This is the published version of a paper published in *The Brunswik Society Newsletter*.

Citation for the original published paper (version of record):


Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=nbn:se:umu:diva-96786
Testing an Evolutionary Theory of Human Rhythm and Groove: Tapping Musicians’ Implicit Knowledge through Egon Brunswik’s Lens Model

Guy Madison
Umeå University, Department of Psychology, Sweden

Contact: guy.madison@umu.se

Music often induces movement. Hearing the old masters in the concert hall or the latest dance music on the radio may both make us tap our feet or bob our head. This behavior tendency is called Groove, defined as the experience of wanting to move when hearing music. Starting out of mere curiosity about this phenomenon, from a musician’s point of view, I have come to think that it is a human universal that might reflect an adaptive value of being able to synchronize actions with conspecifics. My initial goal, however, was simply to crack the code of how to play so as to induce as much groove as possible. To this humble end one needs to know, first, whether groove is personal or general. If a piece of music that makes me experience groove leaves another person indifferent there is little hope of finding this Holy Grail.

The common impression that some artists do a better job than others at inducing groove is correct, though: About 25 percent of the variability in listeners’ ratings of groove could be explained merely by the differences amongst 64 examples of popular music (Madison, 2006). This music was sampled from commercially available audiograms, and was therefore unlikely to vary systematically in groove induction ability. Given the suboptimal stimulus sampling, individual differences in scaling behavior, and measurement error, 25 percent reflects a high degree of consistency across individuals. The next step was to explore which physical properties are associated with groove.

Starting with a naturalistic sample of real music, we measured a number of higher-order rhythmic properties in the sound signals of 100 commercially available music recordings, and related them to listeners’ ratings of groove for each recording (Madison et al., 2011). We constructed a few dozen so-called descriptors of the audio waveform, essentially computational algorithms that focus on various aspects of the signal, from lower-order spectral properties to higher-order ones. For example, systematic micro-timing involves first analyzing the beat and metrical properties, then applying canonical time according to the meter, and finally measuring deviations between these and the actual signal events. Each descriptor yields a single parameter that expresses the magnitude of the property in question. The descriptors called beat salience, event density, fast metrical levels, and systematic micro-timing were
significantly correlated with the listeners’ ratings. Specifically, beat salience reflects the number and loudness of sounds that occur on the beat, event density is the number and loudness of sound events per unit time, fast metrical levels is the number of metrical subdivisions that sound events are articulated in, and systematic micro-timing is the amount of temporal deviations from the metronomic (i.e., canonical) positions of the metrical grid that are recurrent and hence not random. All these descriptors were positively correlated with groove except micro-timing, which was negatively correlated (Madison et al., 2011). Having thus obtained ideas about the physical properties involved from the real music itself and from the listeners themselves, we would seem to be in a position to experimentally test their validity. But a multifaceted and highly complex phenomenon like music requires an even more careful approach to exhaust the possible physical correlates of groove. I tend to refer to the “black box” problem, meaning that music has so many simultaneous properties that it is practically impossible to control them unless it is simplified reductio ad absurdum. Consider at least two apparent problems.

First, a sample of real music might contain only a few or even no pieces actually intended to induce movement. There are certainly many other desirable aspects of music. In fact, derivative styles of what was originally dance music, such as contemporary jazz, samba, tango, and many other from the Latin and Black Atlantic diaspora, are today performed for listening exclusively.

Second, Groove may be trivially associated with a range of properties that are non-essential for the intention to produce groove. This is a variety of the confounding variables problem. For example, speech comprehension hinges on a range of variables under direct control of the speaker and listener, such as stresses, timing, and anticipation, but also on the quality of the voice, impaired hearing, and background noise, that would trivially affect the perception of any auditory signal. Similarly, a musical tradition that, for example, encompasses a focus on the lyrics, a smaller budget, and a desire for natural sounds is likely to simultaneously feature slow tempo, few instruments, and a narrow frequency spectrum. Now, inasmuch as the focus on lyrics might also render groove unimportant we cannot tell if this intention has any effect in and of itself, since it would be confounded with the other features that also lend the music less movement inducing: A slow tempo is difficult to accommodate to body movement (MacDougall & Moore, 2005), few instruments provide less opportunity for explicating fast metrical levels, syncopation, and rhythmical elaboration (Madison et al., 2011; Madison & Sioros, 2014; Sioros et al., 2014), and less low frequency range power fails to engage the vestibular system (e.g., Todd, 2001). In conclusion, “black box” phenomena require special care to avoid, on the one hand, reducing them to something out of their true nature, and, on the other hand, a number of interpretation and design problems related to confounding variables.

One approach to avoid these problems is to examine what musicians would do if asked to increase or decrease groove across a range of different musical structures. Brunswik’s (1952; 1956) lens model is a potent tool for such situations, in particular because it takes the whole communication process into account so that we can confirm that their behavior has the intended effect on listeners’ perception. In other words, we exploit musicians’ experience and expressive skills directly by asking them to play a number of compositions with as much and as little groove as possible, to amend our exploratory examinations of pre-existing musical pieces with high ecological validity. The compositions were monophonic melodies, and the musicians
had to play all notes in the specified order in a steady, specified tempo. They were however allowed to add notes and to change the note values. These depleted conditions were intended to refine musicians’ strategies by focusing on the devices at hand, mainly timing, dynamics, and rhythmization. Four professional musicians performed 12 monophonic melodies; six simple melodies akin to children’s songs that were composed for this study, and six complex ones were adapted from jazz and rock style recordings. The musicians could hear and rehearse the melodies at home, and at the production session each musician individually first recorded all 12 melodies in a deadpan version similar to the one rehearsed. It was only after this that they were instructed to play the same melodies again with the intention of maximizing and minimizing groove. They played a professional Yamaha keyboard through which information about each keystroke was recorded as MIDI data. The 24 performances were subjected to both listener ratings and performance analysis, totaling 96 performances from the four musicians. Ten different performance parameters were computed from the MIDI data across each performance, namely event density, the magnitude of onset, offset, and duration micro-timing, and the proportions of 8th and 16th note onsets, offsets, and durations. Details about methods, analyses, and results are given in Madison and Sioros (2014).

Thirty non-musicians then rated each of the performances on how movement inducing it was on a scale from 0 (“not at all”) to 10 (“entirely”). Movement inducing was defined as “the sensation of wanting to move some part of your body in relation to some aspect of the music”. The ratings were entered through a slider on the computer screen, and ANOVA tests demonstrated significant effects of intention on Groove ratings.

Table 1 summarizes the lens model results from the dichotomous instruction to induce as much or as little groove as possible, over the 10 performance parameters, and to the 11-point groove rating scale.

Table 1
Lens model factors for simple, complex, and both types of melodies

<table>
<thead>
<tr>
<th>Melody type</th>
<th>r_s</th>
<th>G</th>
<th>R_s</th>
<th>R_r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple</td>
<td>0.620</td>
<td>0.83</td>
<td>0.749</td>
<td>0.687</td>
</tr>
<tr>
<td>Complex</td>
<td>0.394</td>
<td>0.34</td>
<td>0.729</td>
<td>0.523</td>
</tr>
<tr>
<td>Both</td>
<td>0.489</td>
<td>0.78</td>
<td>0.621</td>
<td>0.490</td>
</tr>
</tbody>
</table>

Equation 1 shows that the communication achievement (r_a) is the product of the matching factor G, the multiple correlation between musicians’ (the senders) intention and performance parameters R_s, and the multiple correlation between performance parameters and the listeners’ (the receivers) ratings R_r plus an un-modelled component that consists of the correlation between the residuals in the regression models (C) and the residual variation of the models (Hursch, Hammond, & Hursch, 1964):

\[ r_a = G \cdot R_s \cdot R_r + C \cdot \sqrt{(1 - R_s^2)} \cdot \sqrt{(1 - R_r^2)} \]  

(1)
where G is the correlation between the predicted values for the sender and receiver regression models, and C is the correlation between the residual values for the sender and receiver regression models.

The results demonstrate successful communication for the simple melodies but relatively poor for the complex ones. A high G for the simple melodies indicates that performers and listeners share a common code. The two R factors indicate that the cues are not used fully consistently, in particular not by the listeners. Matching primarily limits the communication for the complex melodies, because the performers’ cue utilization is almost equal to that for the simple melodies. The un-modelled component C varied from 0.37-0.44, which is rather low and indicates that we have not excluded cues that performers and listeners would have used consistently. Although not shown here, significant positive correlations between intention and performance cues were found for event density and proportions of 8th and 16th note onsets and 8th offsets (.38 - .57), whereas cue correlations for micro-timing were small and negative.

Overall, these results support those based on real music (Madison et al., 2011) by exhibiting a quite similar pattern of correlations between groove and physical properties. However, musicians’ addition of 8th and 16th note onsets when none were specified in the score show that they intentionally syncopate to increase groove, that is, tend to play stressed notes on relatively weaker positions in the beat or metrical structure (Randel, 1986). Likewise, syncopes and other notes on weak metrical positions were sometimes moved to the beat when musicians decreased groove in the complex melodies. This is a very important qualification of the real music results, where the correlations with density and fast metrical levels could not discriminate between syncopation and the presence of other sounds on weak positions, such as perfectly metrical rhythmic patterns like those typically played by the shaker, hi-hat, tambourine, or rhythm guitar. The poor communication for the complex melodies was attributed to a ceiling effect resulting from their already busy structure, which leaved little room for adding syncopation.

Having thus confirmed, with two quite different approaches, that groove is positively associated with event density and syncopation, and negatively associated with micro-timing (although far from significantly so in the musician study), we proceeded to test these associations experimentally in subsequent studies (Davies, Madison, Silva, and Gouyon, 2013; Sioros, Miron, Davies, Gouyon, & Madison, 2014).

The adaptive perspective on the perception and production of rhythmical patterns presupposes that the abilities and behaviors involved are functional, and we have proposed that their function is to facilitate temporal prediction and synchronization (Madison & Merker, 2005; McNeil, 1995; Merker, Madison, & Eckerdal, 2009). One specific evolutionary scenario posits that producing loud signals by joint vocal exclamations make these signals reach farther and attract larger numbers of conspecifics (Merker, 1999). Several observations support a phylogenetic history. Music is a human universal (Pinker, 2002), and coordinated dance to rhythmically predictable music seems to occur in all cultures (Nettl, 2000). Experiencing rhythmic music is associated with pleasure (Madison, 2006; Todd, 2001; Witek et al., 2014) through activation of brain areas associated with reward and arousal (e.g., Blood & Zatorre, 2001). Finally, passive listening to music and rhythmic sequence activates motor system areas even for tasks without any reference to movement (e.g., Chen, Penhune, & Zatorre, 2008; Grahn & Brett, 2007). Groove
would according to such a scenario constitute the motivational tendency for synchronization, and should conceivably be related to the signal’s effectiveness for synchronization. Consistent with this, faster metrical levels were associated with greater synchronization accuracy (Madison, 2014).

In conclusion, the application of a Brunswikian lens model presented herein provided critical knowledge for extending previous findings into experimentally testable hypotheses. A series of cumulative designs have so far failed to falsify the functional theory of rhythm, which posits that synchronization is associated with an adaptive value (Merker et al., 2009; Madison et al., 2011). The findings are generally consistent with the idea that Groove reflects the behavioral tendency to engage in synchronization, as well as the utility of an auditory signal to facilitate precise synchronization.

Acknowledgements
This research was supported by grant P2008:0887 to Guy Madison from the Bank of Sweden Tercentenary Foundation.

References:


---

**News from the Past Year**

---

**Jeryl L. Mumpower**

*U. S. National Science Foundation, on leave from Texas A&M University, USA*

*Contact: jmumpowe@nsf.com (jmumpower@tamu.edu)*

For the past two years, I have been on leave from my position at Texas A&M University and have served as Division Director for Social and Economic Sciences at the U.S. National Science Foundation. I am now beginning my third and final year in this role. This continues to be an interesting experience and I believe I am engaged in important, valuable work. It has certainly affected my ability to get research done, but I continue to try to be as active as my responsibilities permit.


My colleagues at Texas A&M University, Arnie Vedlitz, Xinsheng Liu, and I just completed a paper, now in review, entitled "Psychometric and Demographic Predictors of the Perceived Risk of Climate Change and Preferred Resource Levels for Climate Change Management Programs." Finally, Tom Stewart, Jim Holzworth and I continue work on our next paper stemming from our research program investigating how people make selection and detection decisions (e.g., how they decide whether to hire someone or whether a patient has a disease) in the face of uncertainty and different feedback conditions.

References: