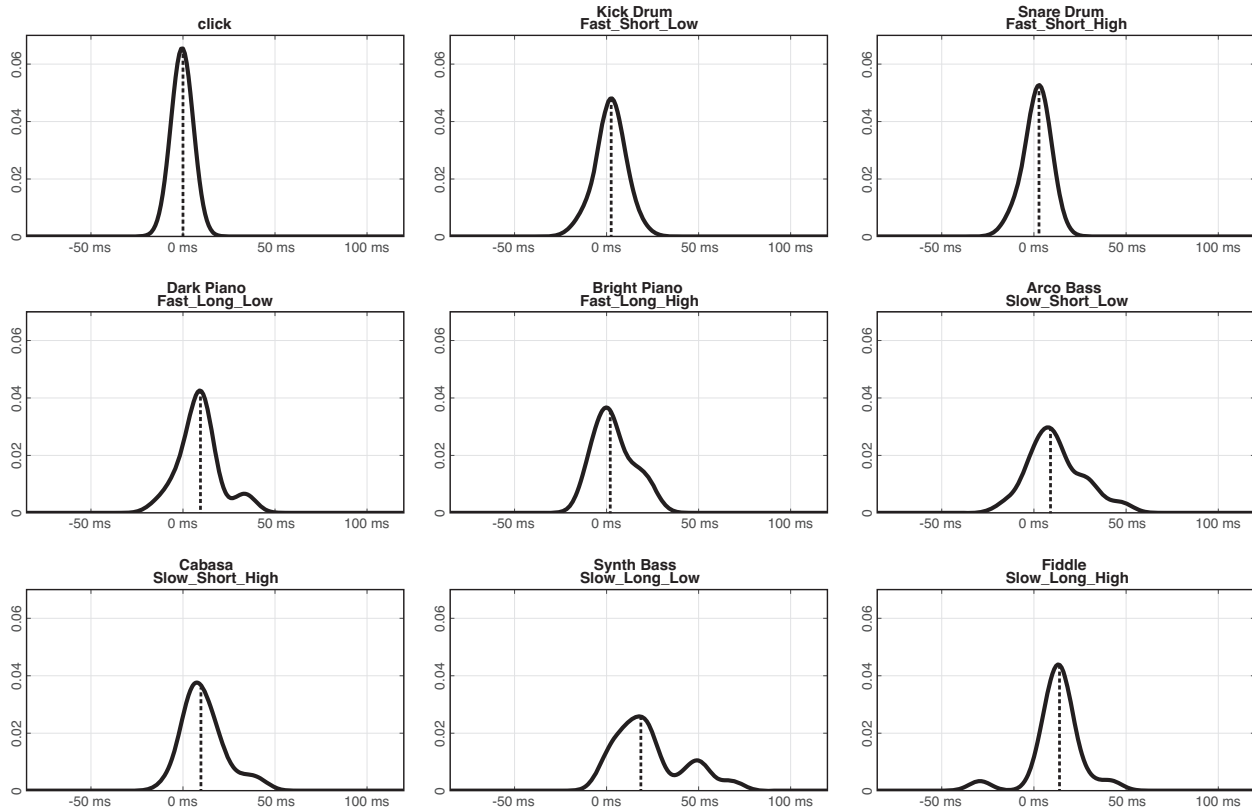


FIGURE 4. Probability density distributions (probability/time) of participant responses for each musical sound used in Danielsen et al. (2019), click alignment task. Descriptors for each sound refer to attack (fast vs. slow), duration (short vs. long), and center frequency (high vs. low). Median indicated by vertical stippled line.

short vs. long duration; and low vs. high center frequency of the passband), with the aim of eliminating any familiarity effects that might obtain with the musical stimuli, as well as to provide a more precise differentiation of these acoustic factors.

The main findings across both experiments were as follows:

- Slow attack and long duration both lead to a later P-center location, but duration has less effect when the attack is fast.
- Low center frequency leads to later P-center location only for musical sounds, and primarily for longer sounds with slow attack.
- Slow attack and long duration also lead to greater variability in the location of the P-center; that is, to wider beat bins.

Danielsen et al. (2019) also presented more detailed/fine-grained portraits of the beat bins for each of the stimuli used in the experiment. As can be seen in Figure 4, which gives the probability density of all

participant responses in the click alignment task, the distributions for most sounds are not symmetrical about their means. The probability density distributions display a systematic pattern of different beat bin shapes, with the combination of slow attack and long duration leading to the flattest shape, indicating a wider tolerance/broader beat bin. Nonparametric statistical tests confirmed this pattern. Slow attack and long duration also produced distributions with complex shapes that suggest these sounds afford multiple locations for beat placement, especially the synthesized bass sound, which has slow attack, long duration, and low spectral centroid.

These results also help to untangle the epistemic problem noted above; that is, how to interpret the variability in participant responses. The characteristic distributions for various classes of stimuli show that the variability in participants' responses is not simply a matter of location + noise, with some sounds leading to noisier responses than others. For while that may seem to be the case with sounds that are short and have fast

onsets (the click and drum sounds), the sounds with longer durations and/or slow onsets have characteristic patterns of skew and kurtosis, and some (the dark piano and the synth bass) have bimodal distributions of participant responses.

The Effect of Expertise

Having established that P-centers/beat bins may vary based upon a systematic combination of acoustic factors, Danielsen et al. (2021) explored the extent to which a listener's musical background affects P-center perception, especially for complex sounds. For this experiment, we recruited musicians with particular expertise in three distinct music genres: Scandinavian traditional fiddle music, jazz, and electronic dance music (EDM)/hip-hop. The fiddlers and jazz musicians were all performers, while the EDM/hip-hop experts were producers who work primarily in a recording studio context. In other words, the fiddlers and jazz musicians shape their microrhythms by varying the articulation, dynamics, and timbral shading in performing on their instruments, while the producers alter these characteristics via the manipulation of audio or MIDI tracks in a DAW environment.

We asked all of them to perform the click alignment and tapping tasks as in our previous experiments, but with a set of sounds that related to each of their musical genres: an acoustic kick drum and electric bass (from jazz), two fiddle sounds (for the Norwegian folk musicians), and a set of synthesized sounds (for the EDM and hip-hop producers). These sounds were distributed across a 2 x 2 factorial design that crossed fast versus slow attack with long versus short duration (the effect of center frequency was not assessed in this experiment; see Table 3). In

addition, we included a set of genre-neutral noise sounds, a subset of the notched noise sounds used in previous experiments.

We found that Genre expertise showed a main effect on both P-center mean location, $F(2, 56) = 9.626$, $p < .001$; $\eta_p^2 = .256$, and variability, $F(2, 56) = 7.964$, $p = .001$; $\eta_p^2 = .221$. Average P-center locations were 26 ms after stimulus onset for producers, 37 ms for jazz musicians, and 40 ms for folk musicians. Pairwise comparisons were significant for producers and jazz musicians ($p = .005$) and producers and folk musicians ($p = .001$); the difference between folk and jazz musicians was not significant. Average P-center variabilities were 15 ms for the producers, 18 ms for the jazz musicians, and 22 ms for the folk musicians. The difference in variability between the producers and Folk musicians was significant ($p = .001$); no other differences were significant. Tellingly, there were no significant differences with respect to P-center location amongst the three participant groups for either the neutral sounds or the electronic sounds; there were small (4–5 ms) but significant or close to significant differences in variability between the folk musicians and the jazz musicians and producers, respectively, showing a higher overall variability for the folk musicians.

The differences between the three participant groups were most pronounced with the organic sounds, most especially the long fiddle sound. Figure 5 illustrates the mean P-center locations for each of the three participant groups in relation to the waveform of the long fiddle sound, and Figure 6 gives histograms of the distribution of all click trials with the long fiddle sound for each of the three expert groups, giving a more fine-grained picture of their responses. It shows that a tri-modal distribution of P-center locations may be latently present in all three groups. While we do not have a large enough

TABLE 3. Sounds Used in Danielsen et al. (2021)

	Click	Fast Short	Fast Long	Slow Short	Slow Long	Fast Short	Fast Long	Slow Short	Slow Long
Sound		Electronic	Electronic	Electronic	Electronic	Organic	Organic	Organic	Organic
Instrument		808 Kick drum	Synth bass	Synth bass	Synth bass	Acoustic kick drum	El Bass	Fiddle	Fiddle
Attack	0 ms	Fast	Fast	Slow	Slow	Fast	Fast	Slow	Slow
Rise Time (ms)		1	1	≈ 74	≈ 122	13	22	≈ 168	≈ 226
Duration (ms)	1	238	519	208	534	180	493	306	589
Pitch (Hz)			65.4	65.4	65.4		55.0	349.2	349.2
Spectral Centroid (Hz)	3532	314	173	313	298	581	406	2317	2405

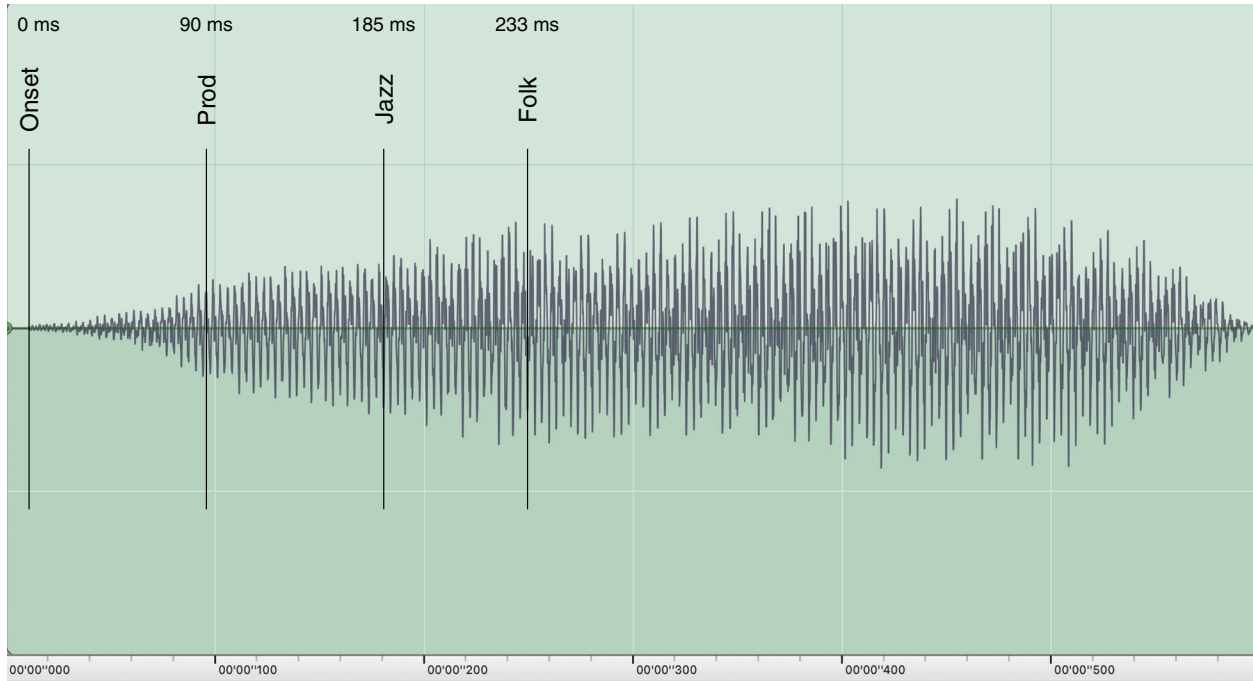


FIGURE 5. Waveform of the long fiddle sound stimulus, showing the mean P-center locations for each of the three participant groups, from Danielsen et al. (2021).

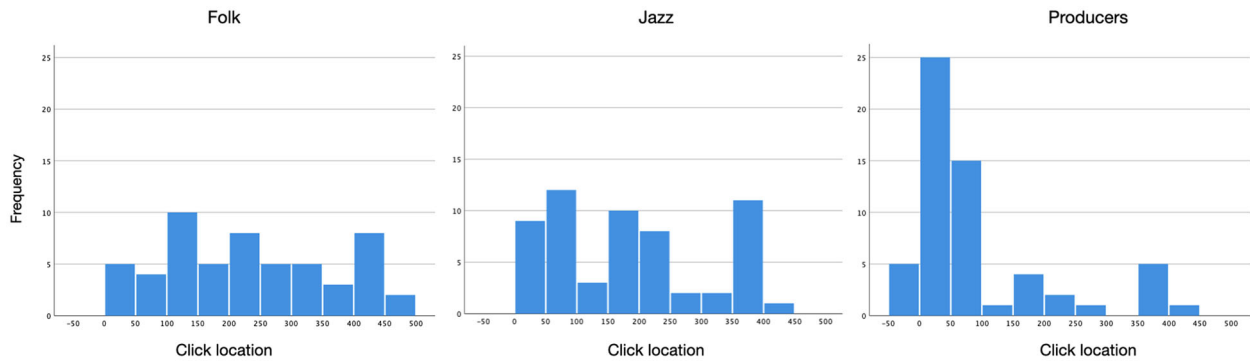


FIGURE 6. Histograms of the distribution of click alignment task responses to the long fiddle sound for each of the three participant groups in Danielsen et al. (2021).

sample to establish multi-modal distributions in our participant sub-populations, as can be seen in Figure 6, the locations of each modal peak correspond to clear inflection points in the amplitude envelope of the sound. One of our initial hypotheses was that the musicians would be most accurate when synchronizing to sounds from their own genres. Interestingly, however, the folk musicians showed greater variability when synchronizing to fiddle sounds from their own genre in comparison to their synchronization with other genres. The extraordinary wide and complex beat bins

we found in response to the long fiddle sound may be related to the aesthetic ideal of performing with flexible timing in Scandinavian fiddle music, as well as broader differences between participants who approach sounds from a performance versus a production mode. In sum:

- Expertise has an effect on what seems to be general, low-level perceptions of sounds, as evidenced by the differences in P-center variability/beat bin width for the neutral sounds.

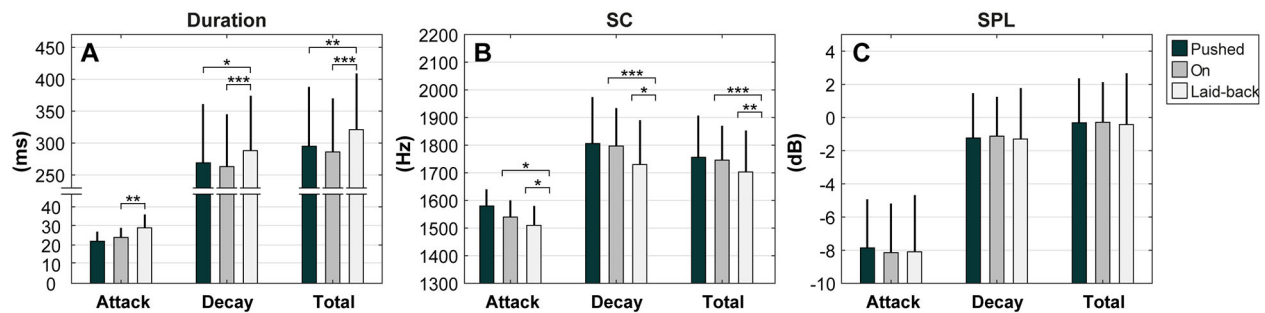


FIGURE 7. Average duration, spectral centroid (SC), and sound pressure level (SPL) of guitar stroke segments across all participants ($N = 21$). All values and error bars represent mean values and one SD , with the exception of attack duration and attack SC, representing median values and median absolute deviation. $*p < .05$; $**p < .01$; $***p < .001$, from Câmara, et al. 2020a.

- Expertise has an effect on how sounds are heard/grasped in terms of their affordance(s) for action/synchronization, as evidenced by the P-center results for organic sounds.
- Expertise has an effect as top-down influence on bottom-up processing in terms of activating genre-specific timing ideals, as evidenced by the P-center and variability results for the long fiddle sound typical of the Scandinavian fiddle music tradition.

Performing Microrhythm

To investigate the extent to which acoustic features other than onset timing are used in the production of microrhythms, Câmara et al. (2020a, 2020b) conducted a series of performance experiments in which expert rhythm-section instrumentalists (drums, guitar, or bass) were instructed to play simple patterns with different microrhythmic “feels,” e.g., in an “on-the-beat,” “laid-back,” or “pushed” manner relative to an external timing reference; that is, either a metronome and/or backing track consisting of the other rhythm section instruments (e.g., bass and guitar for the drums). Data from these experiments included both audio recordings of each musician’s performance, as well as motion capture imaging of their bodily movements. Audio of individual performances were recorded, and time points calculated based on algorithms from the MIR Toolbox (version 1.8) audio analysis package (Lartillot et al., 2008). For the analysis of attack, we developed a new, more precise approach that detects the attack region directly from the audio waveform (Lartillot et al., 2021). The results show that while onset (and/or peak) timing manipulation was the primary cue for creating the different rhythmic feels, musicians also systematically manipulated intensity (sound-pressure level

[SPL]) and/or frequency components (spectral centroid [SC]) of their sounds. Guitarists tended to utilize longer stroke durations (in both attack and decay) and lower brightness (SC), controlled by how hard the different strings are hit, in addition to later onset timing to achieve “laidback” performances. The results of the guitarists’ strategies are summarized in Figure 7.⁵ Bassists utilized greater stroke intensity (SPL) in addition to early onset timing to achieve a “pushed” feel (Câmara, et al., 2020a). Drummers tended to play strokes both earlier/later and with greater dynamic accentuation to distinguish pushed (hi-hat) and laidback (snare) from on-the-beat (synchronous) performances, respectively (Câmara, et al., 2020b).

Musicians’ bodily actions both generate and modify the sound. Research has demonstrated that knowledge of such sound-producing actions is relevant also to the *perception* of the sounds—that is, sound perception implies an understanding of the actions that the listener associates with the production of the sound (e.g., Clarke, 2005; Cox, 2016; Godøy, 2010; Liberman & Mattingly, 1985; Wilson & Knoblich, 2005). While we often observe a musician’s sound-producing actions, other visual cues, such as a performer’s body language, may inform the viewer/listener’s understanding of the underlying metrical structure and associated

⁵ We also looked at the range of individual performance strategies among the drummers, developing a novel interdisciplinary method (Sioros, Câmara, & Danielsen, 2019) that combines fundamental digital signal-processing techniques and music perception principles with statistical methods from bioinformatics (Clarke et al., 2008). The method captures the microtiming relations of the kick, snare, and hi-hat drum onsets among one another. The unique combination of these features in a performance is its “microrhythmic fingerprint.” The clustering results were visualized as phylogenetic trees and present a set of archetypical drumming strategies for each intended timing style (for details, see Câmara et al., 2022).

microrhythms of the music (Blom, 1981; Broughton & Stevens, 2009; Kilchenmann & Senn, 2015; Toivainen et al., 2010). Similarly, in genres where music and dance are closely related, cues for the metric structure are present in the dance (Haugen, 2016, 2021).

Given this project's focus on microrhythm, we were particularly interested in relationships between body motion and playing with a particular timing feel (pushed, laidback), which is related to how individual rhythmic events are shaped by the performer. To this end, we used an infrared motion-capture system consisting of reflective markers attached to the participants' bodies and instruments and multiple cameras surrounding them, ultimately producing a three-dimensional representation of the musician's bodily movements. Câmara et al. (2023) found that laidback strokes were played with a lower [hand/arm] velocity and longer movement duration compared to on-the-beat strokes. This relation corresponds well with the audio-feature results, wherein laidback strokes were found to have slower attacks/longer durations and on-the-beat/pushed strokes had faster attacks/shorter durations. Likewise, Haugen et al. (2023) showed that the performers tended to lean forward when playing pushed as opposed to playing with a laid-back or on-the-beat feel; the difference in posture, while small (1.5–2.0 degrees) is nonetheless biomechanically significant in the context of playing. More broadly, these results show that performers' body posture can be related to their intended timing.

Musicians' Awareness of Microrhythm and Microtiming

As noted above, musicians, recording engineers, and music producers all pay a great deal of attention to the details of microrhythm and microtiming. Knowing this, in the first stage of our project we conducted interviews with expert performers and producers in five selected musical genres (see Table 4). We used a semi-structured interview guide that focused on both general considerations regarding microrhythmic aesthetics in their respective genres and on the specific ways in which the interviewees approached the microlevel temporal and sonic features in their performance practices. Each interview (23 in total) lasted around an hour, and we adjusted the terminology to fit each genre. For details regarding methodology and interviewee selection, see Danielsen et al. (2023).

All our interviewees were concerned with both the shaping of individual sounds, as well as when they should be played/placed relative to other sounds, both

TABLE 4. Overview of Interviewees by Genre and Instrument, from Danielsen et al. (2023)

Genre	<i>n</i>	Instruments/roles
EDM	5	Producers
Hip-Hop	5	Producers
Jazz	6	Vocals, Trumpet, Saxophone, Guitar, Bass, Drums
Samba	5	Vocals, Percussion, Guitar, Drums
Scand. Folk	4	Hardanger Fiddle, Langeleik, Jew's Harp

successively and simultaneously. For example, the jazz guitarist we interviewed said that he might play with his fingers rather than a pick to produce sounds with a softer attack, and that those sounds seem to occur later in time than a sharper sound produced by a pick. Similarly, the jazz drummer adjusted the timbre and the rise time of the sounds of the drum kit via different grips on the sticks and/or striking the cymbal or toms in a specific place. Within the discourse among the jazz musicians, there was a lot of emphasis on the mastery of “time” in the sense of knowing when to play a sound in relation to the collective pulse of the ensemble (Jacobsen & Danielsen, 2023). Accordingly, they stressed that a sharper sound behaves differently than a softer or more muffled sound, and also that timbre in itself can be the basis for the rhythmic flow. A sharp and fast sound produced by a hi-hat, for instance, requires a more precise onset in relation to the perceived pulse than the sound of a double bass, which also needs to be played earlier because of its longer rise time.

Many interviewees thought that sounds with a slow/soft attack have a higher tolerance for alternative temporal positionings that nonetheless appear to be “in time”—that is, they have wide beat bins (Danielsen, 2010, 2018). The folk musicians described soft and slow sounds (or sounds with “secret attacks,” as fiddler Anne Hytta put it) as temporally ambiguous events. In contrast, sharp and fast sounds implied relatively unambiguous rhythmic placement and could be used to highlight an attention-worthy event. The associated balance between ambiguity and clarity was considered an essential part of the rhythmic aesthetic of traditional fiddle music (see also Johansson, 2022). The Scandinavian folk musicians also commented extensively, albeit in more general terms, on the relevance of sound to groove. In the words of fiddler Anne Hytta: “A good fiddle sound is rich and resonant and sharp at the same time, [which] allows for a more differentiated articulation where you mark certain notes more and others less . . . The opposite is a more diffuse sound, which I associate with music that doesn't quite groove.”

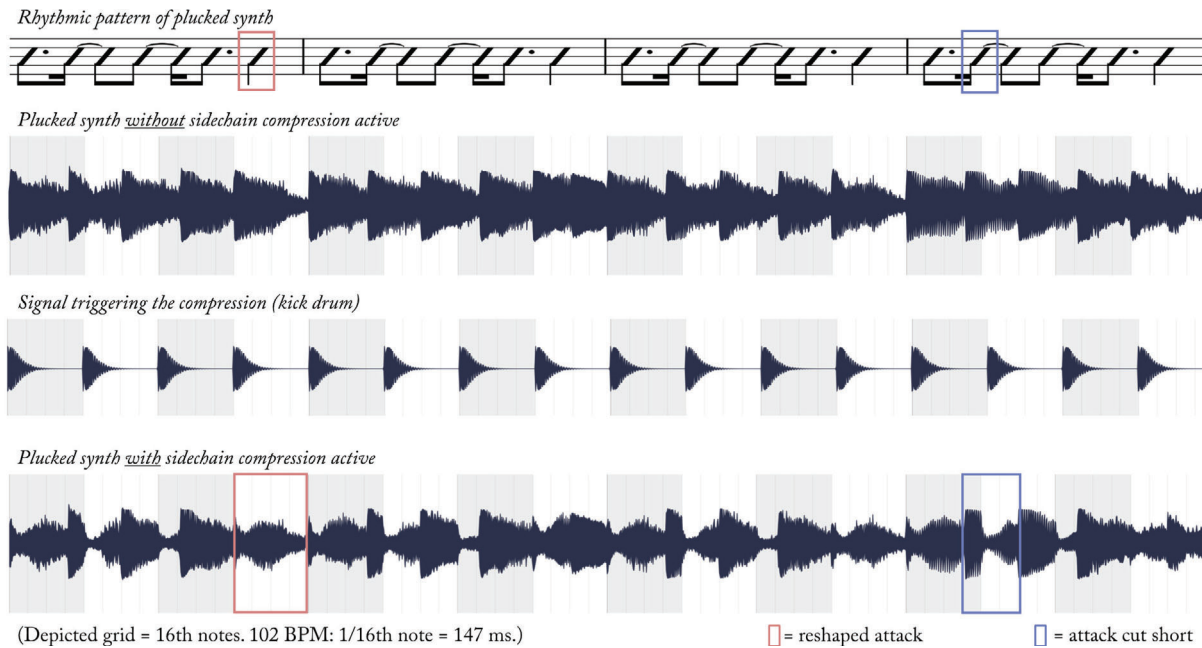


FIGURE 8. Transcription and waveform representations of the plucked synth illustrating the effect of sidechain compression in Seeb’s remix of “I Took a Pill in Ibiza” (0:10–0:29), from Brøvig et al. (2021). The grid of the DAW is marked by alternating white and grey sections. The red rectangle shows how the sidechain compression “ducks” the attack of the plucked synth, reshaping it from sharp to soft. In the blue rectangle the compression reshapes the envelope of the plucked synth.

The EDM and hip-hop producers manipulated the sounds’ envelopes directly, adjusting attack characteristics, as well as overall intensity and dynamics, using filters, volume faders, and sidechain compression. For example, the EDM producer duo Seeb explained that they often used sidechain compression to create dynamic swells off the beat, resulting in a slightly off-the-grid timing pattern (see Figure 8).

The EDM and hip-hop producers also often manipulated the temporal placement/alignment of sounds to “color” the overall sound of their otherwise grid-based grooves, introducing more or less subtle frictions between rhythmic events in a variety of ways. The hip-hop producer Kvam, for example, usually moved the hi-hat attacks slightly behind the snare or kick, and Kholebeatz often delayed the playback of individual tracks by a few milliseconds. A general trend involved placing samples and MIDI events such that they would not be perceived as completely simultaneous, although there was significant variation among substyles regarding this operation, ranging from loose and behind in boom-bap rap to strictly quantized in trap. The producers gave us access to multitrack files of their music, which allowed us to investigate their use of these techniques in practice. In the EDM tracks, we identified the

temporal deviations from the grid (asynchronies in the range of 5–30 milliseconds) caused by sonic manipulation of individual tracks (for analyses of the EDM tracks, see Brøvig-Hanssen et al. (2020) and Brøvig-Hanssen et al. (2021); for analyses of the hip-hop tracks, see Oddekalv 2022a, 2022b).

While many of the performers and (especially) producers we interviewed were quite specific in their descriptions of the ways in which they manipulated the temporal and sonic features of their music, their general discourses about groove were informed by a holistic view of microrhythm, and they tended to talk about groove using bodily and movement-related metaphors. The jazz musicians reflected on how both temporal and sonic features contributed to swing, feel, and drive. The EDM producers appreciated the sense of motion and “breathing” in their grooves. The hip-hop producers consistently referred to how groove manifests itself in movement, so they try to make music “that’s impossible not to nod your head to,” said Kvam. They often used metaphors related to viscosity when describing a certain friction or pushback in a good hip-hop groove. A holistic discourse of groove was also very evident among the samba performers, where terms like *balanço* (balance or swing), *brincadeira* (play), *molho* (sauce), *sabor*

(flavor), suinge (swing), and ola (wave) came up frequently. Among the Norwegian fiddlers, as well, rhythmic qualities were largely identified by means of movement metaphors such as lift, drive, flow, breathing, energy, forwardness, balance, relaxing, and resting.

Discussion

EMBODIED PERCEPTION AND COGNITION OF (MICRO)RHYTHM

Over the last decade or so the conceptual framework of embodied cognition has had an increasing influence in music psychology and music theory (Clarke, 2005; Cox, 2016; Godøy, 2010; Kozak, 2019, 2023; Leman, 2007; Leman & Maes, 2014). Microrhythm and groove is a prime example in this regard and is often approached via bodily metaphors and associated with bodily feelings in literature from both ethnographic (i.e., Berliner, 2009, pp. 349–352; Monson, 1996, pp. 26–29), music-philosophical (i.e., Danielsen, 2006; Roholt, 2014; Witek, 2017), and music-psychological (i.e., Janata et al., 2012; Madison, 2006; Senn et al., 2018) research. Accordingly, one might expect that our awareness of microrhythm, insofar as we have an awareness of sonic/acoustic microstructure and/or microtiming, manifests first and foremost in an awareness of our own pleasure, movements and gestures as we move along with the music (hence the ubiquitous use of the term “feel” to characterize/distinguish musical rhythms, see, for example, discussion on James Brown’s use of the word in Danielsen, 2006, Chapter 10). Paraphrasing Mariusz Kozak we could say that our perception of musical rhythm in general involves our implicit *kinesthetic knowledge* about “how music goes” (Kozak, 2019, p. 5). By actively involving our sensorimotor system—through overt or covert/mentally simulated action—we are, for example, able to turn sounds into beats (Kozak, 2023, p. 41).

As noted above, in our conversations with musicians and music producers, almost all of their descriptions of desirable rhythms and grooves involve metaphors

derived from bodily posture (balance, stability), action (breathing), and movement (head motion, swinging). When describing microrhythm in more detail, their language had a similarly strong imprint of bodily experiences. Table 5 lists how musicians’ descriptions of different microrhythms correlate to specific sets of acoustic properties, perceptual attributes, sound-producing actions. Notably, when jazz/rock/soul guitarists and bass players were asked to play with a “laid-back” feel (note the term itself is a bodily metaphor), the musicians immediately understand what this means. Moreover, when asked what they do to achieve such a laid-back feel, they tend to describe the bodily actions involved (“soft” attacks = less pressure with the pick for plucked sounds), rather than describing the related change in temporal terms (the attack phase of the sound is lengthened); likewise, they note that these sounds have a “floating” feel, indicative of a relatively loose connection to the musical meter, that is, a wider beat bin. Our finding that there is a systematic relationship between so-called ancillary motion; that is, motion that relates to or derives from, for example, emotional intent and the physical interpretation of structural aspects of the music (e.g., Dahl et al., 2010; Davidson, 2007), and the timing instructions given to musicians (pushed is leaning forward), lends further support to the musicians’ understanding of micro-rhythmic feels being highly embodied. In sum, then, both their low-level perceptions and their related motor behaviors seem to be translated into higher level cognitive representations through an embodied framework.

THE CONFOUNDING OF MICRORHYTHM’S “WHAT” AND “WHEN”

An important insight coming from research into embodied cognition is that sound perception implies an understanding of the actions that the listener associates with the sound. Such correspondences have been referred to as *action-sound* couplings (e.g., Godøy, 2010; Jensenius, 2007, 2022). Relatedly, the acoustic

TABLE 5. *Acoustic Properties, Perceptual Properties, Bodily/Gestural Actions, and Participant Descriptions of Laid-back vs. Pushed Microrhythmic Feels (Relative to “On the Beat” Feels) in Jazz/Rock/Soul Electric Guitar, Across Methodological Approaches*

Micro-rhythmic feel	Acoustic properties	Perceptual properties	Sound-producing action	Informant discourse
Laid-back	<ul style="list-style-type: none"> • Longer attack • Longer total duration • Lower spectral centroid 	<ul style="list-style-type: none"> • Late P-center • Wide beat bin • Darker sound 	<ul style="list-style-type: none"> • Slower and longer motion • More upright posture 	<ul style="list-style-type: none"> • Soft attacks • “Floating” feel • Heavy, “fat” sounds
Pushed	<ul style="list-style-type: none"> • Shorter attack • Increased intensity 	<ul style="list-style-type: none"> • Early P-center • Narrow beat bin • Brighter sound 	<ul style="list-style-type: none"> • Faster and shorter motion • Forward leaning Posture • mo 	<ul style="list-style-type: none"> • Sharp attacks • High precision • “Fast” sounds

features that determine the P-center of a musical sound have ecological significance regarding what kind of material and/or action is involved in the production of that sound. Sharp/fast onsets are characteristic of impact sounds; that is, of sounds produced by beating or striking (drums and pianos); slower onsets are characterized by stroking (i.e., bowing), calm breathing or stabilization of vibration, as in a voice, reed, or flute. Loudness is an indication of the energy expended to produce a sound, as well as its proximity. Pitch/spectral centroid is indicative of both the rate of activity of an oscillator/oscillating object as well as its mass, both of which indicate the size of the sound-producing object. Thus, rhythmic microstructure is a strong cue for *what* a sound is, as well as affecting our perception of *when* that sound occurs.

The confounding of the “what” and “when” aspects of microrhythm in our perception and cognition of rhythm is similar to the way small differences in duration or loudness in a series of otherwise similar sounds are regarded as differences in “accent” (Handel, 1989). In fact, already in 1909, Woodrow drew attention to the similar function of relative duration and relative intensity (loudness) in the formation of musical accents (Woodrow, 1909, p. 1), and this has since been confirmed in several more recent studies of perception (see, for example, Povel & Okkerman, 1981; Tekman, 2002; Windsor, 1993).⁶ Thus, it should not surprise us that, as noted immediately above, musicians, for the most part, do not talk about the sonic and temporal aspects of microrhythm separately but recognize that changing the articulation of sound would affect its perceived synchrony relative to other sounds, and vice versa.

The musicians’ tendency to speak in overarching terms like “flow,” “swing,” “feel,” and “groove” also suggests that at some level in our perceptual-cognitive system, we process a rhythmic *gestalt* that integrates both the “what” and “when” aspects of a sound, and this becomes especially apparent when sounds are repeated as part of a pattern. In those contexts, which give rise to a sense of beat and meter, our interviewees related microstructure to a sense not only of flow, but breathing and bodily movement. This has its antecedents in earlier descriptions of beats and meter in terms of “arsis” and “thesis” (upbeat and downbeat), as well as the systole/

diastole of breathing or heartbeat (see London, 2001). In other words, while by their very nature subliminal differences in acoustic features are not directly perceivable, their effects can emerge on higher levels. Our informants’ discourse confirms that the changes in sound and timing we have investigated tend to emerge as differences in an aggregated sense of feel, flow and movement of the larger rhythmic sequence.

MICRORHYTHM AND GROOVE

The confounding of “what” and “when” at the micro level of rhythm has implications for systematic studies of microtiming’s effect on listeners’ experienced groove, measured as ratings of “pleasure” and “urge to move,” which so far have yielded inconsistent results: Some studies show no or detrimental effects of microtiming on groove ratings (e.g., Datsaris et al., 2019; Davies et al., 2013; Madison et al., 2011; Madison & Sioros, 2014) while others show positive effects, at least in certain conditions (Kilchenmann & Senn, 2015; Matsushita & Nomura, 2016; Nelias et al., 2022; Skaansar et al., 2019). The absence of consistent results in systematic microtiming studies is in stark contrast to the discourse on groove among producers and musicians, who invest heavily in how to shape the micro level of their music (see, for example, Berliner, 2009, pp. 349–352; Danielsen et al., 2023; Keil & Feld, 1994; Monson, 1996, pp. 26–29). The findings of the *TIME* project may shed new light on this conundrum.

First, our research shows that, depending on their shape, sounds whose onsets are physically aligned with the grid may in fact involve perceived timing differences due to differences in P-centers relative to otherwise isochronously timed onsets (see, for example, Brøvig et al., 2021, on such effects in EDM). This means that, depending on the sounds used, the no-microtiming (“deadpan”) conditions that have achieved high groove ratings in the studies reporting no positive effects of microtiming, might in fact involve some *perceptual* microtiming.

Second, our results show that “the what” probably has to match “the when” for microtiming to be pleasurable—expressive variations to a series of metronome clicks are not likely to make them very groovy, in other words. Interestingly, when taking a broader scope of micro-rhythmic features into account, either by using stylistically adequate expert performances or stimuli that resemble high-quality musical examples in terms of both temporal and sonic features, microtiming *is* appreciated. Senn et al. (2016), for example, used grooves played by expert performers and found that fully quantized and originally performed microtiming patterns were rated

⁶ Regarding performance as well, “what” and “when” seem to be used in tandem. Accented beats tend to be lengthened in performance (see, for example, Clarke, 1988; Dahl, 2000, 2004; Drake & Palmer, 1993; Gabriellson, 1974, 1999; Waadeland, 2001, 2003, 2006) and when asking a pianist to emphasize one voice in a polyphonic piano performance (i.e., melody lead), it is played both louder and earlier (Goebel, 2001; Palmer, 1996; Repp, 1996).

equally high on groove. This might be explained by both the performed microtiming and the quantized version having an acceptable match between the “what” and “when.” Similarly, Skaansar et al. (2019), when using the R&B/soul-tune *Really Love* by the artist D’Angelo as inspiration for high-complexity groove stimuli, the three highest ranked groove clips of all (15 in total) was this groove with 0 and ± 40 ms asynchrony between kick drum and double bass, which might be explained by the latter groove inducing widened beat bins in the listener. The order of asynchrony of the original (bass 40 ms after kick drum) was ranked higher than the reversed order and the larger asynchronies of ± 80 ms had a clear detrimental effect on groove ratings. Nelias et al. (2022) also found a positive effect of the form of microtiming (downbeat delay) resembling actual practice among swing jazz musicians.

Interestingly, previous research into interaction effects between sonic and temporal features has shown that if the features work in opposite directions, the effect is neutralized and can even be negative (a redundancy loss; see, for example, Melara & Marks, 1990, and Tekman, 2002, on interaction between dynamic accents and perceived duration, also reported by Woodrow, 1909). Such perceptual “mismatch” between the “what” and “when” might, for example, explain the findings of Davies et al. (2013). There they were interested in effects of microtiming on experienced groove in jazz, funk, and samba. All these three musical styles are typically performed by musicians using acoustic and electric instruments, and all three styles have characteristic microtiming patterns. However, the synthetic sounds used in the study rather gave associations to a machinic, “on-the-grid”-oriented aesthetics where asynchronies are absent or rather minute (see Danielsen, 2019). This might have produced a mismatch between the sounds used (the “what”) and the microtiming pattern applied (the “when”) which was detrimental to the groove experience. Given the synthetic sounds used as stimuli, no microtiming would indeed be preferable, and this was also the condition that produced the highest groove ratings.

Our findings from the Scandinavian fiddle music are also interesting in this regard. As reported above, the folk musicians unexpectedly showed *higher* variability than the other expert groups when synchronizing to fiddle sounds from their own genre (Danielsen et al., 2021). With reference to our interview data, this result is consistent with an aesthetic ideal that involves intentional rhythmic-temporal ambiguity, implying that synchronizing to the “what” that fiddle sounds represent does not necessarily involve searching for a precise “when.” Rather, in this musical context, it may be

beneficial to have wider beat bins. This interpretation suggests that the width of the “when,” that is, the beat bin, is indeed a dimension of timing perception that can be informed by what the “what” means to participants with a particular musical enculturation and specialization. A recent EEG study from the *TIME* project provides further support for this, showing that the predicted beat bin of an upcoming sound is partly under top-down control (Leske et al., 2023).

Generally, it seems crucial to avoid the following pitfalls when researching the effect of microrhythm on experienced groove:

- One has to be cognizant of both the “what” and the “when” involved in one’s choice of stimuli (as well as their interdependence).
- P-center location and beat bin width need to be in agreement with the microtimings involved: very small/subtle shifts in timing may not be apparent if the stimulus sounds induce wide beat bins; likewise, larger shifts in timing may be objectionable for sounds with very concise P-centers.
- Different microrhythmic configurations are characteristic of certain genres, and hence listeners may be more or less familiar with (and hence more or less sensitive to) stimuli that present uncharacteristic configurations of microrhythm.

All these aspects must be in place before one can conclude regarding microrhythm’s role in explaining why and how certain groove feels come across as so irresistible while others do not.

Summary and Conclusion

An important premise of the *TIME* project was that musical enculturation and expertise—whether gained through formal training or through immersion in a musical culture—has a profound effect on how and what we hear. To that end, we investigated five different groove-based musical cultures, and the cross-cultural design of the project made it possible to disentangle “nature from nurture” in the ways in which sonic and temporal parameters interact at the micro level of musical rhythm. By comparing different musical genres, we could identify aspects of such interactions that are (most likely) shared by all perceivers (e.g., the effects of the acoustic factors of attack and duration on perceived location and beat-bin width) and at the same time gain important insight into the ways in which such basic perceptual processes are modulated by learning and training, as in the differing perceptions of the fiddle

sound by jazz, folk, and EDM/Hip-Hop musicians (Danielsen et al., 2021).⁷

We also approached the research topic from different methodological angles as the musicological and ethnomusicological experts on the different genres were brought in dialogue not only with each other but also with the researchers who had a quantitative or technological/computational background.⁸ This dialogue helped in designing the experimental parts of the project. The dialogue flowed both ways, as we actively tested the ecological validity of our experimental findings through the interviews and music analyses conducted in the qualitative parts of the project. The exchange between quantitative (experimental) and qualitative investigation thus adopted the form of a hermeneutic circle (Heidegger, 1927/1962), wherein insights from one part shed light on the whole and thereby in turn informed the team's understanding of every other individual part. Through such processes of iterative recontextualization, we integrated the divergent fields, musical cultures and disciplinary perspectives and formed a shared research horizon. Ultimately, a better and deeper sense of both the parts and the whole emerged.

This hermeneutic circle allowed a cross-validation of results across methodological traditions and musical genres, but it also led to the interpretation of quantitative data in unforeseen ways. An example here is the way in which information from the interviews led to additional data analysis. In interviews concerning how they achieved the different timing feels we requested (pushed, on-the-beat, and laidback), the musicians who took part in the performance experiments (Câmara et al., 2020a, 2020b) described a need to adjust their body posture in accordance with the timing feel. The team thus formulated a hypothesis about pushed and laidback timing feels being reflected in “pushed” and “laidback” body postures, and then tested it by looking into the angle of the musicians' upper bodies using the MoCap data from the performance experiments (see Table 2 and discussion above). These data were collected primarily to investigate the sound-producing gestures behind the different timing feels, but our interview results inspired us to investigate the accompanying ancillary gestures as well. Likewise, an innovative approach to attack detection (Lartillot et al., 2021)—that is, a purely signal-processing-related inquiry—came out of a multidisciplinary project not initially centered on signal

processing. The method enables a precise estimation of the attack phase of a sound from an audio recording based on the audio waveform, which is critical to studies of P-center and onset timing. It also shows that perception studies can be an important test of whether signal-processing procedures produce adequate perceptual results.

In sum, our studies present converging evidence of the systematic effects of interaction between temporal and sonic parameters at the micro level of rhythm: musicians are aware of such interactions, can talk about them, and make use of them to create higher-level rhythmic effects/feels; they can be discerned in the acoustic signal and are perceptually salient; they are also understood relative to the bodily gestures involved in producing them. Results from the different investigations thus support the main hypothesis of the project, i.e., that perceived timing is contingent on the microstructure of a sound. We found a strong coupling between attack rise time and duration, and perceived timing: short, percussive sounds with fast attack rise time and short duration have a very narrow beat bin (low variability when synchronizing a click or a tap with the sound) that is located close to the sound's onset, whereas sounds with longer rise time and longer overall duration have a wider beat bin that occurs later relative to the sound's acoustic onset. This pattern accords with the previous findings of research into the perceived timing of sounds in both music and speech (Gordon, 1987; Villing, 2010; Vos & Rasch, 1981; Wright, 2008). But we also found a more complex interaction between microstructure and perceived timing, especially for slow attack sounds with more complex shapes, and also when listeners bring their specific expertise/enculturation to their engagement with those sounds. Thus, in addition to the interaction between “what” a sound is and “when” it is perceived to occur, we would also add “who” is listening to the sound, and “why” are they listening to it—that is, what purpose or goal is involved in their interaction with the sound, whether as a performer or listener. That is, the differences we observed in p-center location and beat bin behavior amongst different groups of expert listeners may be driven by the different rhythmic *affordances* that different groups hear in the “same” sound: for different degrees of tight vs. loose synchronization, succession, and sense of flow.

The interdisciplinary composition of the *TIME* research team is rare in research into rhythm and timing. The assumption was that such a team would maximize the knowledge potential both within and across approaches by fostering a continuous and critical

⁷ The results presented in Danielsen et al. (2021) have been followed up in a second study conducted with classical and jazz singers (see London, Paulsrud, & Danielsen, submitted).

⁸ The team amounted to 16 collaborators overall.

dialogue among these otherwise fragmented research traditions. This, in turn, would increase the potential for novel and valid insights in the project as a whole. With some exceptions (Jakubowski et al., 2022; Polak et al., 2018; see also review in Danielsen et al., 2021), systematic cross-cultural research designs also remain rare (Jacoby et al., 2020). However, the combination of a highly focused research agenda and an interdisciplinary cross-cultural approach is clearly something we will continue to pursue in the future. In our view, it was crucial to producing novel and valid results that hopefully hold true beyond the disciplines and traditions that produced them. These results suggest, first, that future research should take into consideration a wider range of acoustic features involved in the production of groove-based microrhythm, and second, that the “correct” microrhythmic feel varies within and among styles and music cultures. Even though we might feel

like “one nation under a groove” (Funkadelic, 1978), there is always a diversity of people who listen, and a diversity of people who perform.

Author Note

We are grateful to Sverre Albrethsen Reithaug for assistance with preparing the manuscript, and to Rainer Polak for constructive comments to an earlier version of the manuscript. This work was partially supported by the Research Council of Norway through its Centres of Excellence scheme (Project 262762), the TIME project (Grant 249817) and the MIRAGE project (Grant 287152).

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

Publications that originated within the TIME project are marked with an asterisk ().*

ALÉN, O. (1995). Rhythm as duration of sounds in Tumba Francesa. *Ethnomusicology*, 39(1), 55–71. <https://doi.org/10.2307/852200>

BENGTSSON, I., & GABRIELSSON, A. (1980). Methods for analyzing performance of musical rhythm. *Scandinavian Journal of Psychology*, 21(1), 257–268.

BENGTSSON, I., & GABRIELSSON, A. (1983). Analysis and synthesis of musical rhythm. In J. Sundberg (Ed.), *Studies in music performance* (pp. 27–60). Royal Swedish Academy of Music.

BERLINER, P. (2009). *Thinking in jazz: The infinite art of improvisation*. University of Chicago Press.

BJERKE, K. (2010). Timbral relationships and microrhythmic tension: Shaping the groove experience through sound. In A. Danielsen (Ed.), *Musical rhythm in the age of digital reproduction* (pp. 85–104). Ashgate.

BLOM, J.-P. (1981). The dancing fiddle. In J.-P. Blom, S. Nyhus, & R. Sevåg (Eds.), *Slåttar for the Harding fiddle. Norwegian folk music* (Vol. 7, pp. 305–312). Universitetsforlaget.

BROUGHTON, M., & STEVENS, C. (2009). Music, movement and marimba: An investigation of the role of movement and gesture in communicating musical expression to an audience. *Psychology of Music*, 37(2), 137–153.

*BRØVIG, R., SANDVIK, B., AARESKJOLD-DRECKER, J., & DANIELSEN, A. (2021). A grid in flux: Sound and timing in electronic dance music. *Music Theory Spectrum*, 44(1), 1–16. <https://doi.org/10.1093/mts/mtab013>

*BRØVIG-HANSEN, R., SANDVIK, B. E., & AARESKJOLD-DRECKER, J. M. (2020). Dynamic range processing's influence on perceived timing in electronic dance music. *Music Theory Online*, 26(2). <https://doi.org/10.30535/mto.26.2.3>

BUTTERFIELD, M. (2010). Participatory discrepancies and the perception of beats in jazz. *Music Perception*, 27(3), 157–176. <https://doi.org/10.1525/mp.2010.27.3.157>

BUTTERFIELD, M. W. (2011). Why do jazz musicians swing their eighth notes? *Music Theory Spectrum*, 33(1), 3–26. <https://doi.org/10.1525/mts.2011.33.1.3>

*CÂMARA, G. S. (2021) *Timing is everything . . . Or is it? Investigating timing and sound interactions in the performance of groove-based microrhythm* [Doctoral dissertation, University of Oslo, Norway]. Retrieved from <https://www.duo.uio.no/handle/10852/88604>

*CÂMARA, G. S., & DANIELSEN, A. (2018). Groove. In A. Rehding & S. Rings (Eds.), *The Oxford handbook of critical concepts in music theory*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190454746.013.17>

*CÂMARA, G. S., NYMOEN, K., LARTILLOT, O., & DANIELSEN, A. (2020a). Effects of instructed timing on electric guitar and bass sound in groove performance. *Journal of the Acoustic Society of America*, 147(2), 1028–1041. <https://doi.org/10.1121/10.0000724>

*CÂMARA, G. S., NYMOEN, K., LARTILLOT, O., & DANIELSEN, A. (2020b). Timing is everything . . . or is it? Effects of instructed timing style, reference, and pattern on drum kit sound in groove-based performance. *Music Perception*, 38(1), 1–26. <https://doi.org/10.1525/mp.2020.38.1.1>

- *CÂMARA, G. S., SIOROS, G., & DANIELSEN, A. (2022). Mapping timing and intensity strategies in drum-kit performance of a simple back-beat pattern. *Journal of New Music Research*, 51(1), 3–26. <https://doi.org/10.1080/09298215.2022.2150649>
- *CÂMARA, G. S., SIOROS, G., NYMOEN, K., HAUGEN, M. R., & DANIELSEN, A. (2023). Sound-producing actions in guitar performance of groove-based microrhythm. *Empirical Musicology Review*. <https://doi.org/10.31219/osf.io/cdwjr>
- CLARKE, E. F. (1985). Structure and expression in rhythmic performance. In I. Cross, P. Howell, & R. West (Eds.), *Musical structure and cognition* (pp. 209–236). Academic Press.
- CLARKE, E. F. (1988). Generative principles in music performance. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation and composition* (pp. 1–26). Clarendon Press.
- CLARKE, E. F. (1989). The perception of expressive timing in music. *Psychological Research*, 51(1), 2–9.
- CLARKE, E. F. (2005). *Ways of listening: An ecological approach to the perception of musical meaning*. Oxford University Press.
- CLARKE, K. R., SOMERFIELD, P. J., & GORLEY, R. N. (2008). Testing of null hypotheses in exploratory community analyses: Similarity profiles and biota-environment linkage. *Journal of Experimental Marine Biology and Ecology*, 366(1–2), 56–69.
- CLAYTON, M. (2000). *Time in Indian music: Rhythm, metre, and form in North Indian Rag performance*. Oxford University Press.
- COX, A. (2016). *Music and embodied cognition: Listening, moving, feeling, and thinking*. Indiana University Press.
- DAHL, S. (2000). The playing of an accent—Preliminary observations from temporal and kinematic analysis of percussionists. *Journal of New Music Research*, 29, 225–233.
- DAHL, S. (2004). Playing the accent—Comparing striking velocity and timing in an ostinato rhythm performed by four drummers. *Acta Acoustica*, 90(4), 762–776.
- DAHL, S., BEVILACQUA, F., BRESIN, R., CLAYTON, M., LEANTE, L., POGGI, I., & RASAMIMANANA, N. (2010). Gestures in performance. In R. I. Godøy & M. Leman (Eds.), *Musical gestures: Sound, movement, and meaning* (pp. 36–68). Routledge.
- DANIELSEN, A. (2006). *Presence and pleasure: The funk grooves of James Brown and Parliament*. Wesleyan University Press.
- DANIELSEN, A. (2010). Here, there and everywhere: Three accounts of pulse in D’Angelo’s “Left and right.” In A. Danielsen (Ed.), *Musical rhythm in the age of digital reproduction* (pp. 19–35). Ashgate.
- DANIELSEN, A. (2012). The sound of crossover: Micro-rhythm and sonic pleasure in Michael Jackson’s “Don’t stop ’til you get enough.” *Popular Music and Society*, 35(2), 151–168.
- *DANIELSEN, A. (2018). Pulse as dynamic attending: Analysing beat bin metre in neo soul grooves. In C. Scotto, K. M. Smith, & J. Brackett (Eds.), *The Routledge companion to popular music analysis: Expanding approaches*. Routledge. <https://doi.org/10.4324/9781315544700-12>
- DANIELSEN, A. (2019). Glitched and warped: Transformations of rhythm in the age of the Digital Audio Workstation. In M. Grimshaw-Aagaard, M. Walther-Hansen, & M. Knakkegaard (Eds.), *The Oxford handbook of sound and imagination*, (Vol. 2.). Oxford. <https://doi.org/10.1093/oxfordhb/9780190460242.013.27>
- *DANIELSEN, A., JOHANSSON, M., BRØVIG, R., SANDVIK, B., & BØHLER, K. K. (2023). Shaping rhythm: Timing and sound in five groove-based genres. *Popular Music*, 1–22. <https://doi.org/10.1017/S0261143023000041>
- *DANIELSEN, A., NYMOEN, K., ANDERSON, E., CÂMARA, G. S., LANGERØD, M. T., THOMPSON, M. R., & LONDON, J. (2019). Where is the beat in that note? Effects of attack, duration, and frequency on the perceived timing of musical and quasi-musical sounds. *Journal of Experimental Psychology: Human Perception and Performance*, 45(3), 402–418. <https://doi.org/10.1037/xhp0000611>
- *DANIELSEN, A., NYMOEN, K., LANGERØD, M. T., JACOBSEN, E., JOHANSSON, M., & LONDON, J. (2021). Sounds familiar (?): Expertise with specific musical genres modulates timing perception and micro-level synchronization to auditory stimuli. *Attention, Perception, and Psychophysics*. <https://doi.org/10.3758/s13414-021-02393-z>
- DANIELSEN, A., WAADELAND, C. H., SUNDT, H. G., & WITEK, M. (2015). Effects of instructed timing and tempo on snare drum sound in drum kit performance. *Journal of the Acoustical Society of America*, 138(4), 2301–2316. <https://doi.org/10.1121/1.4930950>
- DATSERIS, G., ZIEREIS, A., ALBRECHT, T., HAGMAYER, Y., PRIESEMAN, V., & GEISEL, T. (2019). Microtiming deviations and swing feel in Jazz. *Scientific Reports*, 9, Article 19824.
- DAVIDSON, J. W. (2007). Qualitative insights into the use of expressive body movement in solo piano performance: A case study approach. *Psychology of Music*, 35(3), 381–401. <https://doi.org/10.1177/0305735607072652>
- 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>
- DESAIN, P., & HONING, H. (1989). The quantization of musical time: A connectionist approach. *Computer Music Journal*, 13(3), 56–66. <https://doi.org/10.2307/3680012>
- DOFFMAN, M. (2009). Making it groove! Entrainment, participation and discrepancy in the “conversation” of a jazz trio. *Language and History*, 52(1), 130–147.
- DRAKE, C., & PALMER, C. (1993). Accent structures in music performance. *Music Perception*, 10, 343–378.
- ETANI, T., MIURA, A., KAWASE, S., FUJII, S., KELLER, P. E., VUUST, P., & KUDO, K. (2023). *A review of psychological and neuroscientific research on musical groove* (2006–2022). <https://psyarxiv.com/bmfp6/download?format=pdf>

- FRIBERG, A., & SUNDSTRÖM, A. (2002). Swing ratios and ensemble timing in jazz performance: Evidence for a common rhythmic pattern. *Music Perception*, 19(3), 333–349.
- FUNKADELIC. (1978). *One nation under a groove*. Warner Bros.
- GABRIELSSON, A. (1974). Performance of rhythm patterns. *Scandinavian Journal of Psychology*, 15, 63–72.
- GABRIELSSON, A. (1999). The performance of music. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 501–602). Academic Press.
- GERISCHER, C. (2006). Suingue Baiano: Rhythmic feeling and microrhythmic phenomena in Brazilian percussion. *Ethnomusicology*, 50(1), 99–119.
- GODØY, R. I. (2010). Gestural affordances of musical sound. In R. I. Godøy & M. Leman (Eds.), *Musical gestures: Sound, movement, and meaning* (pp. 103–125). Routledge.
- GOEBL, W. (2001). Melody lead in piano performance: Expressive device or artifact? *Journal of the Acoustical Society of America*, 110(1), 563–572.
- GORDON, J. W. (1987). The perceptual attack time of musical tones. *Journal of the Acoustical Society of America*, 82(1), 88–105. <https://doi.org/10.1121/1.395441>
- GREY, J. M. (1977). Multidimensional perceptual scaling of musical timbres. *Journal of the Acoustical Society of America*, 61, 1270–1277.
- HANDEL, S. (1989). *Listening: An introduction to the perception of auditory events*. MIT Press.
- HANNON, E. E. (2010). Musical enculturation: How young listeners construct musical knowledge through perceptual experience. In S. P. Johnson (Ed.), *Neoconstructivism: The new science of cognitive development* (pp. 132–156). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195331059.003.0007>
- HANNON, E. E., SOLEY, G., & ULLAL-GUPTA, S. (2012). Familiarity overrides complexity in rhythm perception: A cross-cultural comparison of American and Turkish listeners. *Journal of Experimental Psychology: Human Perception and Performance*, 38(3), 543–548. <https://doi.org/10.1037/a0027225>
- HAUGEN, M. R. (2016). Investigating periodic body motions as a tacit reference structure in Norwegian telespringar performance. *Empirical Musicology Review*, 11(3–4), 272–294.
- *HAUGEN, M. R. (2021). Investigating music–dance relationships: A case study of Norwegian telespringar. *Journal of Music Theory*, 65(1), 17–38. <https://doi.org/10.1215/00222909-9124714>
- *HAUGEN, M. R., CÂMARA, G. S., NYMOEN, K., DANIELSEN, A. (2023). Instructed timing and body posture in guitar and bass playing in groove performance. *Musicae Scientiae*. <https://doi.org/10.1177/10298649231182039>
- *HAUGEN, M. R., & DANIELSEN, A. (2020). Effect of tempo on relative note durations in a performed samba groove. *Journal of New Music Research*, 49(4), 349–361.
- HEIDEGGER, M. (1962). *Being and time*. Harper and Row. (Original work published 1927)
- HENRICH, J., HEINE, S. J., & NORENZAYAN, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2–3), 61–83.
- HIRSH, I. J. (1959). Auditory perception of temporal order. *Journal of the Acoustical Society of America*, 31(6), 759–767.
- HODSON, R. (2007). *Interaction, improvisation, and interplay in jazz*. Routledge.
- HOFMANN, A., WESOŁOWSKI, B. C., & GOEBL, W. (2017). The tight-interlocked rhythm section: Production and perception of synchronisation in jazz trio performance. *Journal of New Music Research*, 46(4), 329–341. <https://doi.org/10.1080/09298215.2017.1355394>
- HONING, H. (2013). Structure and interpretation of rhythm in music. In D. Deutsch (Ed.), *Psychology of music* (pp. 369–404). Academic Press.
- HOSKEN, F. (2021). *The pocket: A theory of beats as domains* [Doctoral dissertation, Northwestern University].
- HOVE, M., KELLER, P., & KRUMHANSL, C. (2007). Sensorimotor synchronization with chords containing tone-onset asynchronies. *Attention, Perception and Psychophysics*, 69(5), 699–708.
- IYER, V. (2002). Embodied mind, situated cognition, and expressive microtiming in African-American Music. *Music Perception*, 19(3), 387–414. <https://doi.org/10.1525/mp.2002.19.3.387>
- *JACOBSEN, E., & DANIELSEN, A. (2023). ‘Hard or ‘soft’: Shaping microtiming through sonic features in jazz-related groove performance. *Journal of Jazz Studies*, 14(2), 153–185.
- JACOBY, N., MARGULIS, E. H., CLAYTON, M., HANNON, E., HONING, H., IVERSON, J., KLEIN, T. R., ET AL. (2020). Cross-cultural work in music cognition: Challenges, insights, and recommendations. *Music Perception*, 37(3), 185–195. <https://doi.org/10.1525/mp.2020.37.3.185>
- JAKUBOWSKI, K., POLAK, R., ROCAMORA, M., JURE, L., & JACOBY, N. (2022). Aesthetics of musical timing: Culture and expertise affect preferences for isochrony but not synchrony. *Cognition*, 227, 105205. <https://doi.org/10.1016/j.cognition.2022.105205>
- JANATA, P., TOMIC, S., & HABERMAN, J. (2012). Sensorimotor coupling in music and the psychology of the groove. *Journal of Experimental Psychology: General*, 141(1), 54–75. <https://doi.org/10.1037/a0024208>
- JANKOWSKY, R. C. (2013). Rhythmic elasticity, metric ambiguity, and ritual teleology in Tunisian Stambeli. *Analytical Approaches to World Music*, 3(1). https://www.aawmjournals.com/articles/2014a/jankowsky_AAWM_Vol_3_1.html
- JENSENIUS, A. R. (2007). *Action–sound: Developing methods and tools to study music-related body movement* [Doctoral dissertation, University of Oslo]. UiO DUO Research Archive. <https://www.duo.uio.no/bitstream/handle/10852/27149/1/jensenius-phd.pdf>

- JENSENIUS, A. R. (2022). *Sound actions: Conceptualizing musical instruments*. MIT Press.
- JOHANSSON, M. (2010a). The concept of rhythmic tolerance: Examining flexible grooves in Scandinavian folk fiddling. In A. Danielsen (Ed.), *Musical rhythm in the age of digital reproduction* (pp. 69–84). Ashgate.
- JOHANSSON, M. (2010b). *Rhythm into style: Studying asymmetrical grooves in Norwegian folk music* [Unpublished doctoral dissertation, University of Oslo].
- JOHANSSON, M. (2017). Non-isochronous musical meters: Towards a multidimensional model. *Ethnomusicology*, 61(1), 31–51. <https://doi.org/10.5406/ethnomusicology.61.1.0031>
- *JOHANSSON, M. (2022). Timing–sound interactions: Groove-forming elements in traditional Scandinavian fiddle music. *Puls*, 7.
- KEIL, C., & FELD, S. (1994). *Music grooves: Essays and dialogues*. University of Chicago Press.
- KELLER, P. E. (2014). Ensemble performance: Interpersonal alignment of musical expression. In D. Fabian, R. Timmers, & E. Schubert (Eds.), *Expressiveness in music performance: Empirical approaches across styles and cultures* (pp. 260–282). Oxford University Press. <http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199659647.001.0001/acprof-9780199659647>
- KILCHENMANN, L., & SENN, O. (2015). Microtiming in Swing and Funk affects the body movement behavior of music expert listeners. *Frontiers in Psychology*, 6, 1232. <https://doi.org/10.3389/fpsyg.2015.01232>
- KOZAK, M. (2019). *Enacting musical time*. Oxford University Press.
- KOZAK, M. (2023). Varieties of musical time. In C. Wöllner & J. London (Eds.), *Performing time: Synchrony and temporal flow in music and dance* (pp. 33–45). Oxford University Press.
- KVIFTE, T. (2004). Description of grooves and syntax/process dialectics. *Studia Musicologica Norvegica*, 30, 54–77.
- KVIFTE, T. (2007). Categories and timing: On the perception of meter. *Ethnomusicology*, 51(1), 64–84.
- LAKATOS, S. (2000). A common perceptual space for harmonic and percussive timbres. *Perception and Psychophysics*, 62(7), 1426–1439.
- *LARTILLOT, O., NYMOEN, K., CÂMARA, G. S., & DANIELSEN, A. (2021). Computational localization of attack regions through a direct observation of the audio waveform. *Journal of the Acoustic Society of America*, 149(1), 723–736. <https://doi.org/10.1121/10.0003374>
- LARTILLOT, O., TOIVIAINEN P., & EEROLA, T. (2008). A Matlab toolbox for music information retrieval. In C. Preisach, H. Burkhardt, L. Schmidt-Thieme, & R. Decker (Eds.), *Data analysis, machine learning and applications, studies in classification, data analysis, and knowledge organization* (pp. 261–268). Springer-Verlag.
- LEMAN, M. (2007). *Embodied music cognition and mediation technology*. MIT Press.
- LEMAN, M., & MAES, P.J. (2014). The role of embodiment in the perception of music. *Empirical Musicology Review*, 9(3–4), 236–246.
- *LESKE, S., ENDESTAD, T., VOLEHAUGEN, V., FOLDAL, M. D., BLENKMANN, A. O., SOLBAKK, A. K., & DANIELSEN, A. (2023). Predicting the beat bin—Beta oscillations support top-down prediction of the temporal precision of a rhythmic event. *bioRxiv*, 2023-07.
- LIBERMAN, A. M., & MATTINGLY, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21(1), 1–36. [http://dx.doi.org/10.1016/0010-0277\(85\)90021-6](http://dx.doi.org/10.1016/0010-0277(85)90021-6)
- LONDON, J. (2001). Rhythm. *Grove music online*. Retrieved July 16, 2023 from <https://www.oxfordmusiconline.com/>
- LONDON, J. (2012). *Hearing in time: Psychological aspects of musical meter*. Oxford University Press.
- *LONDON, J., NYMOEN, K., LANGERØD, M. T., THOMPSON, M. R., CODE, D. L. & DANIELSEN, A. (2019). A comparison of methods for investigating the perceptual center of musical sounds. *Attention, Perception and Psychophysics*, 81(6), 2088–2101. <https://doi.org/10.3758/s13414-019-01747-y>
- *LONDON, J., PAULSRUD, T.S., & DANIELSEN, A. (submitted) I just don't hear it that way—Why near transfer between expert musicians is really far away.
- MADISON, G. (2006). Experiencing groove induced by music: Consistency and phenomenology. *Music Perception*, 24(2), 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(5), 1578.
- MADISON, G., & SIOROS, G. (2014). What musicians do to induce the sensation of groove in simple and complex melodies, and how listeners perceive it. *Frontiers in Psychology*, 5, 894.
- MALONE, E. (2022). Two concepts of groove: Musical nuances, rhythm, and genre. *The Journal of Aesthetics and Art Criticism*, 80(3), 345–354. <https://doi.org/10.1093/jaac/kpac020>
- MATSUSHITA, S., & NOMURA, S. (2016). The asymmetrical influence of timing asynchrony of bass guitar and drum sounds on groove. *Music Perception*, 34(2), 123–131. <https://doi.org/10.1525/mp.2016.34.2.123>
- MCADAMS, S., & GIORDANO, B. L. (2008). The perception of musical timbre. In S. Hallam, I. Cross, & M. H. Thaut (Eds.), *Oxford handbook of music psychology*. <https://doi.org/10.1093/oxfordhb/9780199298457.013.0007>
- MELARA, R. D., and MARKS, L. E. (1990). Interaction among auditory dimensions: Timbre, pitch, and loudness. *Attention, Perception and Psychophysics*, 48, 169–178.
- MONSON, I. (1996). *Saying something: Jazz improvisation and interaction*. University of Chicago Press.

- MORTON, J., MARCUS, S., & FRANKISH, C. (1976). Perceptual centers (P-centers). *Psychological Review*, 85(5), 405–408. <https://doi.org/10.1037/0033-295X.83.5.405>
- NELIAS, C., STURM, E. M., ALBRECHT, T., HAGMAYER, Y., & GEISEL, T. (2022). Downbeat delays are a key component of swing in jazz. *Communications Physics*, 5(1), 1–9. <https://doi.org/10.1038/s42005-022-00995-z>
- *NYMOEN, K., DANIELSEN, A., & LONDON, J. (2017). Validating attack phase descriptors obtained by the Timbre Toolbox and MIRtoolbox. In *Proceedings of SMC-17 Sound & Music Computing Conference* (pp. 214–219). Aalto University. <http://urn.nb.no/URN:NBN:no-58820>
- *ODDEKALV, K. A. (2022a). Rytmsiske bumerke [Rhythmic personal marks]. In E. I. Diesen, B. Markussen, & K. A. Oddekav (Eds.), *Flytsoner – studiar i flow og rap-lyrikk*. [Flow zones – studies in flow and rap lyrics.] Scandinavian Academic Press. <https://s3-eu-west-1.amazonaws.com/spartacus.no/production/attachments/2%20rytmsiske%20bumerke.pdf>
- *ODDEKALV, K. A. (2022b). *What makes the shit dope? The techniques and analysis of rap flows* [Unpublished doctoral dissertation, University of Oslo].
- PALMER, C. (1996). On the assignment of structure in music performance. *Music Perception*, 14(1), 23–56.
- POLAK, R. (2010). Rhythmic feel as meter: Non-isochronous beat subdivision in Jembe music from Mali. *Music Theory Online*, 16(4). <https://doi.org/10.30535/mto.16.4.4>
- POLAK, R., JACOBY, N., FISCHINGER, T., GOLDBERG, D., HOLZAPFEL, A., & LONDON, J. (2018). Rhythmic prototypes across cultures: A comparative study of tapping synchronization. *Music Perception*, 36(1), 1–23. <https://doi.org/10.1525/mp.2018.36.1.1>
- POLAK, R., & LONDON, J. (2014). Timing and meter in Mande drumming from Mali. *Music Theory Online*, 20(1). <http://www.mtosmt.org/issues/mto.14.20.1/mto.14.20.1.polak-london.php>
- POVEL, D. J., & OKKERMAN, H. (1981). Accents in equitone sequences. *Attention, Perception, and Psychophysics*, 30, 565–572.
- PRÖGLER, J. A. (1995). Searching for swing: Participatory discrepancies in the jazz rhythm section. *Ethnomusicology*, 39(1), 21–54.
- RASCH, R. A. (1979) Synchronization in performed ensemble music. *Acustica*, 43, 121–131.
- REPP, B. H. (1996). Patterns of note onset asynchronies in expressive piano performance. *Journal of the Acoustical Society of America*, 100(6), 3917–3932.
- REPP, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin Review*, 12, 969–992.
- REPP, B. H., & SU, Y. H. (2013). Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonomic Bulletin and Review*, 20(3), 403–452. <https://doi.org/10.3758/s13423-012-0371-2>
- ROHOLT, T. C. (2014). *Groove: A phenomenology of rhythmic nuance*. Bloomsbury Publishing.
- SETON, J. C. (1989). *A psychophysical investigation of auditory rhythmic beat perception* [Unpublished doctoral dissertation, University of York].
- SENN, O., KILCHENMANN, L., BECHTOLD, T., & HOESL, F. (2018). Groove in drum patterns as a function of both rhythmic properties and listeners' attitudes. *PLOS ONE*, 13(6), e0199604–e0199604. <https://doi.org/10.1371/journal.pone.0199604>
- SENN, O., KILCHENMANN, L., VON GEORGI, R., & BULLERJAHN, C. (2016). The effect of expert performance microtiming on listeners' experience of groove in swing or funk music. *Frontiers in Psychology*, 7, 1487. <https://doi.org/10.3389/fpsyg.2016.01487>
- *SIOROS, G., CÂMARA, G. S., & DANIELSEN, A. (2019). Mapping timing strategies in drum performance. In A. Flexer, G. Peeters, J. Urbano, & A. Volk (Eds.), *Proceedings of the 20th International Society for Music Information Retrieval Conference*. International Society for Music Information Retrieval. https://archives.ismir.net/ismir2019/2019_Proceedings_ISMIR.pdf
- *SKAANSAR, J., LAENG, B., & DANIELSEN, A. (2019). Microtiming and mental effort: Onset asynchronies in musical rhythm modulate pupil size. *Music Perception*, 37, 111–133. <https://doi.org/10.1525/mp.2019.37.2.111>
- STOVER, C. D. (2009). *A theory of flexible rhythmic spaces for diasporic African music* [Unpublished doctoral dissertation, University of Washington].
- TEKMAN, H. G. (2002). Perceptual integration of timing and intensity variations in the perception of musical accents. *Journal of General Psychology*, 129, 181–191. <https://doi.org/10.1080/00221300209603137>
- TOIVIAINEN, P., LUCK, G., & THOMPSON, M. R. (2010). Embodied meter: Hierarchical eigenmodes in music-induced movement. *Music Perception*, 28(1), 59–70. <https://doi.org/10.1525/mp.2010.28.1.59>
- VILLING, R. (2010). *Hearing the moment: Measures and models of the perceptual centre* [Unpublished doctoral dissertation, National University of Ireland Maynooth].
- VOS, J., & RASCH, R. (1981). The perceptual onset of musical tones. *Perception and Psychophysics*, 29(4), 323–335.
- WAADELAND, C. H. (2001). 'It don't mean a thing if it ain't got that swing'—Simulating expressive timing by modulated movements. *Journal of New Music Research*, 30, 23–37. <https://doi.org/10.1076/jnmr.30.1.23.7123>
- WAADELAND, C. H. (2003). Analysis of jazz drummers' movements in performance of swing grooves—A preliminary report. In R. Bresin (Ed.), *Proceedings of SMAC03, Stockholm Music Acoustic Conference* (pp. 573–576). Kungliga Tekniska Högskolan.
- WAADELAND, C. H. (2006). Strategies in empirical studies of swing groove. *Studio Musicologica Norvegica*, 32, 169–191. <https://doi.org/10.18261/ISSN1504-2960-2006-01-11>

- WILSON, M., & KNOBLICH, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, 131(3), 460–473. <https://doi.org/10.1037/0033-2909.131.3.460>
- WINDSOR, W. L. (1993). Dynamic accents and the categorical perception of metre. *Psychology of Music*, 21(2), 127–140.
- WITEK, M. A. (2017). Filling in: Syncopation, pleasure and distributed embodiment in groove. *Music Analysis*, 36(1), 138–160. <https://doi.org/10.1111/musa.12082>
- WOODROW, H. (1909). A quantitative study of rhythm. *Archives of Psychology*, 14, 1–66.
- WRIGHT, M. (2008). *The shape of an instant: Measuring and modelling perceptual attack time with probability density functions* [Unpublished doctoral dissertation, Stanford University].
- ZEINER-HENRIKSEN, H. T. (2010). Moved by the groove: Bass drum sounds and body movements in electronic dance music. In A. Danielsen (Ed.), *Musical rhythm in the age of digital reproduction* (pp. 121–139). Ashgate.