# Subjective Rhythmization: A Replication and an Assessment of Two Theoretical Explanations

Rasmus Bååth
*Lund University, Lund, Sweden*

**Subjective rhythmization is that phenomenon** whereby, when one is listening to a monotone metronome sequence, some sounds are experienced as accented. These subjectively accented sounds group the sequence similarly to how the metrical structure of a piece of music groups the beats. Subjective rhythmization was first investigated by Bolton (1894); the present study aims at replicating and extending that work. Consistent with Bolton's results, all participants reported hearing accent patterns when listening to monotone sequences; the reported group size of an accent pattern was highly dependent on the tempo of the sequence. A power relation captured well the relation between the reported group size and the sequence interstimulus interval. Further, the mean group size reported in the subjective rhythmization task was found to correlate with the timing performance in a slow-tempo tapping task. These results are consistent with the *resonance theory* explanation of subjective rhythmization (Large, 2008).

*Received: June 23, 2014, accepted January 1, 2015.*

**Key words:** subjective rhythmization, subjective accentuation, the tick-tock effect, rhythm perception, resonance theory

Rhythm, the temporal organization of distinct sound events, is an integral part of human speech and music (Patel, 2008). Humans have an astonishing capability both to perceive and to produce rhythms. *Subjective rhythmization* (SR) is one example of this capacity. This is the phenomenon whereby sounds of a monotone metronome sequence are experienced as having different intensity and that these intensity differences follow a regular pattern. In other words, despite the sounds having objectively equal amplitude, they are perceived as subjectively different. Bolton (1894) developed an experimental paradigm for investigating SR; only one other study exists that uses Bolton's original paradigm (Vos, 1973). The current study aims to replicate and extend Bolton's and Vos' work. Extensions include using a wider range of tempi, employing a larger number of participants, and presenting those participants with a number of auxiliary tasks in addition to the SR task. The inclusion of the auxiliary tasks is motivated by three decisive predictions developed from two proposed explanations for SR: the *preferred tempo explanation* (Temperley, 1963) and the *resonance theory explanation* (Large, 2008).

A typical example of SR is when identical ticks of a clock are perceived as "tick tock" (Brochard, Abecasis, Potter, Ragot, & Drake, 2003; van Noorden & Moelants, 1999). For this reason, SR has also been called the *clock illusion* or the *tick-tock effect* (Vlek, Schaefer, Gielen, Farquhar, & Desain, 2011). An alternative way of viewing SR is as the imposition of a subjective meter onto a sequence of sounds, where no meter is enforced through physical intensity or physical pitch differences. It has been pointed out that the term subjective rhythmization is a misnomer and that a more suitable term would be subjective meter (Large, 2008) or subjective accentuation (Temperley, 1963).

Subjective rhythmization was discussed already in the 18th century (Kirnberger, 1776) but not investigated experimentally until Bolton's (1894) seminal work. Bolton used apparatus capable of producing isochronous (temporally equally spaced) sequences of monotone clicks of equal amplitude. By systematically varying the tempi of the sequences he established the following characteristics of SR. Isochronous sequences of identical sounds produce the impression that some sounds are louder or more intense than others. The apparent increases in intensity do not appear randomly but recur every *n*th sound, resulting in the more intense sounds grouping the sequence. Here *n* can range from two up to eight but the most common reported groupings participants reports are two, three, and four: the common metrical groupings of Western music. Group size and tempo are related; participants report smaller groupings at slower tempi and larger groupings at faster tempi. The range of tempi at which SR can be experienced is limited. Bolton found that SR experience ceases when the interstimuli interval (ISI) between consecutive sound onsets rises above 1600 ms, though a later review

*Music Perception*, volume 33, issue 2, pp. 244–254, issn 0730-7829, electronic issn 1533-8312.  DOI: 10.1525/mp.2015.33.2.244

of Bolton's results suggested a slower limit of 1800 ms (Fraisse, 1982).

Only one study, that by Vos (1973), has employed Bolton's (1894) experimental paradigm, despite recent interest in the electrophysiological properties of SR (e.g., Nozaradan, Peretz, Missal, & Mouraux, 2011; Schaefer, Vlek, & Desain, 2011). Vos' study, though limited by a relatively small number of trials and narrow tempo range of the stimuli (ISIs of 150 to 800 ms), produced results in accord with Bolton's. Subsequent analysis of Vos's data by van Noorden and Moelants (1999) emphasizes (1) the dependency between tempo and reported group size, (2) a propensity toward reporting even-numbered groups, and (3) an average interval between each group's onset longer than one second.

### EXPLANATIONS FOR WHY SUBJECTIVE RHYTHMIZATION OCCURS

The literature offers two explanations for SR: one relating to participants' preferred tempo (Temperley, 1963) and one explaining SR using the resonance theory of rhythm perception (Large, 2008).

*The preferred tempo explanation.* When experiencing SR, one hears the sounds of a monotone sequence as grouped, with the first sound in each group being accented. This grouping of the sounds can be viewed as a modification of the period of the sequence, where the *group period* is defined as the period between group onsets. An example of such a modification would be when a participant is given a monotone tone sequence with an ISI of 250 ms and reports a grouping of two, resulting in a group period of 500 ms. The preferred tempo explanation is that participants experience a grouping that results in a group period close to their preferred tempo (Temperley, 1963) so as to facilitate entrainment to the sequence.

A regular observation is that, when participants are asked to tap an isochronous rhythm at a comfortable rate, the resulting tempi tend to cluster around a period of 500-600 ms (Fraisse, 1982). This tempo is called the *spontaneous motor tempo* (SMT) and has been shown to be strongly correlated ($r = .75$) with participants' verbal reports of preferred beat tempo (McAuley, Jones, Holub, Johnston, & Miller, 2006), supporting the existence of an intrinsic preferred rate for event tracking; the SMT may be seen as the tempo where rhythm perception is optimal (Moelants, 2002).

Present knowledge about SR does not favor the preferred tempo explanation, however. Especially problematic is the observation that the group period tends to be above one second (Vos, 1973, as analysed by van Noorden & Moelants, 1999) which is far from the common period of spontaneous motor tempo.

*The resonance theory explanation.* The resonance theory of rhythm perception (Large & Jones, 1999; Large & Kelso, 2002; van Noorden & Moelants, 1999), offers an alternative explanation. According to resonance theory, experiencing the beat of a piece of music or an isochronous sequence of sounds is an emergent phenomenon, caused by neural *oscillatory circuits* that resonate with incoming auditory events. An oscillatory circuit (henceforth *oscillator*) with intrinsic period $T$ entrains to sound events with a similar period. More specifically, sound events with period $T$ cause the amplitude of oscillators with similar periods to increase. The resulting oscillator amplitude indicates the extent to which events of period $T$ occurred in the auditory stream.

Neural resonance is a common theme underlying resonance accounts of rhythm perception. Within this framework though, models differ in whether they model beat perception using a small number of oscillators or a large network of oscillators. The resonance theory explanation of SR assumes the latter, motivated by the observation that the brain encodes information using populations of neurons (Averbeck, Latham, & Pouget, 2006). By assuming multiple oscillators, the account allows for modeling meter perception involving the temporal organization of beats on multiple time scales (Large & Kolen, 1994).

Models using multiple oscillators (e.g., Large, 2000, and Scheirer, 1998) differ in implementation, but the basic mechanism is the same. A network of oscillators, where each oscillator has an intrinsic period, is given an auditory input. The amplitude of an oscillator with period $T$ reflects the extent to which sound events with period $T$ occurred in the auditory stream. The sum of the amplitudes of all oscillators in the network reflects periodicities in the auditory stream. Precisely what periodicities the network is sensitive to depends on the distribution of the intrinsic periods of the oscillators.

Following Large (2008), an explanation for SR using this multiple-oscillator version of resonance theory is based on the notion that an isochronous sequence of sounds with period $T$ will, in addition to entraining oscillators attuned to that period, entrain oscillators at subharmonics of $T$ (i.e., $2 \cdot T$, $3 \cdot T$, $4 \cdot T$, etc.). The summed output from a network of oscillators will contain amplitude fluctuation at the subharmonic frequencies of the given sequence, even if the sequence itself has no fluctuations in amplitude. See Figure 1, where the sound sequence activates both the oscillator with matching period (Oscillator 1) and the oscillator with a period

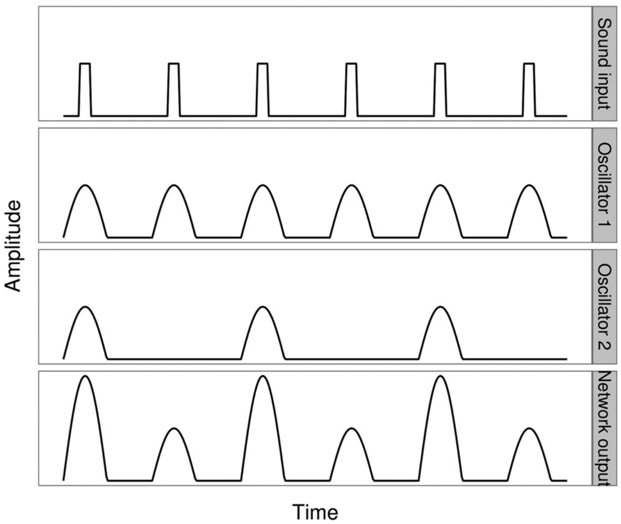


**FIGURE 1.** Schematic plot of subjective rhythmization in a resonance theory framework.

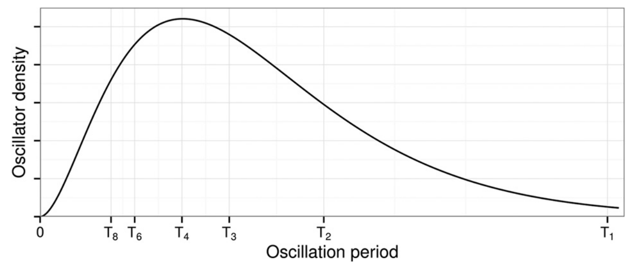


**FIGURE 2.** Schematic plot of the oscillator density as a function of the period.

that is twice as slow (Oscillator 2), resulting in an SR with a grouping of two (Network output). In support of this account, Nozaradan et al. (2011) found that, when participants are asked to listen to an isochronous sound sequence and subjectively impose an accent on every second beat, the resulting electroencephalogram reveals a sustained response at the period of the imposed accent.

Two other aspects of SR can be explained by a multiple oscillator resonance model: (1) why the feeling of SR disappears when the tempo is sufficiently slow and (2) why the size of the perceived groups, and consequently the number of sound onsets between subjective accents, is larger when the tempo is faster. Experiencing SR while listening to a sequence with period $T$ requires oscillators that have at least twice the period of $T$, otherwise there would be no oscillators to mark every second (third, fourth, fifth, etc.) sound of the sequence. The vanishing point of SR then depends on the *slower limit* of rhythm perception; that is, the period where the oscillator density is sufficiently low so that it is not possible to entrain reliably to a rhythm of that period. This is illustrated in Figure 2 where $T_1$ is the longest period to which the model is able to entrain and $T_2$ is the longest period at which SR is still experienced. The size of the perceived groups grow as the period of the sequence becomes shorter because there exist oscillators at higher order subharmonics relative to the period of the sequence. As Figure 2 shows, $T_3$ marks the period for which $T_1$ is the third subharmonic of $T_3$, that is, $T_3$ marks the period for which one finds slow enough oscillators to put an accent on every third beat, resulting in an SR with a grouping of three. Further, $T_4$ and $T_8$ mark the periods for which $T_1$ is the fourth and eight subharmonic, respectively.

*Predictions arising from the two explanations of subjective rhythmization.* The two alternative explanations above make a number of predictions regarding participants' behavior in an SR task as well as relations between that behavior and behavior in other tasks measuring aspects of rhythm perception and production.

The first prediction regards the average group period in the SR task. Remember that the preferred tempo explanation predicts a subject's average group period to be close to her preferred tempo. According to the resonance theory explanation, on the other hand, the group period depends on the slower limit of rhythm perception ($T_1$ in Figure 2), and should fall somewhere between the slower limit of SR ($T_2$) and $T_1$. These two predictions are clearly distinct: preferred tempi, measured using an SMT task, tend to center on a period of 500 ms, while a slower limit of rhythm perception is believed to be above 1500 ms (Repp, 2006). This slower limit can be estimated by way of the *slow motor tempo* task in which a participant is asked to tap as slowly as possible while still maintaining a continuous, regular rhythm (McAuley et al., 2006).

The resonance theory explanation makes a second prediction, regarding the functional relation between the period of the stimulus sequence and the experienced group size in the SR task. As Figure 2 shows, the maximum possible group size $g$ for a sequence with period $T$ depends on the slower limit of rhythm perception $T_1$, so that $g \sim \frac{T_1}{T}$. This can be written more generally as the power function $g \sim \frac{k}{T^a}$, where $k$ equals $T_1$ in the case where the constant exponent $a$ equals 1. Plotted on log-log axes, power laws plot as a straight line with a slope determined by the exponent: $\log(g) \sim \log(k)\ \text{-}a{\cdot}\log(T)$.

The resonance theory explanation makes a third prediction regarding the relation between the SR task and sensorimotor synchronization performance at slow tempi. Within the resonance theory framework, both rhythm perception and rhythm production rely on the same mechanism: the entrainment of neural oscillatory circuits to regularities in the sequence of sounds. Both the slow limit of rhythm perception and rhythm production performance at slow tempi depend on the period at which there cease to be sufficient oscillators to entrain reliably to a sound sequence with corresponding period. The expectation is that participants with a relatively fast slower limit of rhythm perception should struggle to synchronize to a rhythmic stimulus at slow tempi. As noted, the group period in an SR task is expected to be close to a participant's slower limit of rhythm perception; therefore, the *mean group period* can be seen as a proxy variable for that participant's slower limit. One can obtain a measure of synchronization performance at slow tempi by measuring variability in a finger tapping task, where participants are asked to tap in synchrony with isochronous sequences (Repp, 2005). By giving participants both sequences that are comfortably paced and ones that are in the area of the slower limit of rhythm perception, one can factor out variability due to slow tempo from variability due to motor response.

Together, these predictions motivate the inclusion of three auxiliary tasks when extending the SR task introduced by Bolton's (1894): an SMT task, a slow motor tempo task and a taping task using slow pacing sequences.

## Method

### PARTICIPANTS

Nine female and 21 male participants, ranging in age from 19 to 78 years ($M = 31.6$, $SD = 12.8$), were recruited from the Lund community. All were unpaid volunteers. All reported being right handed. Twenty-six reported experience playing a musical instrument, of which ten reported playing or practicing regularly for more than ten years.

### STIMULI AND APPARATUS

The stimuli for the SR task were isochronous sequences of click sounds created with a click-track generator included in the sound editor Audacity (http://audacity team.org/). Each click consisted of a 440 Hz sine wave of 10 ms. Each sequence consisted of 15 s of clicks repeated at a constant ISI. Sequences were presented at eight tempi, corresponding to click ISIs of 150, 200, 300, 600, 900, 1200, 1500, and 2000 ms. The sequence with an ISI of 2000 ms is slower than the proposed slower limit of SR (Fraisse, 1982); participants were expected to report no SR while listening to it. Its inclusion was for detecting any subjects who misinterpreted instructions.

For the SMT task, slow motor tempo task and the tapping task, participants used a custom-built tapping board consisting of a piezoelectric sensor mounted on a 5 $cm^2$ piece of corrugated fiberboard (see Bååth, 2011, for details). Participants tapped on the pad using their right index finger, with their hand resting on a plastic foam cushion. For the tapping task, the stimuli consisted of isochronous sequences of 440 Hz square wave tones of 20 ms. Each sequence consisted of 31 tones. Sequences were presented at five tempi, corresponding to tone ISIs of 600, 1200, 1800, 2400, and 3000 ms. An Arduino microcontroller controlled both generation of sounds and registration of taps. All stimuli were delivered through full-sized head phones (Philips SHP2500).

### PROCEDURE

Participants were tested individually in a quiet room. The experimental tasks comprised an SR task, a tapping task, an SMT task, and a slow motor tempo task, all performed during a single session which, on average, lasted one hour. The order of the SR task and the tapping task was randomized so that the SR task preceded the tapping task for 15 of the 30 participants. The SMT and the slow motor tempo tasks consisted of three trials each. The SMT trials were interleaved between the SR and the tapping task while the slow motor tempo trials were presented last. See Figure 3 for a flowchart of the experimental procedure.

*The subjective rhythmization task.* Each participant was placed in front of a computer with head phones. Prior to the task a 600 ms ISI click sequence was played and the participants were informed that all clicks in the sequence were equally loud and equally spaced. Each participant was asked if she nevertheless experienced a grouping of the clicks or if some clicks were more dominant. The possible groupings of the sequence were explained, from none up to a grouping of eight. The 600 ms ISI click sequence was replayed. At this point, all participants reported experiencing a grouping of the clicks. These instructions conform to those described by Andrews (1905) in his discussion of Bolton's work as a *Test of Involuntary Rhythmisation with Suggestion*.

Participants then began the task proper, which consisted of four blocks of eight trials each: one for each click-sequence ISI level. The order of the trials within each block was randomized. Each participant was asked to attend to each sequence and report the first grouping

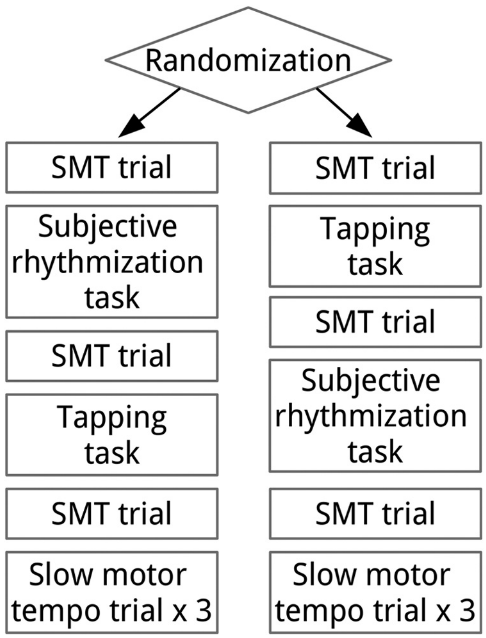


**FIGURE 3.** Flowchart of the experimental procedure.

that she experienced. This was done using a computer interface by selecting the appropriate alternative from a drop-down list with the alternative "No grouping/ groups of one" and alternatives "Groups of two" up to "Groups of eight" (translated from Swedish).[1] The task was self paced and no participant was interrupted while engaged in the task.

*The tapping task.* Each participant sat, wearing head phones, in front of the tapping board and was asked to adjust the volume to a comfortable level while a tone sequence was played. The tapping task consisted of four blocks of five trials each, one for each ISI level. The order of the trials within each block was randomized. A trial consisted of each participant tapping along with a tone sequence, using her dominant hand. Participants were instructed to tap along to each tone sequence, to start tapping as soon as the sequence began, and to stop tapping when the sequence stopped. Participants were requested not to subdivide the beat in any way, for example, by covert counting or by moving the body.

*The spontaneous motor tempo task.* The setup was similar to the tapping task. Prior to each trial, participants were instructed to tap a regular rhythm at a tempo that felt comfortable and natural, and that felt neither too fast nor too slow. Participants were told to start tapping when ready and to continue until given notice. Thirty-one taps were recorded before participants were asked to stop.

*The slow motor tempo task.* The setup was similar to the SMT task, the only difference being that participants were asked to tap at their slowest possible rate while still able to maintain a regular beat. Again, participants were asked to refrain from subdividing taps in any way, either overtly or covertly. These instructions conform to those described by McAuley et al. (2006). For each participant, the first fifteen taps were recorded.

### ANALYSIS

Of primary interest to the present study is participants' SR experience of monotonic tone sequences. That is, the perceptual experience is of interest, while differences in how participants approach the task are seen as a confounding variable. The slower limit of SR has been estimated to lie between an ISI of 1500 and 1800 ms (Fraisse, 1982). Any participant who repeatedly reports experiencing a grouping at an ISI well above this limit is assumed to have misinterpreted instructions. This study included four trials, with an ISI of 2000 ms added to detect such participants. Five of the thirty participants reported experiencing a grouping on all four trials at the 2000 ms ISI level. These participants were removed from further data analysis.

For each participant, the mean group period was estimated using only those trials for which a perceived grouping was reported, using the formula:

$$\frac{\sum_{i=1}^{n} T_i \cdot g_i}{n}$$

. . . where $n$ is the number of trials for which the participant reported experiencing a grouping, $g_i$ is the reported group size for the $i$th trial, and $T_i$ is the corresponding ISI. As an example, consider a participant who reports hearing a grouping of four at an ISI of 300 ms, a grouping of two at an ISI of 600 ms, and a grouping of two at an ISI of 900 ms. The mean group period would then be $(4 \times 300 + 2 \times 600 + 2 \times 900)/3 = 1{,}400$ ms.

For the tapping task, the first four taps in every trial were discarded in order to use only those taps where the participants had had some time to synchronize to the sequence. For each trial, tapping variability was calculated as the standard deviation (SD) of the tone-to-tap asynchronies. The increase in timing variability due to slowing of the tempo was estimated by fitting an ordinary least squares regression to the SD of the asynchronies as a function of ISI. The slope of such a regression

[1] A public version is available at http://sumsar.net/files/sr_task/public_sr_task.html

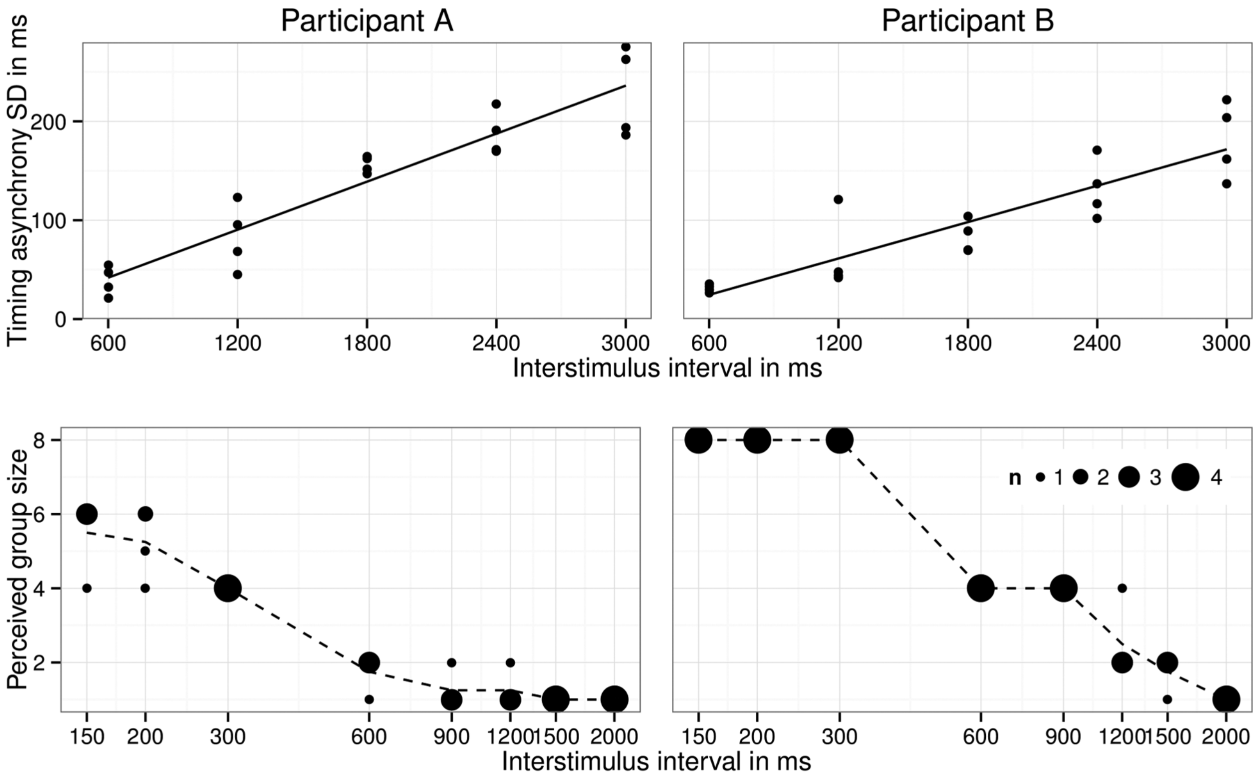


**FIGURE 4.** Timing variability from the tapping task (top row) and reported groupings from the SR task (bottom row) for two participants.

line measures how much worse a participant performs as a result of slowing the tempo; a participant with a small *variability slope* is comparably better at coping with a slow tempo than a participant with a large slope.

Figure 4 shows an example of these measures. Specifically, it shows two participants' reported groupings from the SR task and timing variability from the tapping task. Participant B reported experiencing larger groupings and was better at synchronizing to a slower tempo than participant A, as reflected in the measures of mean group period and variability slope: participant B has a smaller slope and a larger mean group period.

For each participant, the mean spontaneous motor tempo and slow motor tempo were estimated by first calculating the mean intertap interval for each trial, then taking the mean of the three trials for each task. Statistical analysis was performed using the statistical computing environment R (R Core Team, 2012).

## Results

All participants reported hearing groupings when listening to the monotone isochronous sound sequences, despite being told explicitly that the sound sequences were monotone. The most commonly reported groupings were two, four, and eight; three and six were less common; five and seven were rarely reported. Table 1 shows the percentage of responses for each group size and the ISI where that group size was most commonly perceived. A group size of one indicates that the participants reported no grouping.

**TABLE 1.** *Summary of Reported Groupings in the SR Task*

| Grouping | % of Trials | Peak ISI |
|---|---|---|
| 1 | 23.9 | 2000 ms |
| 2 | 29.5 | 600 ms |
| 3 | 3.5 | 300 ms |
| 4 | 29.6 | 200 ms |
| 5 | 0.8 | 150 ms |
| 6 | 2.2 | 200 ms |
| 7 | 0.1 | 150 ms |
| 8 | 10.4 | 150 ms |

The percentage of reported groupings as a function of ISI is shown in Figure 5. Reported group size increases as ISI decreases, both at the group and individual level, i.e., ISI level correlated negatively to reported group size

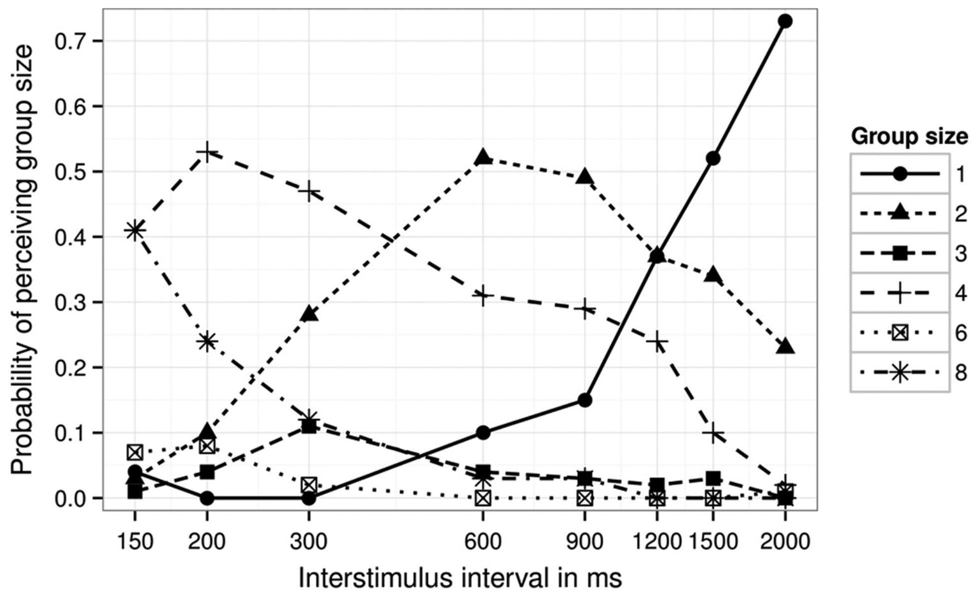


**FIGURE 5.** Percentage of reported groupings as a function of ISI.

for all participants (Spearman's rank correlation with correction for tied values, mean $r_\rho = -.77$, $SD = .15$, $p < .05$ for all participants). For no ISIs did all participants cease to experience a grouping, however, more than half the trials above an ISI of 1500 ms did not result in any experienced groupings.

As a measure of consistency, the probability of reporting the same group size in two different trials with the same ISI was estimated for each participant. Using this measure, participant A in Figure 4 had a consistency of .73, participant B a consistency of .91, and the overall mean consistency was .69 ($SD = .13$). Figure 6 shows the mean consistency across participants at different ISIs. The ISI with the highest consistency was 2000 ms ($M = .75$); the ISI with the lowest consistency was 1200 ms ($M = .65$). Participants were comparably consistent at different ISIs; the standard deviation of the mean consistency across ISIs was 0.03. To put this into perspective, these consistency measures can be compared to those resulting from randomly reporting groupings, according to the group probabilities presented in Table 1. Using this scheme, the consistency is 0.24 (marked by the dashed line in Figure 4), a much lower consistency than any of those calculated using the data.

The mean of the logarithm of reported group sizes was calculated for each participant and ISI level, as shown in Figure 7, where the grand mean is plotted against $\log_2$(ISI). The relationship between reported group size and ISI appears linear, in line with the hypothesis that this relationship would follow a power law. A linear regression between $\log_2$(ISI) and the mean of the logarithm of the reported group sizes for each participant yields the power law relation $g \sim \frac{k}{T^a}$, where estimates of both the factor and exponent are significantly different from zero ($k = 76.7$, $t = 25.1$, $p < .001$; $a = 0.53$, $t = 19.9$, $p < .001$; $R^2 = .67$, $df = 198$).

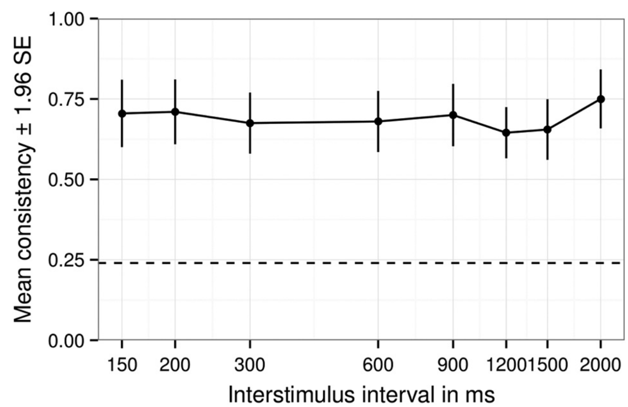


**FIGURE 6.** The mean participant consistency at different ISIs as measured by the probability of reporting the same grouping in two separate trials. The dashed line shows the expected consistency if participants would have reported groupings at random.

The grand mean of the mean group period was 1881 ms ($SD = 656$ ms), the grand mean spontaneous motor tempo was 622 ms ($SD = 157$ ms), and the grand mean

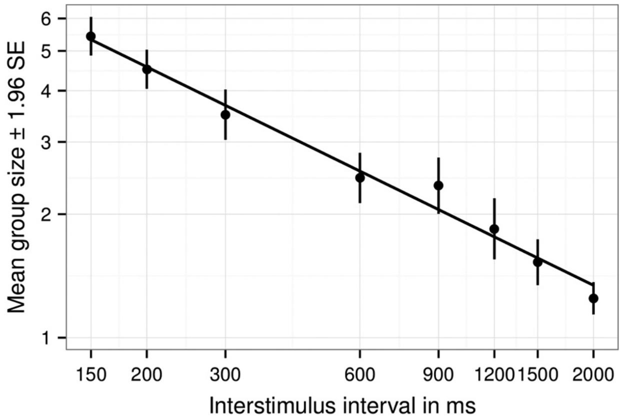


**FIGURE 7.** Log-log plot of mean group size as a function of ISI.

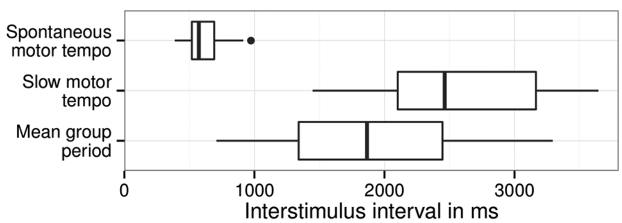


**FIGURE 8.** Distributions of participants' mean spontaneous motor tempo, mean slow motor tempo and mean group period.

slow motor tempo was 2757 ms ($SD$ = 1,100 ms). Figure 8 show the distributions of these three measures. The resonance theory explanation of SR predicted that the mean group period should fall between the slower limit of SR and the slower limit of rhythm perception. The data shown in Figure 8 are in accord with this prediction given that the slower limit of SR is estimated as the ISI where more than half of the trials result in no grouping (ISI 1500 ms) and the slower limit of rhythm perception is estimated by the average slow motor tempo (2757 ms).

Figure 9 shows the relation between a subject's mean group period and variability slope. There was a negative correlation between mean group period and variability slope across participants (Spearman's rank correlation, $r_\rho = -.56$, $p = .0044$, $n = 25$). From a resonance theory perspective, this implies a tendency for participants with a fast slower limit of rhythm perception to have relatively larger timing errors when tapping at slow tempi. There was no significant correlation between years practicing a musical instrument and either mean group period (Spearman's rank correlation with correction for tied values, $r_\rho = -.34$, $p = .09$, $n = 25$) or timing error slope ($r_\rho = -.02$, $p = .91$, $n = 25$).

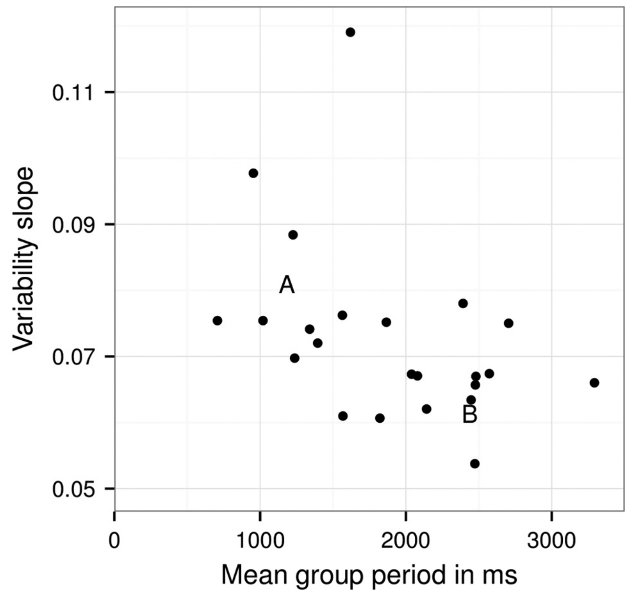


**FIGURE 9.** Variability slope plotted against mean group period for each participant. Participant A and B from Figure 4 are marked by the corresponding letter.

## Discussion

Subjective rhythmization (SR) is the phenomenon whereby the sounds of a monotone metronome sequence are experienced as having different intensity, with the experienced intensity differences following a regular pattern. The present study aimed to replicate and extend the two studies employing the original SR experimental paradigm (Bolton, 1894; Vos, 1973). The extensions were the use of a wider tempo range, the inclusion of multiple trials per tempo level, and the administration of supplemental rhythm production tasks, motivated by two theoretical explanations of SR: the preferred tempo (Temperley, 1963) and the resonance theory (Large, 2008) explanations.

The results confirmed four findings of the earlier studies. First, most participants do report that they experience SR. In the current study all participants reported experiencing SR. While this could be due to the music training of many of the participants, it supports the position of SR as a robust phenomenon that a large part of the population experiences.

Second, the experience of SR is strongly affected by the tempo of the sound sequence, as shown by a strong negative correlation between sound sequence ISI and reported group size. Participants were highly consistent with regard to the group sizes reported at particular ISIs; the probability of reporting the same group size on any two trials with the same ISI averaged .69. Putting this into perspective, the probability of choosing the same response on two different trials would be 0.24 if choosing randomly according to the group probabilities in Table 1. Participants were also comparably consistent across ISIs, that is, although the impression of SR is strongly affected by tempo, consistency of responses is not.

Third, all group sizes are not reported with equal frequency. Groups of two, four, and eight were reported most often, followed by groups of three and six. Groups of five and seven were reported on less than 1% of the trials. This is the ordering one would expect from a Western music-theoretical perspective (van Noorden & Moelants, 1999). To date, no SR study has been conducted in a country with a non-Western musical tradition. It remains to be determined to what degree SR is affected by cultural factors. As culture has been shown to play an important role in rhythm perception (Hannon, Soley, & Ullal, 2012), a prediction is that groups of five and seven would be more commonly reported by participants accustomed to odd meters prevalent in, e.g., the traditional music of the Balkan Peninsula.

Fourth, when the tempo of the sequences is sufficiently slow, participants do not experience SR. This slower limit of SR, while not probed by Vos (1973), was estimated to an ISI of 1500 ms by Bolton (1894). The current study found no such sharp limit but instead found large inter-individual variability. However, at an ISI of 1500 ms more than half the trials resulted in no experienced SR, comparable to Bolton's figure.

The current study focused on how the experience of SR varies as a function of tempo but many other factors might also be influential. Time perception differs depending on the pitch of the stimulus (Hove, Marie, Bruce, & Trainor, 2014), so it is possible that pitch affects the experience of SR. Another factor that is likely to influence SR is the task instructions, even though the comparability of the results from the current study with the previous studies by Bolton (1894) and Vos (1973) shows that SR is at least somewhat robust to variation in the task instructions. That said, differences in how subjects approach the task might still heavily influence the experience of SR. The study by Nozaradan et al. (2011) is already an example of this, as whether participants were asked to actively imagine a subjective duple meter or not influenced their subsequent EEG readings. It remains an open question to what degree, and in what way, the experience of SR depends on the task instructions and on qualities of the stimulus such as amplitude, pitch, and timbre.

As an aside, my experience is that perceived groupings can be changed somewhat at will, for example, listening to a monotone sequence with an ISI of 600 ms I often start out hearing an accent on every second sound. By focusing, however, I can switch the accent to every fourth sound. If SR can generally be affected by such top-down control it would not imply that SR is a purely top-down phenomena. Rather, such a finding would resonate with research regarding visual illusions, such as the Necker cube, known to be affected by both bottom-up and top-down processes (Long & Toppino, 2004).

### EXPLANATIONS OF SUBJECTIVE RHYTHMIZATION

The literature offers two explanations for SR: the preferred tempo (Temperley, 1963) and the resonance theory (Large, 2008). Resonance theory is a dynamical systems framework for modeling rhythm perception and production. The resonance theory explanation of SR is based on the notion that an isochronous sequence, in addition to entraining oscillatory units responsive to the fundamental period, entrains subharmonic oscillators, thus producing the subjective accents characteristic of SR (Large, 2008). This explanation of SR gives rise to three predictions: (1) the mean group period of the reported groupings should fall between the slower limit of SR and the slower limit of rhythm production, (2) the relation between the size of the reported grouping and ISI of the sound sequence should follow a power relation, and, (3) a participant's mean group period should relate to tapping performance at slow tempi. Within the resonance theory framework, (1) follows from assuming a slower limit of rhythm perception, with the mean group period being seen as a proxy variable for this limit; (2) follows from a slower limit of rhythm perception limiting the highest possible grouping that can be perceived for any given ISI; (3) follows from the assumption that rhythm perception and rhythm production both share the same underlying mechanism.

The results of the present study are in line with the predictions developed on the basis of resonance theory. The results do not support the preferred tempo explanation, whereby the mean group period should be close to participants' spontaneous motor tempo. Instead, the mean group period was closer to the participants' slow motor tempo (see Figure 8), in line with prediction (1).

Resonance theory assumes that rhythm perception and rhythm production share a common neural substrate. Thus, there should be a relation between a participant's performance in rhythm perception tasks and rhythm production tasks. The present study did indeed find such a relation as there was a correlation between what a participant reported in the SR task and her timing performance in the tapping task. Specifically, participants that reported large groupings in the SR task tended to have smaller timing variability when tapping at a slow tempo relative to tapping at a moderate tempo. From a resonance theory perspective this is explained by that the slower limit of rhythm perception influences both timing variability at a slow tempo and what grouping is perceived in an SR task.

The relation between the reported group sizes and the ISI of the sound sequences was found to follow a power relation closely (see Figure 7). Resonance theory explains this by that of the group size perceived at a certain ISI depends on the participant's slower limit of rhythm perception. That slower limit governs the ISI at which the participant starts to experience a given group size. The relation between group size and ISI was well captured by the expression $g \sim \frac{k}{T^a}$, where $g$ is the perceived grouping, $T$ is the ISI of the sequence, and $k$ and $a$ are constants. The results are not compatible with a sharp slower limit of rhythm perception. A sharp limit would imply that a participant should experience a grouping of two at half the ISI of the slower limit, a grouping of four at a fourth of the slower limit, etc. Such behavior would result in $a = 1$, with $k$ equal to the slower limit. The estimate of the current study was $a = 0.53$ implying that participants tend to report smaller group sizes at faster tempi compared to what a sharp limit predicts. This can be accommodated within a resonance theory framework by treating rhythm perception as an ability that, instead of having a sharp limit, deteriorates gradually as the tempo gets slower.

Overall, the current results are well explained by the resonance theory of rhythm perception. This is not to say that other models could not explain the phenomena of SR. However, the current results do suggest that any such account would need to include both a slow limit of rhythm perception and a close connection between rhythm perception and rhythm production. Subjective rhythmization is closely related to meter perception; the ability of subjects to experience widely different accent patterns while listening to the same sequences draws attention to the difference between a rhythm sequence as stimulus and as percept. Of course, it is not uncommon that different people experience the same piece of music differently. What is perhaps surprising is that, even while listening to the most simple monotone metronome sequence, what is experienced is *still* in the ear and mind of the listener.

## Author Note

This research was supported by the Linnaeus environment Thinking in Time: Cognition, Communication and Learning, Swedish Research Council grant number 349-2007-8695. I thank Geoffrey R. Patching, Joel Parthemore and Christian Balkenius for useful comments on various aspects of this work.

*Correspondence concerning this article should be addressed to* Rasmus Bååth, Cognitive Science, LUX Lund University, Box 192, 221 00 Lund, Sweden. E-mail: rasmus.baath@lucs.lu.se

## References

Andrews, B. (1905). Auditory tests. *The American Journal of Psychology*, *16*, 302-326.

Averbeck, B. B., Latham, P. E., & Pouget, A. (2006). Neural correlations, population coding and computation. *Nature reviews. Neuroscience*, *7*(5), 358-366. doi: 10.1038/nrn1888

Bååth, R. (2011). *Construction of a low latency tapping board* (LUCS Minor Report 17). Retrieved from the website of Lund University Cognitive Science: http://www.lucs.lu.se/LUCS/M017/rasmus_baath_2011_construction_of_an_arduino_based_tapping_board.pdf

Bolton, T. (1894). Rhythm. *The American Journal of Psychology*, *6*(2), 145-238.

Brochard, R., Abecasis, D., Potter, D., Ragot, R., & Drake, C. (2003). The "ticktock" of our internal clock: Direct brain evidence of subjective accents in isochronous sequences. *Psychological Science*, *14*(4), 362-366. doi: 10.1111/1467-9280.24441

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

Hannon, E. E., Soley, G., & Ullal, S. (2012). Familiarity overrides complexity in rhythm perception: A cross-cultural comparison of American and Turkish listeners. *Journal of Experimental Psychology. Human Perception and Performance*, *38*(3), 543–548. doi: 10.1037/a0027225

Hove, M. J., Marie, C., Bruce, I. C., & Trainor, L. J. (2014). Superior time perception for lower musical pitch explains why bass-ranged instruments lay down musical rhythms. *Proceedings of the National Academy of Sciences*, *111*, 10383-10388.

Kirnberger, J. P. (1776). *Die Kunst des reinen Satzes in der Musik: aus sicheren Grundsätzen hergeleitet und mit deutlichen Beyspielen erläutert* [The art of pure movement in music: Derived from sound principles and explained with clear examples]. Berlin: Decker and Hartung.

LARGE, E. (2000). On synchronizing movements to music. *Human Movement Science*, *19*(4), 527-566. doi: 10.1016/S0167-9457(00)00026-9

LARGE, E. (2008). Resonating to musical rhythm: Theory and experiment. In S. Grondin (Ed.), *The psychology of time* (pp. 189-231). Bingley: Emerald Group Publishing. doi: 10.1016/B978-0-08046-977-5.00006-5

LARGE, E., & JONES, M. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, *106*(1), 119-159. doi: 10.1037/0033-295X.106.1.119

LARGE, E., & KELSO, J. A. S. (2002). Tracking simple and complex sequences. *Psychological Research*, 3-17. doi: 10.1007/s004260100069

LARGE, E., & KOLEN, J. (1994). Resonance and the perception of musical meter. *Connection Science*, *6*(2), 177-208. doi: 10.1080/09540099408915723

LONG, G. M., & TOPPINO, T. C. (2004, September). Enduring interest in perceptual ambiguity: alternating views of reversible figures. *Psychological Bulletin*, *130*(5), 748-68. doi: 10.1037/0033-2909.130.5.748

MCAULEY, J. D., JONES, M. R., HOLUB, S., JOHNSTON, H. M., & MILLER, N. S. (2006). The time of our lives: Life span development of timing and event tracking. *Journal of Experimental Psychology: General*, *135*(3), 348-367. doi: 10.1037/0096-3445.135.3.348

MOELANTS, D. (2002). Preferred tempo reconsidered. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, & J. Renwick (Eds.), *Proceedings of the Seventh International Conference on Music Perception and Cognition* (pp. 572-574). Sydney: Causal Productions.

NOZARADAN, S., PERETZ, I., MISSAL, M., & MOURAUX, A. (2011). Tagging the neuronal entrainment to beat and meter. *The Journal of Neuroscience*, *31*(28), 10234-10240. doi: 10.1523/JNEUROSCI.0411-11.2011

PATEL, A. (2008). *Music, language, and the brain*. New York: Oxford University Press.

R CORE TEAM. (2012). *R: A language and environment for statistical computing* [Computer software manual]. Vienna, Austria: R Foundation for Statistical Computing.

REPP, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin and Review*, *12*(6), 969-992. doi: 10.3758/BF03206433

REPP, B. H. (2006). Rate limits of sensorimotor synchronization. *Advances in Cognitive Psychology*, *2*(2), 163-181.

SCHAEFER, R. S., VLEK, R. J., & DESAIN, P. (2011). Decomposing rhythm processing: Electroencephalography of perceived and self-imposed rhythmic patterns. *Psychological Research*, *75*(2), 95-106. doi: 10.1007/s00426-010-0293-4

SCHEIRER, E. D. (1998). Tempo and beat analysis of acoustic musical signals. *Journal of the Acoustical Society of America*, *103*, 588-601. doi: 10.1121/1.421129

TEMPERLEY, N. (1963). Personal tempo and subjective accentuation. *Journal of General Psychology*, *68*(2), 267-287. doi: 10.1080/00221309.1963.9920534

VAN NOORDEN, L., & MOELANTS, D. (1999). Resonance in the perception of musical pulse. *Journal of New Music Research*, *28*(1), 43-66. doi: 10.1076/jnmr.28.1.43.3122

VLEK, R. J., SCHAEFER, R. S., GIELEN, C. C. A. M., FARQUHAR, J. D. R., & DESAIN, P. (2011). Sequenced subjective accents for brain-computer interfaces. *Journal of Neural Engineering*, *8*(3), 1-12. doi: 10.1088/1741-2560/8/3/036002

VOS, P. (1973). *Waarneming van metrische toonreeksen* [The perception of metric tone sequences]. Nijmegen: Stichting Studentenpers.