# Similarity and Families of Musical Rhythms

Erica Cao
*Princeton University*

Max Lotstein
*Universität Freiburg, Freiburg im Breisgau, Germany*

Philip N. Johnson-Laird
*Princeton University*

**What determines the similarity of musical** rhythms? According to the "family" theory, which this paper presents, one factor is the temporal sequence of the onsets of notes: rhythms with the same pattern of interonset intervals tend to sound similar. Another factor is meter. It determines whether or not rhythms are members of the same family, where families depend only on three types of possibility for each metrical unit. If the beat is the relevant metrical unit, these three possibilities are: 1) a note starts on a beat and therefore reinforces the meter, 2) a syncopation anticipates the beat and lasts through its onset and therefore disturbs the meter, and 3) all other events such as rests or ties that start on the beat provided no syncopation anticipates them. Two experiments showed that similarity between rhythms depends on both their temporal patterns of onsets and their families, which combined give a better account than edit distance – a metric of the distance apart of two strings of symbols. Two further experiments examined the errors that participants made in reproducing rhythms by tapping them. Errors more often yielded rhythms in the same family as the originals than rhythms in a different family.

*Received: March 1, 2013, accepted October 10, 2013.*

**Key words:** meter, perceptual groups, rhythm, similarity, syncopation

Music can forego rhythm, as in certain works of Ligeti. It can contain a sequence of random onsets, as in certain works of Cage. It can contain systematic but highly unpredictable onsets, as in certain works of Babbitt. It can be based solely on a pulse, as in certain aboriginal melodies (Cook, 2012, p. 308). In common practice, however, music is based on meter. And meter has the advantage that it allows musicians to use syncopations and to vary the predictability of rhythms in a way that is not otherwise possible. Granted such variations, a question that concerns the cognitive foundations of music is what makes musical rhythms similar to one another. The aim of the present article is to answer this question.

The article begins with pertinent results of previous studies of meter, rhythm, and syncopation. It then presents what we refer to as the "family" theory of rhythms. A precursor was described in Johnson-Laird (1991), but the new theory, which we have implemented in a computer program, yields some novel but testable predictions about rhythms. The article then reports four experiments corroborating these predictions. Finally, it draws some general conclusions about rhythm.

## Previous Studies of Meter, Rhythm, and Syncopation

### Meter

Meter in common practice depends on organizing a sequence of isochronous events into strong and weak beats, and this principle extends to smaller units, so as to create a hierarchy of such organizations (see, e.g., Handel, 1992; Johnson-Laird, 1991; Jones & Boltz, 1989; Yeston, 1976). Music that depends on a hierarchy of pulses has a meter. The largest metrical unit is the measure, the basic unit is the beat (the level known as the *tactus*), smaller units are subdivisions of the beat, and so on for a small finite number of levels. This sort of structure can be captured in a context-free grammar (see, e.g., Lerdahl & Jackendoff, 1983; Longuet-Higgins, 1979; Longuet-Higgins & Lee, 1984; Steedman, 1977). Strictly speaking, however, context-free grammars call for more computational power than is needed for metrical structure because they imply no limits on the number of levels in the hierarchy. Meter is restricted to loops of operations that can generate a hierarchy that calls for no more than a handful of levels, and that cycle repeatedly through a small number of beats (London, 2004, 2008). These computational requirements therefore do not fit neatly into the well-known Chomsky hierarchy of grammars and the corresponding automata of increasing computational power (see Hopcroft & Ullman, 1979, Chapter 2). Instead, meter calls for a device that can carry out a finite

*Music Perception*, volume 31, issue 5, pp. 444–469, issn 0730-7829, electronic issn 1533-8312.  DOI: 10.1525/mp.2014.31.5.444

number of operations – a finite-state automaton – but that nevertheless creates a structure akin to one produced by a context-free grammar (cf. Large & Kolen, 1994).

The intuitive cognitive system has no access to unlimited working memory and so it too is at best a finite-state automaton (Evans, 2008; Johnson-Laird, 1983; Kahneman, 2011; Khemlani & Johnson-Laird, 2012; Stanovich, 1999). This characteristic limits the number of beats likely to yield a perceptible meter. The system copes easily with meters of two, three, or four beats to the measure, which can be created and perceived without the need for counting (Miller, 1956). It tends to group beats into sets of two or three for meters with a greater numbers of beats. It also tends to impose a metrical organization on perfectly isochronous pulses (Bolton, 1894; Fraisse, 1982). Even seven-month-old infants are sensitive to the difference between duple and triple meters in simple examples (Hannon & Johnson, 2005). They are also sensitive to violations of many types of meter, whereas adults have become sensitive only to violations in those meters common to their culture (Hannon & Trehub, 2005) – a phenomenon analogous to the decline in infants' phonetic mastery, in which they become sensitive only to the phonemes and transitions from one phoneme to another that occur in their native language (Aslin, Jusczyk, & Pisoni, 1998). Likewise, individuals reproduce a metrical rhythm more accurately than a nonmetrical one (Summers, Hawkins, & Mayers, 1986).

Some music such as jazz and rock makes beats explicit, i.e., a perceptible musical event occurs on each beat, such as a note in the bass or in the percussion. Much metrical music, however, lacks manifest beats. Its meter is made perceptible in part by cues, such as an accent on any note that falls on the first beat of the measure, which the intuitive system uses to infer the meter (Hannon, Snyder, Eerola, & Krumhansl, 2004; Large, 2008; Palmer & Krumhansl, 1990). The perception of meter is accordingly based on a cognitive framework that listeners can impose on music given that they have inferred beats. Meter is akin to subjective contours in visual figures (Shepard, 1999): it is cued by features in the stimulus but goes beyond them to fill in missing beats (Lee, 1991). The task of programming computers to make these inferences is not trivial (Desain & Honing, 1999, 2003; Handel & Oshinsky, 1981; Honing, 2012; Large & Palmer, 2002). The computational problem is to detect cues to an isochronous sequence, possibly embedded among other sounds, to set up a sequential prediction that can be maintained even in the sporadic absence of cues, and to modify the prediction to accommodate changes in tempo (Parncutt, 1994). Some simple principles are that a strong beat in the meter should not coincide with a rest or tied note, but should occur on isolated notes and on the first and last notes in a contiguous sequence (Povel & Essens, 1985). Meter may even depend on musical genre and style (Collins, 2006).

Evidence shows that two- to three-day-old infants can induce beats (Zentner & Eerola, 2010), and accordingly that the system has strong innate determinants (Patel, Iversen, Bregman, & Schulz, 2009). Because disturbances to the beat elicit an evoked potential in the brain independently from an individual's attention to the rhythm (Honing, Ladinig, Háden, & Winkler, 2009), the intuitive cognitive system can induce beats without conscious deliberation. It has access to internal mechanisms that generate beats. Such a mechanism may be a direct pulse-generator based on a neuronal oscillator, or it may depend on assessing the decay of information over time (Grondin, 2010; Large & Snyder, 2009).

The metrical organization of audible events is an essential musical principle to which the intuitive cognitive system is attuned, because even nonmusicians have a tacit ability for such an organization. A simple demonstration of this fact is to tap the following sequence,

It is difficult not to tap the onsets in a metrical way, and it is difficult for listeners not to perceive the tapping as metrical, too. In a series of unpublished experiments, Jung Min Lee, Malcolm Bauer, and the third author, showed that musically naive individuals have an intuitive sense of meter. In one experiment, the experimenter counted out aloud, "One two three four," to set up an expectation of four beats to the measure, and then the participants heard one of the rhythms in Figure 1 as a sequence of claps. The participants' task was to judge whether the final note was longer than, shorter than, or equal in duration to the note that preceded it. Of course, all the claps were of roughly the same duration, and so the only basis for judgments was that common time led the participants to expect a note at the start of the next measure after the clapping. The interval between the final clap and this expected note was the same as the previous interonset interval for the first rhythm in Figure 1, shorter than it for the second rhythm, and longer than it for the third rhythm. The participants responded reliably in the way that corroborated this metrical interpretation, and the difference between musicians (83% of metrical responses) and nonmusicians (79% of metrical responses) was not reliable. Meter organizes musical expectations and musical performance (Shaffer, Clarke,

**FIGURE 1.** Three rhythms that were clapped and that yielded different judgments about the duration of the final note based on the expectation of a second measure in common time.

& Todd, 1985), and it is based on an intuitive cognitive system that musicians and nonmusicians both possess.

A basic element of musicianship is the ability to determine the number of beats in the measure and the start of the measure for a given piece of music. But, even for musicians this task can be difficult. One reason is that composers from Bach onwards play with metrical ambiguity (Krebs, 1999; Samarotto, 1999; Steedman, 1977). The opening of Bach's *Fugue in C major* from the "48" is a good example. Listeners readily tap the beat, but they cannot identify the start of a measure until they have heard several measures. Likewise, composers can phrase a rhythm so that it conflicts with meter, as in the opening of Ravel's *Rapsodie Espagnole: I. Prélude à la nuit,* in which the beams on notes indicate a meter of 2/4 and cross the bar line of measures in 3/4. Ravel took this procedure still further in the second movement of his *Piano Trio* in which a staccato scherzo written in 3/4 meshes with the legato phrases written in 4/2 with a coincidence between the two meters after every three measures in the latter meter.

*Rhythm.* William James (1890) acknowledged that if someone claps the opening phrase of a familiar tune, such as *The Star-Spangled Banner*, listeners can identify it. The intervals between the onsets of successive notes (interonset intervals) are the main component of a rhythm. Hence, two rhythms sound similar even if one is a sequence of staccato notes each followed by a rest, and the other is a series of legato notes that each last until the onset of the next note (see, e.g., Clarke, 1984; Povel, 1984). What matters is the sequence of the interonset intervals, or, as we refer to them, the "pattern of onsets." This view dovetails with a Gestalt approach in which notes are the figure against a background of rests (Cooper & Meyer, 1960; London, 2004).

In contrast, rhythms that have no strong meter tend to be perceived in terms of sets of contiguous onsets separated by silences. Povel and Essens (1985) refer to a sequence of contiguous onsets as a "perceptual group." Consider a rhythm such as:

```
11100110001110
```

which we have written in binary notation, where "1"denotes an onset on a temporal unit and "0" denotes a rest or a tie on an equivalent temporal unit, so each 1 and 0 marks the start of an equal interval of time. This rhythm contains three perceptual groups, which Povel and Essens represent as: 323, where the first "3" represents the three contiguous notes prior to the first rest or tie, "2" represents the next perceptual group, and "3" represents the final perceptual group (cf. Bamberger, 1982, 2003). When individuals judged whether two nonmetrical rhythms were the same or different, these perceptual groups had more of an effect on their judgments than the time intervals between them (Handel, 1992). When two rhythms had different perceptual groups, they were easier to discriminate than when they did not (Ross & Houtsma, 1994). In the first case, one rhythm had an instance of 10 where the other rhythm had an instance of 01; whereas in the second case, one rhythm had an additional 0 between two perceptual groups. Handel (1998) argued that meter affects discrimination between rhythms only within the more fundamental organization of perceptual groups, but other researchers have found that meter can improve the discrimination of rhythms (Hébert & Cuddy, 2002). Perceptual groups are not based on meter, and so they make no distinction between notes on the beat and syncopations – a factor that is crucial to the present investigation.

One factor sometimes ignored in studies of perceptual groups is the "edit" distance between rhythms. This metric is due to Levenshtein (1966, see also Damereau, 1964), and it depends on the number of edits required to get from one string of symbols to another, where an edit consists in deleting a symbol, inserting a symbol, or substituting one symbol for another. A program in Common Lisp computes this distance and is available at: http://psych.princeton.edu/psychology/research/johnson_laird/music.php

Edit distance is a natural metric for comparing rhythms with different perceptual groups. Consider, for instance, this pair of rhythms (from Ross & Houtsma's, 1994, study):

```
11100110001110
```

and:

```
11010110001110
```

The second rhythm is derived from the first by swapping the first adjacent 10 into 01, and so the perceptual groups of the first rhythm (323) differ from those in the second rhythm (2123). The edit distance between the two rhythms equals 2, because 0 is substituted for 1 in the third character, and 1 is substituted for 0 in the fourth character. Likewise, the edit distance from 323 to 2123 also equals 2. If each edit is of equal psychological importance, the smaller the edit distance between two stimuli, the more similar they should be, and therefore the harder they should be to discriminate (see Orpen & Huron, 1992, for the relevance of the metric to various aspects of music including rhythms). Edit distance also has a wide application in cognitive psychology (see, e.g., Ragni, Khemlani, & Johnson-Laird, 2014, for its use with deductive reasoning), and Toussaint and his colleagues have examined in depth its applicability to rhythms (e.g., Toussaint, 2013). It yields sensible results when it is applied to human judgments about the similarities between various types of cyclical rhythms, including the underlying patterns in various Afro-Cuban and Middle-Eastern rhythms (Toussaint, Campbell, & Brown, 2011). These authors allow, however, that it may have shortcomings because it weights equally each of the three types of edit. Indeed, neither perceptual groups nor edit distance take into account meter (Post & Toussaint, 2011, p. 177), and so the measure does not distinguish between notes and syncopations. Consider, for instance, these two rhythms:

1010101010

and:

01010101010

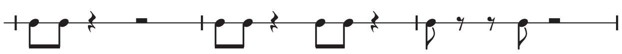
Their edit distance apart is only 1, because a single 0 has been inserted at the start of the second rhythm, and they have identical perceptual groups. Yet, if "1" and "0" each last for an eighth note and the two rhythms both start on the first beat of a measure in common time, then the two rhythms are quite different. The first rhythm is a sequence of notes on the beat, whereas the second rhythm is a sequence of syncopations. Our experiments enabled us to make a crucial comparison between our theory and the perceptual groups.

Once listeners have induced a meter, it has a striking effect on the perception of rhythms. A temporal pattern of onsets is perceived, not as demarcating intervals in the continuum, but as an instance of a category, such as an eighth note, a quarter note, and an eighth note, with allowances for small deviations from their mechanically exact performance (see, e.g., Clarke, 1999). Hence, when experiments explore the space of possible temporal intervals between the onsets of notes, their results show that individuals map them into a much smaller space of categories based on simple note durations (e.g., Desain & Honing, 2003). This phenomenon of categorical perception occurs in domains other than music, such as the perception of speech sounds (Harnad, 1987). Many musical rhythms are composed from a simple repertoire of patterns of onsets – of which the most frequent in binary notion are: 111, 1101, 1011, 1110 (Huron, 2006). Their frequency of occurrence in turn affects listeners' tolerance in categorizing instances of them despite vagaries in their timings.

Phrases in a melody have a rhythm. For singers, the end of a phrase is a place to take a breath. Listeners need to get their breath back too, i.e., to have time to construe the events in the phrase. Phrases are therefore analogous to sentences in discourse. Phrases are demarcated by various cues, e.g., the interval from the onset of the last note in a phrase to the onset of the first note in the next phrase tends to be longer than any interval between adjacent notes within the phrase (see, e.g., Stoffer, 1985). Some phrases have subphrases within them. Lerdahl and Jackendoff (1983) treat this hierarchy as dependent on context-free rules that ensure that everything from the piece as a whole down to the lowest subphrase is treated as a "grouping." The preferred groupings, however, depend on rules such as the principle above demarcating phrases, and Deliège (1987) provided corroboratory evidence in a study in which participants made conscious deliberations to demarcate groupings. Phrases can differ from one genre to another. For example, according to Lerdahl and Jackendoff, phrases in classical music usually divide into two equal halves, but the improvisations of Charlie Parker show no such preference.

*Syncopation.* In informal terms, a *syncopated* note has an onset on a metrical unit of lesser importance than one that occurs prior to the onset of the next note, and so it tends to disturb the meter for a moment. The importance of a metrical unit is given by its height in the tree generated by a metrical grammar (see, e.g., Longuet-Higgins & Lee, 1984). In the context of common time, the second rhythm in Figure 1 ends with a syncopated note given that no note occurs on the first beat of the next measure. The syncopated note occurs on an unimportant metrical unit – the second half of the fourth beat – and, because no note occurs on the first beat of the next measure, it lasts through a more important metrical unit. In theory, the nearer an onset is to a more important metrical unit, the greater the degree of syncopation provided that its anticipation of this unit is perceptible, or else the note will merely sound as if its onset is a slight perturbation in the performance of an unsyncopated note.

In a sequence of isochronous pulses with no metrical structure, systematic syncopations are impossible. If notes occur after one pulse and before another pulse, they cannot do so in a systematic way unless the interval from one pulse to the next is divided in a regular way, or unless the pulses themselves are grouped together into strong and weak ones. Either of these procedures yields a meter. Syncopation accordingly depends on the location of a rhythm with respect to the meter. As an example, consider the opening phrase of Duke Ellington's *C Jam Blues*. It has the following the rhythm:



Every second note is a syncopation: its onset anticipates the following beat, and occurs on a metrically less important unit than the tactus. However, the location of the start of the measure is critical. The same temporal pattern of onsets could occur in this metrical framework:

This shift of the rhythm in relation to the meter eliminates all but one of the syncopations. The first note and the note at the end of the first measure are not syncopations, but anacruses or "pick up" notes preceding the note on the first beat of the next measure. But, the note at the end of the second measure is a syncopation because it is not followed by a note on the first beat of the final measure. If you set up the beat and meter and then tap the two rhythms, they sound different, and the difference is attributable solely to their relations to the metrical framework.

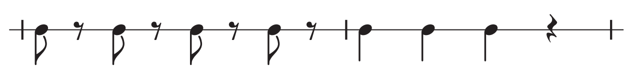
Notes with onsets on units of metrical importance reinforce the meter, whereas syncopations disturb the meter (Volk, 2008). This effect is detectable in judgment (Lee, 1991) and in EEG patterns (Ladinig, Honing, Háden, & Winkler, 2009). Likewise, syncopations yield more errors in tapping on the beat, and in reproducing a rhythm (Fitch & Rosenfeld, 2007). They induce a weaker internal "clock" (Povel & Essens, 1985). The perceptions of rhythm and of meter are therefore mutually interdependent: the intuitive system infers meter from a rhythm, but the nature of the rhythm itself depends on the meter (e.g., Essens, 1995; Essens & Povel, 1985; Palmer & Krumhansl, 1990; Shmulevich & Povel, 2000). A study of evoked potentials in the brain suggests that the two types of perception have different time courses (Geiser, Ziegler, Jancke, & Meyer, 2009).

## The Family Theory of Rhythms

Any genre of music has a large number of possible rhythms that vary in the degree to which they resemble one another. Many listeners can hear these resemblances, and many musicians can improvise variants of the same rhythm. These similarities reflect the cognition of rhythms, and so it makes sense to ask what determines the similarity of rhythms. Many factors are likely to have an effect, including similarity in meter, tempo, subdivisions of the beat, number of notes, density of notes, and manner of attack. The claim is hardly controversial. But, suppose all these factors are held constant and suppose also that listeners have induced the beat and meter – what determines the similarity of rhythms? The family theory posits two critical factors.

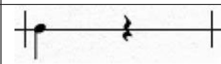
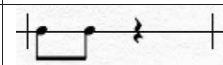
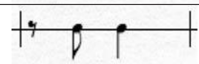
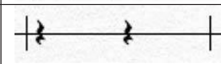
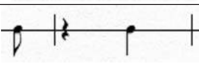
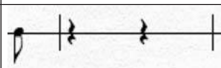
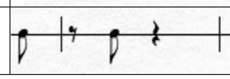
The first factor is whether or not the two rhythms contain an identical temporal sequence of interonset intervals, i.e., the same sequence of temporal intervals from the onset of one note to the next (London, 2004). If they have the same "pattern of onsets," which we use to refer to this factor in relation to the same metrical structure, then they are highly similar even if one rhythm is legato and the other is staccato with rests between each adjacent pair of notes. Hence, the following two rhythms are highly similar:

and:



The second factor arises from the effect of shifting a rhythm in relation to the meter, as in the two versions above of *C-Jam blues*. The two versions have the same pattern of onsets, and yet they sound different. This second factor accordingly concerns a more abstract level of similarity between rhythms – their relations to metrical structure. We describe this factor first at the level of the tactus (the individual beats), and then we describe how it applies to metrical units smaller than the tactus. The family theory stipulates that there are only three types of musical events that matter with respect to meter. The most important events are syncopated notes – they carry the most information, because they are relatively unpredictable, anticipate the beat, and disturb

**TABLE 1.** *Examples of the Nine Different Families of a Single Measure in a Meter of 2/4.*

| | | Second beat of measure | | |
|---|---|---|---|---|
| | | N | O | S |
| First beat of measure | N |  |  |  |
| | O |  |  |  |
| | S |  |  |  |

*Note*: S denotes a syncopation, N denotes a note on the beat, and O denotes all other musical events. S is located on the beat whose onset the syncopation precedes and lasts through.

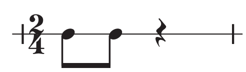
the meter for a moment. The next most important events are notes with onsets on a beat, because they maintain the meter. The least important events are all other possible events – ties and rests on the onset of the beat, as long as a syncopation does not precede them. They neither disturb the meter nor maintain it. The three types of event are mutually exclusive and exhaustive. We label them using these abbreviations:

S for syncopations,
N for notes on the beat,
O for other events.

Each beat in any metrical rhythm can be categorized as an instance of one, and only one, of the three categories, S, N, or O, and so it is the beat which a syncopation anticipates that is classified as S even if this beat coincides with a rest or a tie. This point bears emphasis, because it has subtle implications.

Table 1 illustrates how to classify each of two beats in the meter 2/4 into one of the three categories. Given that listeners have induced the beat and meter, the table shows examples of all nine possible pairs of familial events on two beats (N N, N O, N S, and so on). And it illustrates the role of context for the assignment of S. We have implemented the algorithm for classifying rhythms according to their family in a computer program written in Common Lisp, and the source code is available at: http://psych.princeton.edu/psychology/research/johnson_laird/music.php

The computer program assigns an occurrence of N whenever the onset of a note coincides with the start of a beat, and the categorization is not affected by notes, if any, that occur after this onset and prior to the next beat. The program classifies a beat as a syncopation, S, provided that two conditions hold: first, no note starts on the beat, and, second, a note occurs after the start of the prior beat. It is this note, or the last of a sequence of such notes, that is the syncopated *note*. But, it is the *beat* that the note anticipates that is categorized as S. For example, consider this sequence of notes in 2/4:



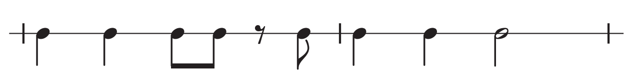
The first eighth note is an instance of N on the beat. The second eighth note starts after the onset of the first beat, and no note has an onset on the second beat, which is a rest. Hence, the second eighth note is a syncopation anticipating the second beat, and so the rhythm is in the family:

| N S |

The beat corresponding to a rest is accordingly assigned the classification of S because it is the beat with an onset that the syncopation anticipates and whose onset it lasts through. This example is in the top right-hand corner of Table 1. The lowest row in Table 1 illustrates three cases in which the first beat is S, because a note starts prior to this first beat and lasts through its onset. Any beat that is neither S nor N is, by definition, O. As Table 1 illustrates, typical beats that are classified as O are cases in which no note starts on the beat because of a rest, and no syncopation anticipates the beat. The computer program takes as input a rhythm in a numerical notation, e.g.,:

```
| 1/4 1/4 1/8 1/8 rest-1/8 1/8 | 1/4 1/4 1/2 |
```

that corresponds to:



Its output is the family to which the rhythm belongs:

| N N N S | N N N O |

Many other rhythms belong to the same family, see, e.g., the stave labeled "Target 7 middle" in Figure 3 below for another example. Once again, this example illustrates a case in which S is assigned to a beat that is a rest, because a syncopation anticipates the beat and lasts through its onset.

The program allows analysis at levels smaller than the beat. If this example from the preceding paragraph:

`| 1/4 1/4 1/8 1/8 rest-1/8 1/8 | 1/4 1/4 1/2 |`

is analyzed at durations of an eighth note, the result is:

`| N O N O N O S O | N O N O N O O O |`

In this case, the analysis is not in the least revealing – compared to an analysis at the level of quarter-note beats, it merely interpolates additional instances of O. But, consider this example in which the meter is 4/4:

`| 1/16 1/8 1/16 1/16 1/8 1/16 1/16 1/8 1/16 1/4 |`

At the level of quarter-note beats its family is: | N N N N |. But, an analysis at the level of eighth notes is more revealing because it shows that syncopations occur within the beats in 4/4:

`| N S N S N S N O |`

Figures 3, 4, 6, and 7 below illustrate many examples of rhythms in the same families, and of rhythms in different families.

*The predictions of the family theory.* The family theory makes five principal predictions. First, if two rhythms share the same pattern of onsets, then they should tend to be judged as similar.

Second, if two rhythms are from the same family then, other things being equal, they should be judged as more similar than two rhythms from different families.

Third, if individuals try to reproduce a rhythm, say, by tapping it, then their errors should tend to yield rhythms in the same family as the original target as opposed to rhythms in a different family.

Fourth, errors in reproduction should be more likely to occur in the case of a syncopation, S, than in the case of a note on a beat, N. This trend reflects the departures from the metrical norm: N reinforces the beat because a note occurs on the beat, whereas S disturbs the beat, because a note anticipates a beat and lasts through its onset.

Fifth, the fewer the notes in a rhythm, the easier it should be to reproduce, because it places a lesser load on the processing capacity of working memory. But, the family theory predicts an interaction. If an additional N eliminates an S, then this extra note should be easier to reproduce than other types of addition (see the introduction to Experiment 4).

## Experiment 1

Previous research has shown that individuals are able to assess the similarity of rhythms (e.g., Pitt & Monahan, 1987). Likewise, in an earlier study (Gabrielsson, 1973), participants were able to rate the similarity of tape-recorded rhythms and melodies on a scale from 0 to 100, and multidimensional scaling revealed a variety of dimensions underlying their judgments, including meter, tempo and the density of onsets, the relative uniformity of rhythms, accent on the first beat, and several other such variables. Most of these variables are global in that they apply across sets of rhythms. Our concern, however, is with the similarity between particular rhythms, and so Experiment 1 examined particular pairs of rhythms. A pilot study showed that an identical pattern of onsets of three or more notes swamped any effect of family alone. Likewise, a design based on random samples of rhythms from the same family is not feasible because it is necessary to control for the same numbers of notes in the rhythms, e.g., the family: | N N N N |, can consist of 4 quarter notes or 8 eighth notes. The experiment accordingly minimized these differences. The participants' task was to listen to one pair of rhythms and then to another pair, and to judge which pair was of a greater similarity. All the rhythms were two measures in length, and Figure 2 shows an example of the four versions of pairs of rhythms used to construct the materials. A crucial region of each rhythm was manipulated in the experiment, and it is marked with a slur in the figure. The first two measures in the figure are the target rhythm, and the remaining pairs of measures are comparisons. The second two measures in the figure have a crucial region identical to that of the target,

FIGURE 2. An example of four rhythms from which pairs of rhythms were constructed for Experiment 1. The crucial region, which is manipulated in the experiment, is marked with a slur, and the first two measures are the target rhythm, the second two measures are a comparison rhythm with an identical crucial region, the third two measures are a comparison rhythm with a crucial region in the same family as the target, and the fourth two measures are a comparison rhythm with a crucial region in a different family to the target.


**FIGURE 3.** The 13 sets of rhythms in Experiment 1. The first rhythm is the target, and the other three are comparisons. The slur marks the crucial region of each rhythm, which is identical to the target in the first comparison, in the same family as the target in the second comparison, and in a different family to the target in the third comparison. The regions outside the crucial area are the same for all three comparisons, but in a different family from the target.

the third two measures have a crucial region in the same family as that of the target, and the fourth two measures have a crucial region in a different family to the target. Because the comparison rhythms are identical to one another outside the crucial area, any difference in their similarity to the target must be a consequence of the differences in the crucial areas.

Participants in a pilot study were concerned about making quantitative judgments of the similarity between rhythms. Hence, in this experiment, the participants heard two pairs of rhythms on each trial (a total of four rhythms) and had to judge which pair had the more similar rhythms. The experiment compared the three possible types of trial; the order of the two pairs in each trial was counterbalanced:

Target: same family in critical region vs. Target: different family in critical region.
Target: identical critical region vs. Target: same family in critical region.
Target: identical critical region vs. Target: different family in critical region.

The first pair in each of these comparisons should be judged as the one with more similar rhythms. The family theory's crucial prediction, however, is an increase in the differences over the three comparisons in the order in which they are shown above. In particular, a rhythm with a crucial region identical to the target's should be rated as more similar than one with a crucial region in the same family, which in turn should be rated as more similar than one with a crucial region in a different family.

### METHOD

*Participants.* Twenty-one undergraduates from Princeton University, 10 men and 11 women, ranging in age from 18 to 22 years, volunteered to participate in the experiment. They were recruited without regard to their music training, but most had a musical background, and they were all familiar with Western music.

*Design.* Figure 3 presents the rhythms used in the experiment: targets and three comparison rhythms for each of them, with crucial regions of a few notes: identical, same family, and different family. The crucial regions were either in the middle or towards the end of a rhythm. There was a mistake in the preparation of rhythm 2 with the crucial region towards the end, and so it is not shown in Figure 3, and we dropped it from the analysis of the experiment. In every place except for the crucial region, the comparisons were the same but differed from the target. The four rhythms in each set in Figure 3 yielded the three trials made up of two pairs of rhythms, and each participant carried out the three trials for all of the rhythms. The resulting 42 trials (including three with the rhythm that was dropped from the analysis) were presented in a different random order to each participant.

*Materials.* The rhythms were created in Sibelius and a sound file was generated for each pair of rhythms. Four beats (one measure of 4/4) of rest separated the two rhythms from each other. Rhythms used a synthesized bongo, an unpitched percussion instrument with a comfortable timbre, a moderate tempo of 100 bpm, a loudness of about 50 dB, and an attack of 10 ms with a slight reverberation, and so it is easy to discriminate intervals between even the shortest interonset intervals (150 ms for $1/16^{\text{th}}$ notes). The difference between rhythms is therefore dependent on the interonset intervals rather than the duration of the notes themselves, which is constant. An example of a typical pair of rhythms can be heard on the webpage: http://psych.princeton.edu/psychology/research/johnson_laird/music.php

*Procedure.* The instructions to the participants explained that their task on each trial was to judge which of two pairs of rhythms sounded more similar. They were told that the first rhythm in each pair was identical, and so the comparison hinged on the second rhythm in each pair. They were also told that there was no time limit, but that they could play each pair only twice. The experiment was presented on a laptop computer under the control of a program written in E-Prime. The participants saw in appropriate positions on the computer's screen two labels: "Left pair" and "Right pair." Under each label was a screen button labeled "Play," and when a participant touched the button, they heard the corresponding pair of rhythms. The instructions stated that they had to play the left pair first, and then the right pair. Beneath the labels for each pair was another circular button that the participants touched to make their response, and beneath these buttons there was a single text box for them to take notes to help them with their judgments. A button labeled "Next" was at the bottom of the screen, and as soon as the participants touched it, the program presented the next trial. As the participants were told, the first two trials were for practice only.

### RESULTS

The family theory predicts the following increasing trend over the three comparisons in the degree to which the first pair is judged as more similar than the second pair:

**TABLE 2.** *Percentages of Predicted Responses in which the First Pair of Rhythms was Rated as More Similar Than the Second Pair in Experiment 1.*

| | The three types of comparison | | |
|---|---|---|---|
| The 14 materials sets of materials | Target: same family vs. Target: different family | Target: identical vs. Target: same family | Target: identical vs. Target: different family |
| Target 1, middle | 55 | 60 | 60 |
| Target 1, end | 65 | 40 | 55 |
| Target 2, middle | 65 | 60 | 65 |
| Target 3, middle | 40 | 65 | 75 |
| Target 3, end | 60 | 65 | 70 |
| Target 4, middle | 85 | 80 | 70 |
| Target 4, end | 30 | 75 | 80 |
| Target 5, middle | 55 | 60 | 60 |
| Target 5, end | 55 | 60 | 75 |
| Target 6, middle | 40 | 65 | 75 |
| Target 6, end | 75 | 70 | 70 |
| Target 7, middle | 70 | 70 | 60 |
| Target 7, end | 45 | 65 | 60 |
| Overall | 57 | 64 | 67 |

Target: same family in critical region vs. Target: different family in critical region: 57%
Target: identical critical region vs. Target: same family in critical region: 64%
Target: identical critical region vs. Target: different family in critical region: 67%

The position of the crucial region in the middle or towards the end had no reliable effect on the participants' choices, and so we have pooled the data from them in the percentages above. The predicted trend over the three conditions above is highly reliable. The small differences between the percentages are slightly misleading because the stochastic increase over the three conditions corroborated the predicted trend in a highly reliable way in by-participants analysis (Page's $L = 258, z = 2.85, p < .0025$). Table 2 presents the percentages over the three conditions for each of the 13 sets of materials, and a by-materials analysis of these data was almost significant (Page's $L = 164, z = 1.57, p < .06$). Of the 20 participants, 17 had a higher proportion of predicted responses than chance (Binomial test $p < .002$). As here, we used nonparametric ("distribution free") statistical tests throughout the present paper because they obviate problems of distribution, and because they allow us to test the reliability of predicted ordinal trends. Unlike analysis of variance, which can test for linear trends, quadratic trends, and so forth, Page's $L$ is a test for a monotonic increase from one condition to another, i.e., a stochastic ordinal trend. In general, nonparametric tests are less powerful than parametric tests such as analysis of variance (see, e.g., Siegel & Castellan, 1988, Sec. 3.4.1), and so they are less likely to lead to an incorrect rejection of the null hypothesis (a Type I error).

The experiment showed that a rhythm and a variant in the same family are more similar to one another than the rhythm and a variant in a different family, but most similar of all is a variant that has the identical rhythm in the crucial region. One possible confound is that the manipulation of the crucial regions in the rhythms also affected the perceptual groups in the rhythms (of the sort described in the Introduction). But, for five of the six rhythms in which the crucial regions are towards the end, the change in the crucial region has no effect on the subsequent perceptual groups in the rhythms in the same family or different family. Hence, this factor cannot account for the difference between them. Another potential confound is that the variant in the same family has one more note than the target, whereas the variant in the different family has one less note than the target. The difference occurs because the direct way to make a change that remains in the same family is to add an eighth-note to a sub-beat, whereas the direct way to make a change that does not remain in the same family change is to omit a note. We doubted whether this confound was responsible for the difference among the conditions, but we carried out a second experiment to check this possibility and to make a more stringent examination of perceptual groups.

## Experiment 2

The previous experiment showed that individuals are sensitive to the same temporal pattern of onsets in


**FIGURE 4.** The six sets of rhythms used in Experiment 2. The first rhythm in each row is the target, and the other four are comparisons. The lines over notes in the staves demarcate a pattern of onsets that is the same as one in the target rhythms. SFSP denotes same family with a same pattern of onsets, SFDP denotes same family with a different pattern of onsets, DFSP denotes different family with same pattern of onsets, and DFDP denotes different family with a different pattern of onsets.

different rhythms. Hence, the present experiment was designed to examine participants' *ratings* of the similarity of rhythms with matching patterns of onsets, or with matching families, or both. The experiment used sets of five types of rhythm: a target rhythm and four comparison rhythms that each departed from the target in different ways. As an example, consider the five rhythms on the first stave of Figure 4 above. They each contain seven notes spread over two measures. The target rhythm is leftmost. Next to it is a rhythm labeled SFSP (for same family and same pattern of onsets). Both rhythms are in the family: | N N N N | N, and both contain within them the same pattern of onsets (two triples of eighth notes separated by a rest of an eighth) shown in the horizontal line over the SFSP rhythm. Next, the SFDP (same family different pattern) rhythm is in the same family as the target, but does not contain any pattern of onsets of at least three notes that are identical to those in the target. The DFSP (different family same pattern) rhythm is in the family: | N N S N | N, which differs from the family of the target, but the rhythm has a sequence of five onsets identical to one that occurs in the target. Finally, the DFDP (different family different pattern) rhythm is in a different family – | N N S N | N, again – from the target, and it also has no sequence of at least three onsets in it that are the same as the target. In sum, the design manipulates family and pattern of onsets independently from one another. Figure 4 above presents the six sets of such rhythms used in the experiment.

The procedure was similar to the one in Experiment 1 in that each trial used two pairs of rhythms, but differed in that the participants rated the similarity of the rhythms in the first pair and then the similarity of the rhythms in the second pair. The reason for presenting two rhythms on each trial was, as in Experiment 1, to provide a context in which it was easier for the participants to make ratings of similarity. The experiment allowed us to contrast the family theory with the predictions of perceptual groups (e.g., Handel, 1998; Povel & Essens, 1985). Perceptual groups predict that the edit distance apart of any pair of rhythms should determine their similarity. In contrast, the family theory predicts the following trend regardless of edit distance: two identical target rhythms should be most similar, then two rhythms in the same family including a same pattern of onsets (SFSP); then either two rhythms including a same pattern of onsets (DFSP) or two rhythms in the same family (SFDP), and finally two rhythms that are in different families that do not include a same pattern of onsets (DFDP). As this partial trend shows, the theory makes no prediction about the relative similarities of same family and same pattern of onsets, and so a further purpose of the experiment was to determine which, if any, has the greater effect on ratings of similarity.

## METHOD

*Participants.* Twenty students (including undergraduates, graduates, and others) were recruited on a voluntary basis from Princeton University and the Berklee School of Music. Their mean age was 21.4 years and there were 12 females and 8 males. The mean number of years of music training was 9 years. All participants were familiar with Western music, and about half were equally familiar with non-Western music.

*Design.* All the participants carried out 24 trials in which they rated the similarity of one pair of rhythms and then of another pair of rhythms. The target rhythm was first in every pair, and there were 6 different targets and their comparisons, and four types of trial depending on which pairs of rhythms occurred on a trial:

| | | |
|---|---|---|
| 1. | target vs. target | target vs. same family and pattern |
| 2. | target vs. same family and pattern | target vs. same family only |
| 3. | target vs. same family only | target vs. same pattern only |
| 4. | target vs. same pattern only | target vs. neither same family nor pattern |

The order of the two pairs in a trial was counterbalanced, and the 24 trials were presented in a different random order to each of the participants.

*Materials.* Figure 4 presents the full set of 6 target rhythms and their comparisons based on the following families, which all started and ended with N: NNNN | N, NNNN | N, NONN | N, NNON | N, NSNN | N, and NNSN | N. The comparisons were of four sorts defined only in terms of the crucial region, which was either in the middle or at the end of a rhythm: same family and pattern (as the target), same pattern only, same family only, and neither the same pattern nor the same family.

*Procedure.* The participants were tested individually in a quiet room. The computer presented a description of the task, and the key instructions were as follows:

> On each trial, you will be presented with four rhythms. Your task is to rate how similar the two rhythms in the first pair are, and then how similar the two rhythms in the second pair are. The first rhythm in each pair is the same.

The instructions then explained that the participants were to make their ratings by moving a slider on the screen, where the left end demarcated identity and the right end demarcated maximum dissimilarity.

There were two practice trials to give participants practice with the task before the experiment began. The practice trials included a pair that was identical and a pair with neither same family nor pattern in order to give participants experience with the range of possibilities. After they listened to each pair of rhythms, the participants rated their similarity by moving the slider on the computer screen to the appropriate scale point, i.e., they rated their similarity on a seven-point Likert scale from 1 (*most similar*) to 7 (*most dissimilar*). The computer recorded the participant's similarity rating, and the participant pressed a button to go on to the next trial. There was no time limit, and so the participants responded in their own time. At the end of the experiment, they filled out a brief questionnaire concerning their musical experience and their thoughts about what made rhythms similar.

## RESULTS

Table 3 presents the mean ratings for the five comparisons for each of the six sets of rhythms. Figure 5 presents the overall mean ratings for the similarity between the target and the five rhythms. The means showed that identical targets were rated as highly similar (1.34 on the seven-point scale). The results reliably corroborated the predicted partial trend in similarity between the target and comparison rhythms: same family and same pattern of onsets (3.16) < same pattern of onsets only (4.19) or same family only (5.25) < different family and different pattern of onsets (6.02; Page's $L = 598$, $p < .0000005$). The family theory did not predict any difference between same family only and same pattern of onsets only, but the participants rated rhythms with the same pattern of onsets as more similar to the target than those with the same family as the target (Wilcoxon test, $z = 3.85$, $p < .0005$, two-tail).

The theory of perceptual groups takes no account of meter, and so similarity should depend on edit distance, i.e., the closer the distance between the two rhythms in a pair the greater should be their rated similarity (see our earlier account). Table 4 presents the edit distances for the five comparisons with each of the six rhythms and shows that it is possible for rhythms to occur in all four of the main comparisons yet have the same edit distance from the target (see target rhythm 1). Because both the family theory and perceptual groups predict that a comparison of a rhythm with itself should yield ratings of maximal similarity, we computed the rank-

**TABLE 3.** *The Rated Similarity in Experiment 2 of Five Types of Comparisons with Six Different Rhythms.*

| | The five types of comparison | | | | |
|---|---|---|---|---|---|
| The six sets of rhythms | Target | Same family Same pattern (SFSP) | Different family Same Pattern (DFSP) | Same family Different Pattern (SFDP) | Different family Different pattern (DFDP) |
| Target 1 | 1.50 | 3.83 | 4.65 | 6.08 | 6.25 |
| Target 2 | 1.00 | 3.30 | 4.10 | 4.78 | 6.35 |
| Target 3 | 1.15 | 3.18 | 4.08 | 4.93 | 5.70 |
| Target 4 | 1.10 | 2.78 | 4.20 | 5.58 | 6.15 |
| Target 5 | 1.85 | 2.95 | 4.44 | 4.83 | 5.85 |
| Target 6 | 1.45 | 2.78 | 3.88 | 5.35 | 5.80 |
| Overall | 1.34 | 3.16 | 4.19 | 5.25 | 6.02 |

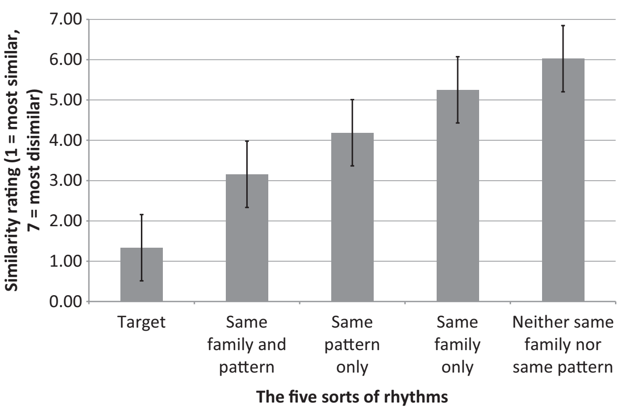



**FIGURE 5.** The mean ratings of similarity (1 = *most similar*, 7 = *most dissimilar*) and standard error bars for the five types of comparison in Experiment 2.

order correlation with edit distance for the four comparisons when the target-to-target ratings were dropped from the analysis (Kendall's $\tau = .64$, $p < .001$). We also computed the rank-order correlation with the family theory for the same four comparisons (Kendall's $\tau = 0.96$, $p < .0000001$). The results show that the family theory yields a very high fit to the ratings of similarities – a fit so high shows that the predicted trend generalizes over the six different sets of rhythms.

Overall, the results supported the family theory, showing that rhythms in the same family as the target yield an increase in similarity over those in a different family, both for rhythms in which the patterns of onsets are the same as the target and for cases in which the pattern of onsets are different. The family theory yields a better fit than perceptual groups. One obvious limitation of the experiment, however, is that that task of rating the similarity of rhythms is artificial: musicians and listeners could spend a lifetime of listening to music without ever carrying it out. In our next experiments, we accordingly turned to a more natural task.

## Experiment 3

The aim of the experiment was to test the family theory using the reproduction of rhythms. The participants listened to a rhythm and then attempted to tap the same rhythm at a tempo well within the constraints on accurate performance (Repp, 2005). Half the rhythms ended in a syncopation and half did not. Five of the six rhythms ending in a syncopation had an onset on the last eighth subdivision (sub-beat) of the measure, and so the syncopation anticipated the first beat of a second virtual measure (see Figure 6b below). This sort of syncopation is common in modern jazz. The family theory predicts that errors should be more likely to occur in trying to reproduce a rhythm ending in a syncopation. The theory also predicts that errors should be more likely to yield a rhythm in the same family as the original rhythm than in a different family. On half of the trials, the participants carried out the task immediately after hearing the rhythm, but on the other half of the trials they heard a distracting musical sequence between the original rhythm and their performance of the task. This distraction should increase errors.

### METHOD

*Participants.* Twenty-one individuals from the same population as those in Experiment 1 (11 male and 10 female) volunteered to participate in this experiment, which took about 20 min. Four of them had no music training, five had some music training, and 12 had extensive music training, including music theory. All were familiar with Western music. No music training was required to participate in the experiment.

*Design.* Each participant carried out the task for twelve rhythms. The rhythms were presented in one of four different orders, and an order was assigned randomly to a participant, but with the constraint that equal

**TABLE 4.** *Edit Distance Between the Target Rhythm and Each Comparison Rhythm in Experiment 2.*

| | Comparison rhythm | | | | |
|---|---|---|---|---|---|
| | Target | Same pattern of onsets and same family (SFSP) | Same pattern of onsets only (DFSP) | Same family only (SFDP) | Neither same pattern of onsets nor same family (DFDP) |
| Target rhythm 1 | 0 | 4 | 4 | 4 | 4 |
| Target rhythm 2 | 0 | 2 | 2 | 4 | 5 |
| Target rhythm 3 | 0 | 2 | 4 | 4 | 6 |
| Target rhythm 4 | 0 | 2 | 6 | 4 | 6 |
| Target rhythm 5 | 0 | 2 | 4 | 4 | 5 |
| Target rhythm 6 | 0 | 2 | 4 | 2 | 4 |



numbers of the four orders occurred in the experiment as a whole. For each participant, six of the trials were randomly assigned to have a musical distraction intervening between the rhythm and the task of reproducing it, and the remaining six trials had no such distraction. In the experiment as a whole, each rhythm occurred about equally often in the distraction and no-distraction conditions.

*Materials.* Figure 6 below presents the 12 rhythms in the experiment. Each rhythm lasted for one measure of four beats, and contained five or six notes – well within the constraints of working memory (e.g., Miller, 1956). Six of the rhythms ended with a syncopation, and six did not. The rhythms were created in Sibelius using the same procedure as in Experiment 1. Four clicks of a metronome using "high wood blocks" preceded each rhythm in order to set the beat, but did not last through the duration of the rhythm, which was played on a synthesized bongo similar to the one in the previous experiments but with no reverberation and an attack of slightly less than 10 ms. Before the participants tried to reproduce a rhythm, they heard another four clicks of the metronome, which continued during their input. The rhythm that they tapped was also played in the synthesized bongo. The six musical distractions each consisted of a 30 s sequence of upbeat and highly rhythmic music, such as *Supercalifragilisticexpialidocious* and the *Vitamin String Quartet*, which had faster tempi than the experimental rhythms. They were created using the Audacity editing program, and immediately followed the rhythm to be reproduced in order to prevent the participants from using auditory rehearsal of the rhythm. The resulting files, 12 with the distraction and 12 without the distraction, were used to construct the sound files in the E-Prime program, which controlled the experiment.

*Procedure.* The instructions to the participants explained that their task was to listen to a rhythm, which was sometimes followed by an irrelevant musical extract, and then to tap the rhythm as exactly as possible on a computer key. When the musical distraction occurred on a trial, they had to listen to it before they tried to reproduce the target rhythm. There was no time limit on when the participants had to make their response, so they were free to rehearse the rhythm before they tried to reproduce it. The first three trials of the experiment were treated as practice trials, though the participants were not aware of the status of these trials. One of these trials included a musical distraction.

One computer was used to present the rhythms, and another computer was used to record the participants' reproductions: they tapped the rhythm on the G4 key of a MIDI keyboard, which was connected to the second computer running the Sibelius Flexitime program. This program makes an automatic transcription of music, and the resulting notation was displayed on the computer's screen so that the participants could see it, though they could not go back to edit their responses. The Flexitime system and its settings are described on http://www.sibelius.com/helpcenter/index.html. The settings under "Flexitime options" were as follows: the flexibility of tempo was medium, and the minimum duration was an eighth note. Before the start of the experiment, participants were instructed on how to use Flexitime, and they practiced tapping rhythms in order to become familiar with the procedure. Because the minimum duration of Flexitime was set to an eighth note, a reasonable amount of flexibility existed so that if participants were slightly off the metronome beat, their response was still correctly transcribed. The settings therefore allowed for a compromise between strict accuracy and reasonable performance. Flexitime's accuracy was confirmed in the participants' survey responses. None of the participants voiced any major concerns in response to the question posed in a questionnaire after the experiment, "Did you have any difficulties using the notation program? Did you notice

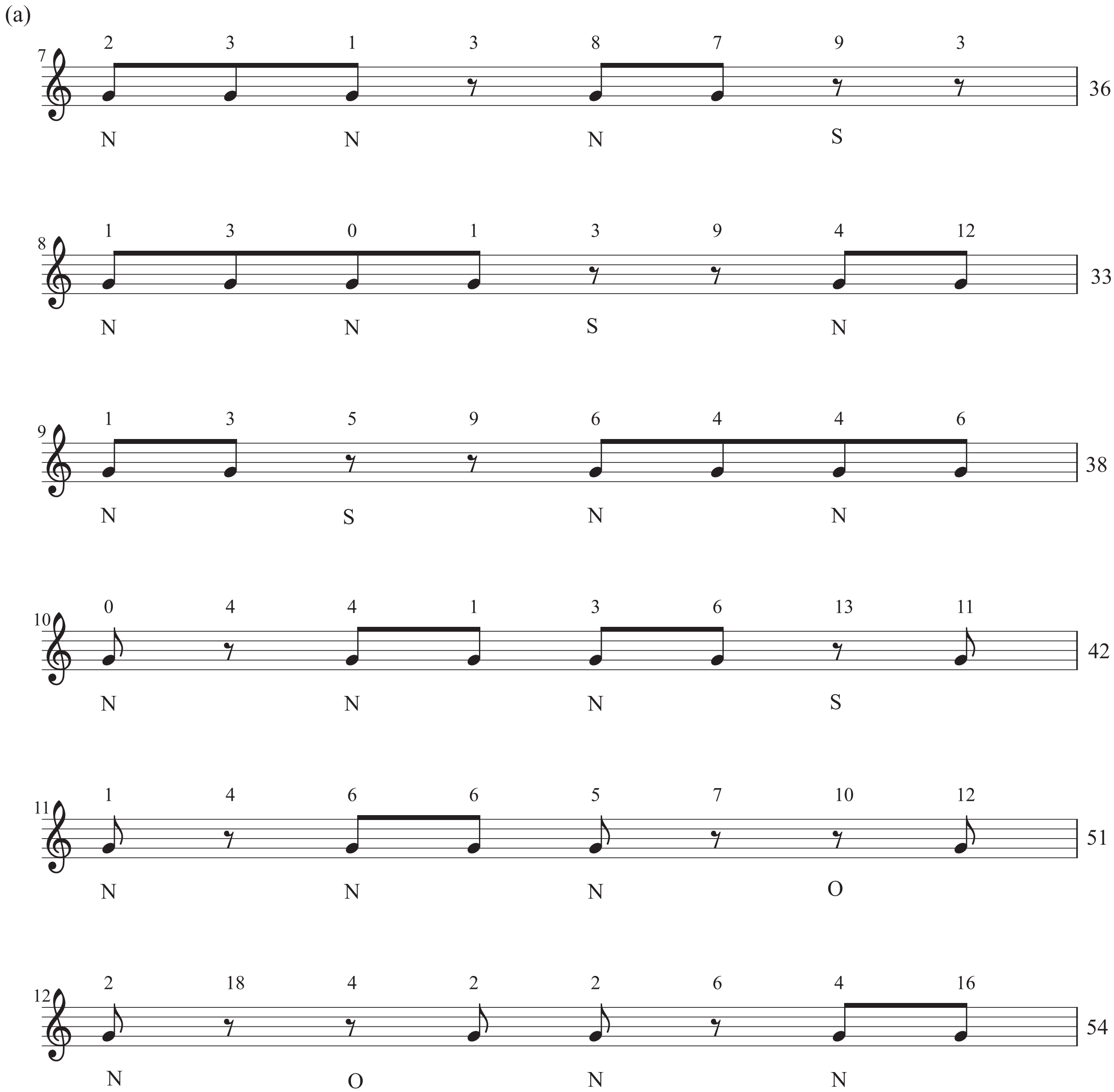



**FIGURE 6.** The 12 rhythms in Experiment 3, their familial status in which N denotes a note with an onset on the beat, S denotes a syncopation anticipating the beat, and O denotes all other events, the distribution of errors over each eighth beat division of the measure, and the total number of errors for each rhythm (shown on the right-hand side of the rhythms). (a) The six rhythms that do not end in a syncopation, and (b) the six rhythms that end in a syncopation, which yields S on the first beat of the next virtual measure (see the next page).

any notation errors from the rhythm you intended to tap and the rhythm that was notated by the program?"

RESULTS

A total of 15 out of the 252 reproduced rhythms were excluded from analysis because the participant made no response, or the rhythm was outside the scope of the Flexitime settings. Because edit distance is insensitive to meter, we used a different measure of accuracy. We scored each sub-beat (at the eighth note level) in an original rhythm for whether it was correctly reproduced or not. If participants performed a note in place of a rest on a sub-beat, or vice versa, it counted as a single error. For example, the two errors on the fourth eighth note

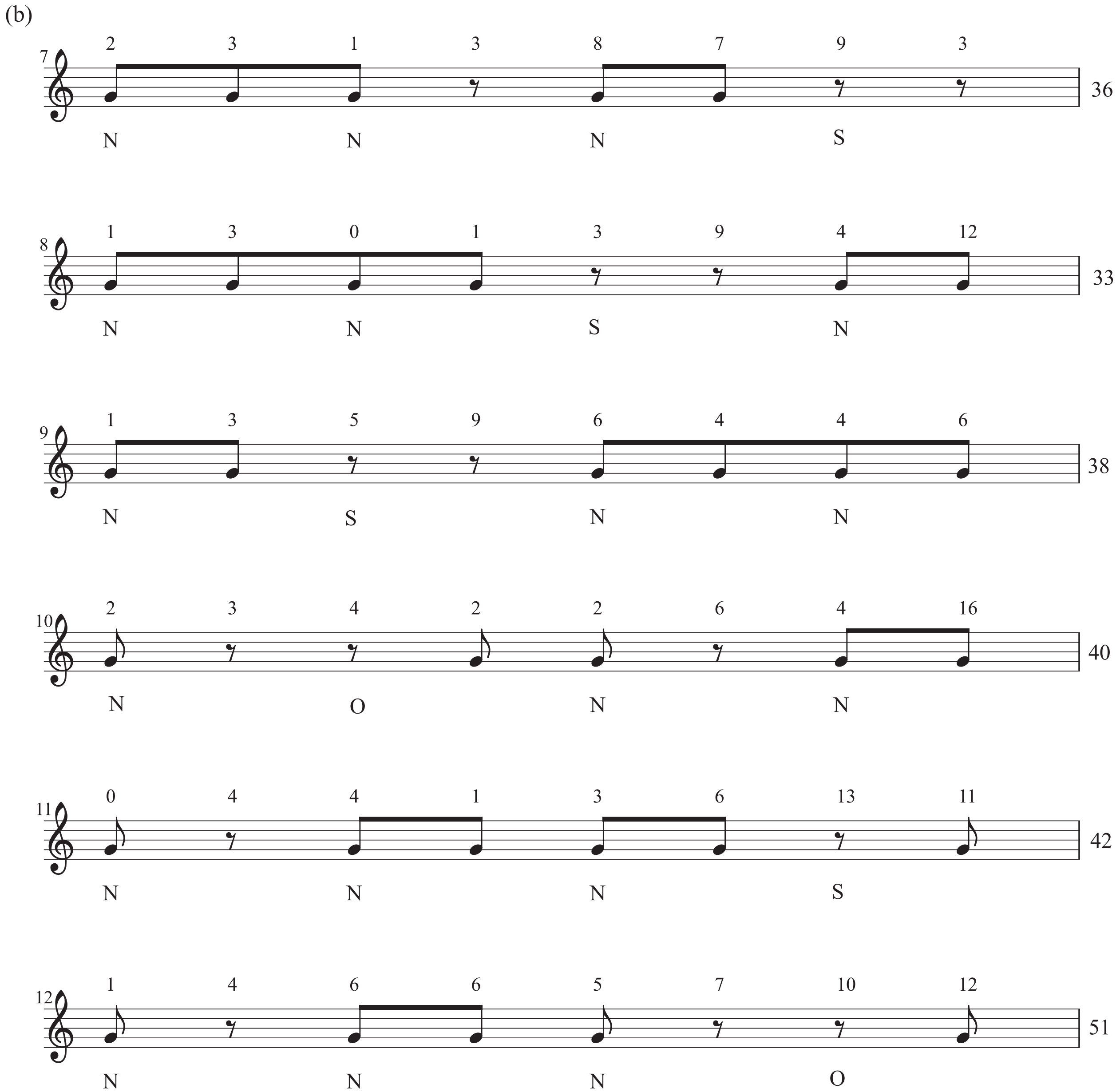



**FIGURE 6.** Continued.

in rhythm 1 of Figure 6a indicate that on two trials the participants erred in using a rest instead of a note. Table 5 presents the percentages of trials on which the rhythms were correctly reproduced for the two types of rhythm – those ending with a syncopation, S, and those ending with an onset on the fourth beat, N, and in the two conditions – with or without the intervening musical distraction. As Table 5 shows, there were more correct trials without the musical distraction (Wilcoxon, $z = 3.08$, $p < .001$), and there were more correct trials for rhythms that ended on a note rather than a syncopation (Wilcoxon, $z = 3.33$, $p < .0004$). The two variables did not interact reliably (Wilcoxon $z = 1.19$, $p > .10$). A by-materials analysis of the rhythms in Figure 6 confirmed that errors were more likely to occur with the six rhythms ending in a syncopation than in the six rhythms that do not end in a syncopation (Mann-Whitney test, $z > 2.80$, $p < .005$). A single trial could yield more than one error, and Table 6 presents the total numbers of errors (and the mean number of errors per trial) yielding

**TABLE 5.** *Percentages of Correct Trials for Rhythms in Experiment 3.*

| | No intervening distraction | Intervening distraction | Overall |
|---|---|---|---|
| Rhythms with no syncopation at the end (1-6) | 73 | 66 | 69 |
| Rhythms with a syncopation at the end (7-12) | 47 | 10 | 28 |
| Overall | 60 | 38 | 49 |

*Note*: The balances of the percentages are trials with at least one error.

**TABLE 6.** *Total Numbers of Errors (Mean Numbers per Trial in Parentheses) in Experiment 3.*

| | Errors in the same family as original | Errors in a different family from the original |
|---|---|---|
| No intervening distraction | 74 (0.62) | 42 (0.35) |
| Intervening distraction | 87 (0.74) | 69 (0.59) |
| Overall | 161 (0.68) | 111 (0.47) |

a rhythm in the same family as the original and the mean number of errors per trial yielding a rhythm in a different family from the original. Overall, the participants made more errors in the same family (a mean of 0.68 per trial) than in a different family (a mean of 0.47 per trial, Wilcoxon, $z = 3.08$, $p < .001$). Not surprisingly, more errors were made after a distraction than in immediate responses (Wilcoxon, $z = 2.0$, $p < .025$). The trend was for a bigger bias towards errors in the same family when no intervening distraction occurs, but the interaction was not reliable (Wilcoxon, $z = .81$, $p > .2$).

Figure 6a presents the six rhythms that did not end in a syncopation, and the total number of errors on each eighth of a beat in the measure for them, and Figure 6b presents the six rhythms that ended in a syncopation, and the number of errors for them. The rhythms without a syncopation at the end yielded a mean of .99 errors, whereas those with a syncopation at the end yielded a mean of 1.42 errors (Wilcoxon test, $z = 3.74$, $p < .0001$). The increase in errors from the start to the end of the measure was highly reliable (Page's $L = 3889$, $z = 6.695$, $p < .0000003$). The results showed that the twelve rhythms differed reliably in the difficulty of reproducing them (Friedman nonparametric analysis of variance, $\chi^2 = 65.11$, $df = 11$, $p < .00000001$). Despite their difference in music training, the participants did not differ reliably in their ability to reproduce rhythms (Friedman nonparametric analysis of variance, $\chi_2 = 26.40$, $df = 20$, $p > .10$). The experiment overall showed a bias towards errors yielding a rhythm in the same family as the original rather than a different family. The results also suggested that the occurrence of a syncopation, which perturbs meter, increased the likelihood of an error.

## Experiment 4

The previous experiment showed that syncopations, S, with their tendency to disturb the meter, elicited more errors in the reproduction of rhythms than notes on the beat, N. We therefore carried out a replication of the previous experiment in which we systematically manipulated the rhythms that the participants had to reproduce. There were two related sets of nine rhythms. The rhythms in the first set lasted for a single measure, and the rhythms in the second set had an identical first measure to those in the first set but were followed by a single note on the first beat of a second measure. In the nine rhythms that lasted one measure, three varied whether the fourth beat was N, O, or S, three varied whether the third beat was N, O, or S, and ended with N on the fourth beat, and three ended with an eighth note on the final eighth sub-beat, i.e., a syncopation, S, anticipating the first beat of an imaginary second measure. For the rhythms in the second set matching these last three rhythms, the note on the first beat of the second measure eliminates the syncopation, S, at the end of the first measure. Figure 7 below presents the full set of rhythms used in the experiment, and Table 7 below presents the families of the rhythms. The participants acted as their own controls and tried to reproduce the 18 rhythms presented to each of them in a different random order. All the rhythms were followed by an intervening musical distraction prior to the participants' attempts to reproduce them.

The family theory makes two main predictions. First, as before, errors should tend to occur in the same family as the original rhythms rather than in a different family. Second, there should be an interaction between the two main variables in Table 7. Rhythms in which the fourth beat was N or S (the second and third rows in Table 7) should be easier in the one-measure case than in the two-measure case, which adds an additional note to be remembered. But, this effect should be smaller when a syncopation in the one-measure rhythm is eliminated in the two-


FIGURE 7. The 18 rhythms in Experiment 4, their familial status in which N denotes a note with an onset on the beat, S denotes a syncopation anticipating the beat, and O denotes all other events, the distribution of errors over each eighth beat division of the measure, and the total number of errors for each rhythm (shown in bold on the right-hand side of the rhythms). The left-hand side of the figure shows the 9 one-measure rhythms, where the numbers in parentheses at the end of the measure are errors in which participants produced a note at the start of the virtual second measure, and the right-hand side of the figure shows the 9 matching rhythms that have an additional beat on a second measure.

**TABLE 7.** *Percentages of Correct Trials in Experiment 4.*

| | One-measure rhythms | Two-measure rhythms |
|---|---|---|
| The families of the two rhythms (one measure and two measure) | | |
| \| N N N v \| S vs. \|N N N v \| N | 60 | 55 |
| \| N N v N \| vs. \| N N v N \| N | 58 | 49 |
| \| N N v S \| vs. \| N N v S \| N | 58 | 25 |

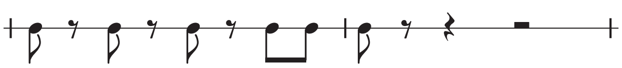
*Note*: Balances of the percentages are trials with at least one error. Families are defined in terms of N: a note with an onset on the beat, S: a syncopation, with an onset prior to the beat, and O: other types of event, such as a rest, and 'v' denotes a variable ranging over N, S, and O for different rhythms.

measure rhythm. For example, consider a rhythm such as (the first one-measure rhythm in Figure 7):

It is in the family:

| N N N N | S

because the final eighth note is a syncopation anticipating the first beat of the next (virtual) measure. The following rhythm adds an extra note that occurs on the first beat of the second measure (see the first two-measure rhythm in Figure 7):



This extra note eliminates the syncopation, and the rhythm is instead in the family:

| N N N N | N.

The additional note should therefore add less to the difficulty in reproducing the rhythm than other types of additional notes in the two-measure rhythms.

## METHOD

*Participants.* Twenty-four undergraduates from the same population sampled in the previous experiment volunteered to participate in the experiment, which lasted for about half an hour. Their mean number of years of music training was 10.5 years, and ranged from 0 years to 17 years, but all of the participants were familiar with Western music.

*Materials.* Figure 7 above presents the 18 rhythms used in the experiment. To keep the rhythms as varied as possible, the four possible combinations for N on the first two beats: quarter quarter, quarter eighth-eighth, eighth-eighth quarter, and eighth-eighth eighth-eighth, were distributed across the rhythms. The rhythms and the distraction files were created in the same way as in the previous experiment.

*Procedure.* The key instructions were:

> On each trial, you will hear a short musical rhythm. The rhythm will then be immediately followed by a short music clip. After listening to *both* the rhythm and music clip, your task is to reproduce the rhythm (but *not* the music clip) by tapping it into a music notation program. Each rhythm will be preceded by 4 clicks of the metronome. The four clicks are simply there to set the beat—they are not part of the rhythm. You will hear a total of 21 rhythms, and so the study lasts for only about 30 minutes.

Three simple practice trials were included in the beginning of the experiment, but as far as the participants knew, they were part of the experiment proper. We changed one setting for the Flexitime program from the previous experiment: the flexibility of the tempo was set to "none," because the "medium" setting had allowed some participants to slow down in their tapping, and we aimed for all the participants to tap at the same rate throughout the experiment.

## RESULTS

Two out of the 432 reproduced rhythms were not included in the analysis because participants made no response. Table 7 presents the percentages of trials in which the participants correctly reproduced the rhythms. The participants made accurate reproductions of the one-measure rhythms more often than of the two-measure rhythms (Wilcoxon test, $z = 2.71$, $p < .005$). A by-materials analysis of the 9 matched pairs of rhythms in Figure 7 corroborated this difference (Wilcoxon test, $z > 1.95$, $p < .03$). There was a reliable interaction between the number of measures and the three types of family of rhythms. As Table 7 shows, the one-measure rhythms yielded comparable numbers of accurate reproductions, whereas the two-measure rhythms differed reliably (Friedman nonparametric analysis of variance, $\chi r^2 = 18.91$, $df = 2$, $p < .00025$,

**TABLE 8.** *Total Numbers of Errors in the Same Families as the Originals and in Different Families from the Originals in Experiment 4.*

| The families of the two rhythms (one measure and two measure) | One-measure rhythms | | | Two-measure rhythms | | |
|---|---|---|---|---|---|---|
| | Different family errors | Same family errors | Totals | Different family errors | Same family errors | Totals |
| \| N N N v \| S<br>\|N N N v \| N | 30 | 39 | 69 | 26 | 28 | 54 |
| \| N N v N \|<br>\| N N v N \| N | 39 | 43 | 82 | 41 | 54 | 95 |
| \| N N v S \|<br>\| N N v S \| N | 38 | 36 | 74 | 46 | 60 | 106 |

*Note*: There were three different rhythms in each of the six families (a total of 18 rhythms), but the onsets in the first measures were identical in the matching one- and two-measure rhythms. 'N' denotes a note with an onset on the beat, 'S' denotes a syncopation with an onset prior to the following beat, 'O' denotes any other sort of onset, and 'v' had as its value N, S, O, in the three rhythms, respectively.

two-tailed). In particular, one-measure rhythms ending in an S or an N on the fourth beat were less accurately reproduced in the two-measure versions because of the additional note. However, one-measure rhythms ending in an S on the first beat of a virtual second measure (the first row in Table 7) showed less of an increase in difficulty in the second-measure case. In these cases, the additional note in the two-measure version changed the first beat of the second measure from an S into an N. This predicted interaction was reliable (Wilcoxon test, $z = 2.61$, $p < .005$). Even though only three rhythms eliminated S, the by-materials interaction was marginally significant (Mann-Whitney test, $z = 1.42$, $p < .08$).

As in the previous experiment, a single trial could yield more than one error, and Table 8 presents the number of errors yielding a rhythm in the same family as the original and the number of errors yielding a rhythm in a different family from the original. In the analyses of errors, we did not include sub-beat 9 of the two-measure rhythms, because a fair comparison needs the same number of sub-beats in both rhythms. Not surprisingly, the total errors in Table 8 show a complementary pattern to the percentages of accuracies in Table 7, and so there were reliably fewer errors for one-measure rhythms than for two-measure rhythms (Wilcoxon test, $z = 1.72$, $p < .05$). The participants made more errors in the same family as the original rhythms (mean = .60) than errors in a different family (mean = .51), and the difference was significant (Wilcoxon, $z = 1.64$, $p < .05$). Errors that changed the family of a rhythm were more likely to convert a note into a syncopation or vice versa (58% of responses) than to be of the other four sorts (42%), and this difference was almost significant (Wilcoxon test, $z = 1.57$, $p < .06$). An analysis of errors allowed us to make an independent estimate of the effects of N, O, or S, as the value of the variable, v, in the design of the experiment (see Table 7). The total numbers of errors were, respectively, 111, 179, and 190, and the predicted trend was reliable (Page's $L = 308.0$, $z = 2.89$, $p < .0025$).

Figure 7 presents the nine pairs of rhythms and the total numbers of errors on each of the nine sub-beats, and Figure 8 presents the total numbers of errors on each of the nine sub-beats for the one- and two-measure rhythms. As Figure 7 shows, a large proportion of errors occurred on the ninth sub-beat of the two-measure rhythms, consisting for the most part of failures to produce the note on the beat. If we ignore this sub-beat, a reliable trend of an increase in errors occurs both over one-measure rhythms (Page's $L = 5876.0$, $z = 4.58$, $p < .00001$) and over two-measure rhythms (Page's $L = 6311.5$, $z = 8.77$, $p < .0000005$). As Figure 8 makes clear, the errors for one-measure rhythms tend to peak around sub-beat 6, whereas the trend continues to increase to the end of the two-measure rhythms.

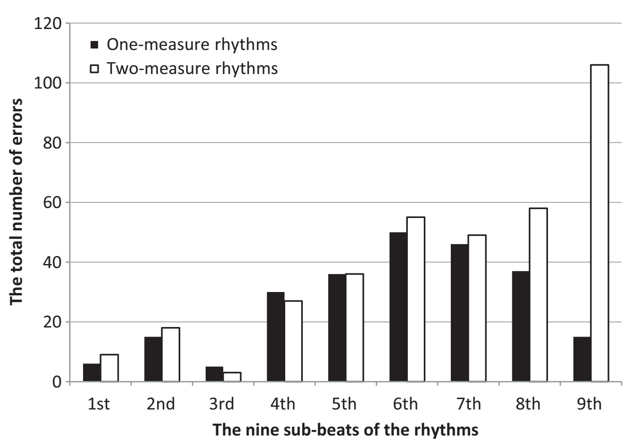



**FIGURE 8.** The total numbers of errors in Experiment 4 on each of the sub-beats of the nine pairs of one measure and the two measure rhythms.

The participants differed reliably in their ability to reproduce rhythms accurately (Friedman $\chi r2 = 42.73$, $df = 23$, $p < .025$, two-tailed). The best participant made 100% correct reproductions, whereas the worst participant made only 6% correct reproductions. The cause of this difference in ability is unknown. But, as a reviewer suggested, it may reflect discrepancies between the participants' preferred tempi and the actual tempi of the rhythms. It did not correlate with the number of years of music training ($R^2 = .02$).

## General Discussion

The family theory of musical rhythms presupposes two distinct cognitive systems in the perception and production of music – an intuitive system and a deliberative system. The former, which is known as "System 1," delivers rapid and automatic intuitions, and, because it has no access to a working memory for intermediate results, it cannot carry out recursive processes such as counting beyond single digits. It does not even have the power of a finite-state automaton (see Hopcroft & Ullman, 1979), because it can carry out a loop of mental operations for only a small number of times. This restriction is perhaps responsible for the appeal of cyclical accounts of meter (e.g., London, 2004). In contrast, the deliberative system, which is known as "System 2," underlies slower and intentional cognitions, and, as it can use working memory, it can carry out recursive processes, such as full-scale counting. The distinction between the two systems is familiar in studies of higher cognition (e.g., Evans, 2008; Johnson-Laird, 1983; Kahneman, 2011; Khemlani & Johnson-Laird, 2012). Some skeptics, however, have argued that perhaps there is only a single system (e.g., Kruglanski, 2013), and others have raised the question of whether the theory is testable (e.g., Keren, 2013). These issues go far beyond the scope of the present paper, but the family theory is clearly refutable. System 1 is crucial for the perception of rhythms – unlike performers, listeners tend not to count measures – and also for the improvisation of rhythms (Johnson-Laird, 2002). Sensitivity to patterns of onsets and to families of rhythms also depends on System 1, because individuals with no training in music are sensitive to them at least in simple cases (see, e.g., Experiment 3). If individuals were sensitive to patterns of onsets and families of rhythms only as a result of explicit musical training and the use of counting, rhythm would depend on System 2. Conversely, if individuals could identify meters of any number of beats without having to count, then no need would exist to use System 2 for this task.

The family theory postulates that the similarity between two rhythms depends both on their patterns of onsets and on their respective families. If two rhythms contain a sequence of the same pattern of onsets, then they should tend to sound similar even if one rhythm is staccato and the other is legato. Likewise, if they contain a sequence in the same family, then they should tend to sound similar. Family, however, depends on meter, because only meter allows syncopations to occur in a systematic way (see the Introduction). Meter yields three types of musical event: 1) notes reinforcing the meter because they occur on metrically important units (N), 2) notes unsettling the meter because they are syncopations with onsets anticipating metrically important units (S), and 3) other events that neither reinforce nor disturb the meter (O). Our experiments focused on the metrical unit of beats, and at this level the theory assigns to each beat one of three categories of event: notes on the beat (N), syncopations anticipating the beat and lasting through its onset (S), and all other types of event that occur between one beat and the next (O).

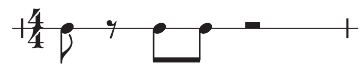
Our experiments corroborated the family theory, both for judgments of the similarity of rhythms and for the reproduction of rhythms. When two rhythms share a critical region, they are most similar if their critical regions contain the same pattern of onsets over three notes, quite similar if they are in the same family, and least similar if they contain a different pattern of onsets in a different family (Experiment 1). Analogous results occurred with comparisons between entire rhythms, as opposed to just brief critical regions within them. As one would expect, the greatest similarity is between a target rhythm and itself. There then follows a robust trend in which a rhythm that includes the same pattern of onsets, and is in the same family as a target rhythm, is rated as more similar than one that has only the same pattern of onsets, which in turn is more similar than one that is only in the same family, and the least similar rhythm shares neither a pattern of onsets nor the family of the target (Experiment 2). This experiment showed that pattern of onsets is more important than same family – a difference about which the theory had made no prediction. We suppose, however, that it reflects the greater importance of a factor common to all rhythms – the temporal pattern of onsets – than a factor that is relevant only to those musical rhythms based on a meter.

Experiment 2 also allowed us to examine the potential role of perceptual groups in judgments of similarity; that is, groups defined by the sequences of contiguous onsets in rhythms without regard to meter (see, e.g., Handel, 1992; Povel & Essens, 1985). Its natural measure of similarity is the 'edit' distance apart of a pair of rhythms (e.g., Toussaint et al., 2011). It correlated reliably with similarity: the smaller the edit distance between two rhythms, the greater their ratings of similarity (in Experiment 2). However, the family theory yielded a much larger correlation which, unlike edit distance, was hardly affected when we dropped the comparison of a rhythm with itself from the analysis. We attribute the difference to the fact that edit distance weights all edits equally, and therefore does not distinguish between syncopations and notes on the beat.

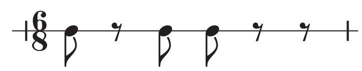
Few occasions occur in listening to music when individuals judge the similarity between rhythms, a process likely to elicit System 2 in addition to System 1. We therefore examined the more common task of reproducing a rhythm. The family theory predicts its difficulty. When individuals reproduce a rhythm, they are more accurate when a rhythm does not end with a syncopation than when it does (Experiment 3) – a result that corroborates earlier findings that syncopations elicit errors in reproduction (Fitch & Rosenfeld, 2007). When individuals err, the resulting rhythm is also more likely to be in the same family as the original rhythm than in a different family. One unexpected result that merits a closer examination is that musical experience did not correlate with accuracy in reproducing rhythms. This result bears out the importance of system 1 in carrying out the task.

The likelihood of an error in the reproduction of a rhythm depends in part on the number of notes in the rhythm – more notes place a greater load on working memory. Hence, one-measure rhythms become more difficult to reproduce when they are followed by a note (N) at the start of a second measure (Experiment 4). The family theory predicts an exception to this principle. When a one-measure rhythm ending with a syncopated note (S) becomes a two-measure rhythm with an actual note at the start of the second measure, this note eliminates the syncopation (S). In consequence, the additional note has a lesser effect on the difficulty of reproducing the rhythm. In general, errors in the reproduction of rhythms were more likely to yield a rhythm in the same family than in a different family.

The principal shortcoming of our experimental studies is perhaps inevitable. We examined only short and simple rhythms. Given the set of possible rhythms based on quarter and eighth notes, the sets of rhythms in our studies are reasonably representative of those in common time. However, there are two potential ways in which our results might fail to generalize, though we have no substantial reason for skepticism. First, they may not generalize to other meters. At the level of beats, the status of notes (N), syncopations (S), and other events (O) does not change with switches to duple or triple meters. But, a change from an even to an odd number of units within the beat can change the family of a rhythm. In common time:



is an instance of | N N S O |. It remains in an analogous family in 3/4: | N N S | O. But, the same pattern of onsets in 6/8:



is in the family | N N |. Second, the findings may fail to scale up to long and complex rhythms. We surmise that empirical studies of such rhythms are less likely to yield a clear pattern of errors, or a need for different rhythmic principles to those in the family theory. We did contemplate an experiment in which musicians improvised variations on a given rhythm, and this sort of experiment may show that the constraints of family are less important when individuals strive to be creative.

Are there fundamental rhythmic principles outside the scope of the family theory insofar as the present results are concerned? Perceptual groups play an important role in the discrimination of nonmetrical rhythms (e.g., Handel, 1992). We have shown that they have some power in predicting similarity among metrical rhythms, but not as much power as the family theory, perhaps because perceptual groups take no account of meter and therefore do not distinguish between notes on the beat and syncopations. Some music outside common practice may have non-isochronous meters (Kvifte, 2007; London, 2004). We have no definitive evidence on this matter, but if such meters do occur then the concept of a syncopation may itself stand in need of revision. And the family theory may need to be revised for the undifferentiated category of other (O) events, which lump together ties, rests, and anacruses ('pick up' notes). The theory might give a better account of the perception or imitation of rhythms if this category were to distinguish between anacruses and cases in which no notes occur. This matter could certainly be addressed in empirical studies that contrast the two types of case in a systematic way.

Some genres of music differ in the families that they exemplify. A phrase in the family: O | N N N N | N S O, is much more likely to occur in modern jazz than in a Christmas carol. But, genres differ in rhythmic ways outside the family theory. Jazz performances usually "swing," and waltzes performed in Vienna have a characteristic "lilt." These are systematic distortions of rhythms within a metrical framework (see, e.g., Gabrielsson, Bengtsson, & Gabrielsson, 1983; Honing & de Haas, 2008). Another characteristic of some music is polyrhythm – the co-occurrences of distinct rhythms. Much of common practice music consists of a melody and its accompaniment. West African music, such as the drumming of the Ewe, likewise consists of a solo improvisation with a complex rhythmic accompaniment (Locke, 1998). This distinction is analogous to the figure-ground contrast in vision (Toussaint et al., 2011). Some music, however, depends on a concurrence of important rhythms, as in counterpoint. The family theory allows an analysis of separate rhythms, but the resulting complex rhythm often results in nothing more than the sequence, N N N N . . . , in the case of metrical music. Such music can be handled by System 1. But, some music calls for System 2, and readily goes beyond a useful analysis in terms of musical families of rhythms, e.g., some of the player-piano works of Conlon Nancarrow are based on simultaneous but distinct meters or pulses that diverge in tempo, as in his *Study 21,* with one pulse speeding up as the other slows down.

No-one knows the role of families, if any, in the processes underlying improvisation and composition, in part because of the difficulty of examining them experimentally. The machinery for creating rhythms is likely to depend on an interplay between System 1 and System 2. Improvisation does not appear to rely on a working memory for the results of intermediate computations, but instead on just a small memory buffer, which serves as "echoic" memory for what has just been improvised (Johnson-Laird, 2002). It accordingly depends on System 1 and long-term memory. A finite-state automaton can generate its output, perhaps using just a matrix of transitions (see Huron, 2006) or a hidden markov model (see Mavromatis, 2009). Composition, however, makes use of deliberation (System 2), and so it can contrive rhythms with strict long-term dependencies, i.e., ones that consistently follow certain rules. Because System 1 is not deliberate, in some cases it follows and in other cases it violates long-term dependencies purely according to chance. As computational power increases, its effect is to prevent such violations. Another matter outside the present account is the complexity of rhythms (Shmulevich & Povel, 2000). Musicians and listeners agree that rhythms differ in complexity, but no one has succeeded in pinning down an appropriate psychological measure of complexity, e.g., Thul and Toussaint (2008) considered 32 different metrics. The difficulty in establishing a cognitively plausible measure is to find a reliable and valid method for studying human judgments of complexity.

To encapsulate our argument: meter matters. Rhythms are identifiable by the patterns of their onsets, but what distinguishes them in common practice is meter. In turn, meter itself depends on the mental organization of events in a way that is hierarchical, though the recursion is limited by the system of intuitions (System 1) underlying music. One function of meter is to make possible rhythms that are neither too predictable nor too unpredictable. It allows them to vary in predictability and it may be the only systematic way to do so within the computational bounds of the intuitive System 1. As a consequence of meter, System 1 is also sensitive to three main types of musical event, which define families of rhythms: events that reinforce the meter (onsets on the beat), events that are neutral with respect to meter (rests, ties, and anacruses), and events that disturb the meter (syncopations). As our experiments have shown, these three types of event are reflected both in judgments of the similarity between rhythms and in the errors that occur when individuals try to reproduce a rhythm. Only musical rhythms are metrically organized, and meter appears to be a convention much like tonality – a cultural invention within the innate constraints of musical cognition.

## Author Note

This research was carried out with support to the third author from Princeton University. We are grateful to Marco Ragni for advice about Levenshtein's edit distance. We thank Lucy Cao, Sangeet Khemlani, Gary Marcus, and Godfried Toussaint, for helpful discussions. We are also grateful to Eric Clarke, Stephen Handel, and two anonymous referees, for their cogent criticisms of previous drafts of this article.

*Correspondence concerning this article should be addressed to* P. N. Johnson-Laird, Visiting Scholar, Department of Psychology, Room 409, New York University, 6 Washington Place, New York, NY 10003. E-mail: phil@princeton.edu

## References


Aslin, R. N., Jusczyk, P. W., & Pisoni, D. B. (1998). Speech and auditory processing during infancy: Constraints on and precursors to language. In D. Kuhn & R. Siegler (Eds.), *Handbook of child psychology: Cognition, perception, and language* (pp. 147-254). New York: Wiley.

Bamberger, J. (1982). Growing up prodigies: The midlife crisis. *New Directions for Child Development*, *17*, 61-77.

Bamberger, J. (2003). Music as embodied mathematics: A study of a mutually informing affinity. *International Journal of Computers for Mathematical Learning*, *8*, 123-160.

Bolton, T. L. (1894). Rhythm. *American Journal of Psychology*, *6*, 145–238.

Clarke, E. F. (1984). Structure and expression in rhythmic performance. In P. Howell, I. Cross, & R. West (Eds.), *Musical structure and cognition* (pp. 209–236) New York: Academic Press.

Clarke, E. F. (1999). Rhythm and timing in music. In D. Deutsch (Ed.), *The psychology of music* (pp. 473–500). San Diego, CA: Academic Press.

Collins, N. (2006). Towards a style-specific basis for computational beat tracking. In M. Baroni, A.R. Addessi, R. Caterina, & M. Costa (Eds.), *Proceedings of the 9th International Conference on Music Perception and Cognition* (pp. 461-467). Bologna, Italy: ICMPC and ESCOM.

Cook, N. (2012). *Harmony, perspective, and triadic cognition*. Cambridge, UK: Cambridge University Press.

Cooper, G., & Meyer, L. B. (1960). *The rhythmic structure of music*. Chicago, IL: University of Chicago Press.

Damerau, F. J. (1964). A technique for computer detection and correction of spelling errors. *Communications of the Association CM*, *7*, 171-176.

Deliège, I. (1987). Grouping conditions in listening to music: An approach to Lerdahl & Jackendoff's grouping preference rules. *Music Perception*, *4*, 325-360.

Desain, P., & Honing, H. (1999). Computational models of beat induction: The rule-based approach. *Journal of New Music Research*, 28, 29-42.

Desain, P., & Honing, H. (2003). The formation of rhythmic categories and metric priming. *Perception*, *32*, 341-365.

Essens, P. (1995). Structuring temporal sequences: Comparison of models and factors of complexity. *Perception and Psychophysics*, *57*, 519-532.

Essens, P. J., & Povel, D. J. (1985). Metrical and nonmetrical representations of temporal patterns. *Perception and Psychophysics*, *37*, 1-7.

Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgment and social cognition. *Annual Review of Psychology*, 59, 255-278.

Fitch, W. T., & Rosenfeld, A. J. (2007). Perception and production of syncopated rhythms. *Music Perception*, *25*, 43-58.

Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 149-180). London, UK: Academic Press.

Gabrielsson, A. (1973). Similarity ratings and dimension analyses of auditory rhythm patterns. *Scandinavian Journal of Psychology*, *14*, 138-160. doi:10.1111/j.1467-9450.1973.tb00105.x

Gabrielsson, A., Bengtsson, I., & Gabrielsson, B. (1983). Performance of musical rhythm in 3/4 and 6/8 meter. *Scandanavian Journal of Psychology*, *24*, 193-213.

Geiser, E., Ziegler, E., Jancke, L., & Meyer, M. (2009). Early electrophysiological correlates of meter and rhythm processing in music perception. *Cortex*, *45*, 93-102. doi:10.1016/j.cortex.2007.09.010

Grondin, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception, and Psychophysics, 72*, 561-582.

Handel, S. (1992). The differentiation of rhythmic structure. *Perception and Psychophysics*, *52*, 497-507.

Handel, S. (1998). The interplay between metric and figural rhythmic organization. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1546-1561.

Handel, S., & Oshinsky, J. S. (1981). The meter of syncopated auditory polyrhythms. *Perception and Psychophysics*, *30*, 1-9.

Hannon, E. E., & Johnson, S. P. (2005). Infants use meter to categorize rhythms and melodies: Implications for musical structure learning. *Cognitive Psychology*, *50*, 354-377.

Hannon, E. E., Snyder, J. S., Eerola, T., Krumhansl, C. L. (2004). The role of melodic and temporal cues in perceiving musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 956-974. doi: 10.1037/0096-1523.30.5.956

Hannon, E. E., & Trehub, S. E. (2005). Metrical categories in infancy and adulthood. *Psychological Science*, *16*, 48-55.

Harnad, S. (1987). *Categorical perception: The groundwork of cognition*. Cambridge, UK: Cambridge University Press.

Hébert, S., & Cuddy, L. L. (2002). Detection of metric structure in auditory figural patterns. *Perception and Psychophysics*, *64*, 909-918.

Honing, H. (2012). Without it no music: Beat induction as a fundamental musical trait. *Annals of the New York Academy of Sciences*, *1252*, 85-91.

Honing, H., & De Haas, W. B. (2008). Swing once more: Relating timing and tempo in expert jazz drumming. *Music Perception*, *25*, 471-476.

Honing, H., Ladinig, O., Háden, G. P., & Winkler, I. (2009). Is beat induction innate or learned? Probing emergent meter perception in adults and newborns using event-related brain potentials. *Annals of the New York Academy of Sciences*, *1169*, 93-96.

Hopcroft, J. E., & Ullman, J. D. (1979). *Formal languages and their relation to automata*. Reading, MA: Addison-Wesley.

Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation.* Cambridge, MA: MIT Press.

James, W. (1890). *The principles of psychology* (Vol. 1). New York: Henry Holt.

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, MA: Harvard University Press; Cambridge: Cambridge University Press.

Johnson-Laird, P. N. (1991). Rhythm and meter: A theory at the computational level. *Psychomusicology*, *10*, 88-106.

Johnson-Laird, P. N. (2002). How jazz musicians improvise. *Music Perception*, *19*, 415-442.

Jones, M. R., & Boltz, M. (1989). Dynamic amending and responses to time. *Psychological Review*, *96*, 459-491.

Kahneman, D. (2011). *Thinking fast and slow*. New York: Farrar, Strauss, Giroux.

Keren, G. (2013). A tale of two systems: A scientific advance or a theoretical stone soup? Commentary on Evans & Stanovich (2013). *Perspectives on Psychological Science*, *8*, 257-262.

Khemlani, S., & Johnson-Laird, P. N. (2012). The processes of inference. *Argument and Computation*, *4(1)*, 4-20.

Krebs, H. (1999). *Fantasy pieces: Metrical dissonance in the music of Robert Schumann*. New York: Oxford University Press.

Kruglanski, A.W. (2013). Only one? The default interventionist perspective as a unimodel –commentary on Evans & Stanovich (2013). *Perspectives on Psychological Science*, *8*, 242-247.

Kvifte, T. (2007). Categories and timing: On the perception of meter. *Ethnomusicology*, *51*, 64-84.

Ladinig, O., Honing, H., Háden, G., & Winkler, I. (2009). Probing attentive and pre-attentive emergent meter in adult listeners with no extensive music training. *Music Perception*, *26*, 377-386.

Large, E. W. (2008). Resonating to musical rhythm: Theory and experiment. In S. Grondin (Ed.), *The psychology of time* (pp. 189–231). Bingley, UK: Emerald.

Large, E. W., & Kolen, J. F. (1994). Resonance and the perception of musical meter. *Connection Science*, *6*, 177-208.

Large, E. W., & Palmer, C. (2002). Perceiving temporal regularity. *Cognitive Science*, *26*, 1-37.

Large, E. W., & Snyder, J. S. (2009). Pulse and meter as neural resonance. *Annals of the New York Academy of Sciences, 1169*, 46-57.

Lee, C. S. (1991). The perception of metrical structure: Experimental evidence and a model. In P. Howell, R. West, & I. Cross (Eds.), *Representing musical structure* (pp. 59-127). London: Academic Press.

Lerdahl, F., & Jackendoff, R., (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Levenshtein, V. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, *10*, 707-710.

Locke, D. (1998). *Drum Gahu: An introduction to African rhythm*. Gilsum, NH: White Cliffs Media.

London, J. (2004). *Hearing in time*. New York: Oxford University Press.

London, J. (2008). Rhythms in twentieth-century theory. In T. Christensen (Ed.), *Cambridge history of music theory* (pp. 695–725). New York: Cambridge University Press.

Longuet-Higgins, H. C. (1979). The perception of music. *Proceedings of the Royal Society (London), B205*, 307-322.

Longuet-Higgins, H. C., & Lee, C. S. (1984). The rhythmic interpretation of monophonic music. *Music Perception*, *1*, 424-441.

Mavromatis, P. (2009). Minimum description length modelling of musical structure. *Journal of Mathematics and Music*, *3*, 117-136.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *63*, 81-97.

Orpen, K. S., & Huron, D. (1992). Measurement of similarity in music: A quantitative approach for nonparametric representations. *Computers in Music Research*, *4,* 1-44.

Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 728-741.

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

Patel, A. D., Iversen, J. R., Bregman, M. R., & Schulz, I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Current Biology*, *19*, 827-830.

Pitt, M. A., & Monahan, C. B. (1987). The perceived similarity of auditory polyrhythms. *Perception and Psychophysics*, *41*, 534-546.

Post, O., & Toussaint, G. (2011). The edit distance as a measure of perceived rhythmic similarity. *Empirical Musicology Review*, *6*, 164-179.

Povel, D-J. (1984). A theoretical framework for rhythm perception. *Psychological Research*, *45*, 315-337.

Povel, D-J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception*, *2*, 411-440.

RAGNI, M., KHEMLANI, S., & JOHNSON-LAIRD, P. N. (2014). The evaluation of the consistency of quantified assertions. *Memory and Cognition*, *42*, 53-66.

REPP, B. H. (2005). Sensorimotor synchronization: A review of tapping literature. *Psychonomic Bulletin and Review*, *12*, 969-992.

ROSS, J., & HOUTSMA, A. J. M. (1994). Discrimination of auditory temporal patterns. *Perception and Psychophysics*, *56*, 19-26.

SAMAROTTO, F. (1999). Strange dimensions: Regularity and irregularity in deep levels of rhythmic reduction. In C. Schachter & H. Siegel (Eds.), *Schenker studies 2* (pp. 222-238). Cambridge, UK: Cambridge University Press.

SHAFFER, L. H., CLARKE, E. F., & TODD, N. P. (1985). Metre and rhythm in piano playing. *Cognition*, *20*, 61-77.

SHEPARD, R. (1999). Cognitive psychology and music. In P. Cook (Ed.), *Music, cognition, and computerized sound* (pp. 21-35). Cambridge, MA: MIT Press.

SHMULEVICH, I., & POVEL, D-J. (2000). Measures of temporal pattern complexity. *Journal of New Music Research*, *29*, 61-69.

SIEGEL, S., & CASTELLAN, N. J. (1988). *Nonparametric statistics for the behavioral sciences* (2nd ed.). New York: McGraw-Hill.

STANOVICH, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah, NJ: Erlbaum.

STEEDMAN, M. J. (1977). The perception of musical rhythms and metre. *Perception*, *6*, 555-569.

STOFFER, T. H. (1985). Representation of phrase structure in the perception of music. *Music Perception*, *3*, 191-220.

SUMMERS, J. J., HAWKINS, S. R., & MAYERS, H. (1986). Imitation and production of interval ratios, *Perception and Psychophysics*, 39, 437-444.

THUL, E., & TOUSSAINT, G. T. (2008). Rhythm complexity measures: A comparison of mathematical models of human perception and performance. In M. Baroni, A. R. Addessi, R. Caterina, & M. Costa (Eds.), *Proceedings of the Ninth Conference of International Society of Music Information Retrieval* (pp. 663-668). Bologna, Italy: ICMPC and ESCOM.

TOUSSAINT, G. T. (2013). *The geometry of musical rhythm: What makes a "good" rhythm good?* Boca Raton, FL: Chapman and Hall/CRC Press.

TOUSSAINT, G. T., CAMPBELL, N., & BROWN, N. (2011). Computational models of symbolic rhythm similarity: Correlation with human judgments. *Analytical Approaches To World Music, 1*, 380-430.

VOLK, A. (2008). The study of syncopation using inner metric analysis: Linking theoretical and experimental analysis of metre in music. *Journal of New Music Research*, *37*, 259-273.

YESTON, M. (1976). *The stratification of musical rhythm*. New Haven, CT: Yale University Press.

ZENTNER, M., & EEROLA, T. (2010). Rhythmic engagement with music in infancy. *Proceedings of the National Academy of Sciences*, *107*, 5768-5773.