The role of executive control in rhythmic timing at different tempi

Rasmus Bååth\textsuperscript{1,2} · Trond Arild Tjøstheim\textsuperscript{1} · Martin Lingonblad\textsuperscript{1}

Like visual perception can be divided into such subcategories as color perception, motion perception, and depth perception, so too can time perception. Some aspects of time perception are interval timing, temporal motor coordination, rhythm perception, and meter perception. It is possible to further subdivide time perception by modality and time scale. Much debated is whether all aspects of time perception share a common mechanism and, if not, what aspects of which mechanisms they do share (for a review see Ivry & Schlerf 2008). One influential class of models assumes that timing is governed by a pacemaker-accumulator type mechanism (Ivry, 1996), while more recent theoretical development are dynamical systems models that assume that timing and rhythm perception depend on oscillatory neural circuits (Large & Jones, 1999; Large, 2010). The former have been used successfully to model interval timing but has not proven a good model of responses to more complex stimuli such as musical rhythms, while the latter have been used to model rhythm and meter perception but have not been applied to interval timing (Grondin, 2010). The two mechanisms – pacemaker-accumulator type and oscillatory based – need not stand in opposition; models exist that incorporate both (Teki et al., 2012).

Some have suggested that time perception relies on different mechanisms, depending on time scale. Lewis and Miall (2006) report evidence that different neural mechanism are responsible for timing intervals shorter versus longer than one second. The timing of sub-second intervals has been termed \textit{automatic timing} and that of supra-second intervals has been termed \textit{cognitive timing}. These terms reflect that automatic timing recruits circuits within the motor system and auditory cortex, while cognitive timing depends more on circuits within the prefrontal and parietal cortices (Lewis & Miall, 2003). Interval timing is but one
aspect of time perception, and a three second window has been suggested as the limit of temporal integration (Pöppel, 2004; Mates et al., 1994). For rhythmic timing, for example, synchronizing finger taps to a metronome sequence, there is evidence supporting a shift in mechanisms between a one- and a two second interstimulus interval (ISI), i.e., the time interval between each beat in a rhythmic sequence, where a short ISI implies a fast tempo and vice versa. The Weber fraction – a measure of relative timing error – increases markedly from synchronizing to a one-second ISI to synchronizing to a two-second ISI (Grondin, 2012) and so does the perceived difficulty of synchronizing (Bååth & Madison, 2012). A related notion is the slower limit of rhythm perception, suggested to lie between a 1.5 second and a 3 second ISI (Repp, 2006).

Brown (1997) hypothesizes that the mechanism responsible for rhythmic timing above a two second ISI requires more attentional and executive resources, support for which has been presented by Miyake et al. (2004). They showed that participants’ ability to synchronize to a metronome sequence while simultaneously performing a memory task is more impaired at ISIs above two seconds compared to shorter ISIs. However, a similar study by Holm et al. (2013) showed no evidence of any interaction effect between the ISIs of the sequences and whether participants performed an executive function distractor task or not. Both studies used a dual-task setup: a standard experimental paradigm that aims to discern whether two tasks depend on the same limited cognitive capacity, such as executive control or short term memory (Pashler, 1994). One of the tasks – sometimes referred to as the main task – is the task under study; the second task – here called the distractor task – is assumed a priori to tax a certain cognitive capacity. Participants are either asked to perform the main task and the distractor task simultaneously or to perform solely the main task. Performance between the two conditions is then compared. If the distractor task interferes with performance in the main task, this is taken to indicate that both tasks rely – at least to some extent – on the same limited cognitive capacity.

Our study investigated whether rhythmic timing requires more attentional resources when the tempo is slow compared to when it is fast, where a fast tempo is loosely defined as an ISI shorter than 1500 ms and a slow tempo as an ISI longer than 1500 ms. In keeping with Miyake et al. (2004) and Holm et al. (2013), we used a dual-task paradigm, with a rhythmic timing task as the main task and a distractor task selected to require attentional resources and executive control.

More specifically, the main task was a sensorimotor synchronization task in where participants were asked to tap their index finger in time with metronome sequences. The tempo of the sequences included ISIs of 600 ms to 3000 ms. The distractor task was a novel variation on the n-back task. The n-back task was chosen because it is commonly used to assess executive function (Baddeley, 2003) and because its design facilitates straightforward varying of attentional resource and executive control demands (Smith and Jonides, 1999; Chatham et al., 2011). In line with the capacity sharing explanation of dual task interference (Pashler, 1994), we hypothesize that simultaneously performing the timing task and the n-back task would require participants to share limited cognitive resources between these two task resulting in degraded performance in both tasks.

The difficulty with using the standard n-back task in a dual-task setup is that it requires participants to make responses throughout the task, either verbally or by key press. These motor responses might well interfere with the motor responses in the sensorimotor synchronization task, making it difficult to infer whether any task interference is due to attentional interference or motor interference. Therefore, a novel variant of the n-back task was used, here called the covert n-back task, where the participant makes no overt responses during the task.

If the distractor task should be found to impair rhythm timing more at slow tempi than at fast tempi then this would accord with models that assume different timing mechanisms being recruited depending on time scale (Lewis & Miall, 2003). It would also accord with models that assume a dedicated rhythm-perception mechanism and a slower limit for rhythm perception, for example, Large’s (2008) proposed resonance model of rhythm perception.

Method

Participants

Twenty-four participants were recruited via public advertising (11 women and 13 men, mean age: 27 years, SD: 6 years). Seventeen participants reported having experience playing a musical instrument and the mean reported number of years of regular practice was 13 (SD: 10).

Material

The main task was a sensorimotor synchronization task. A covert response n-back task was used as distractor task.

Sensorimotor synchronization task

Participants were asked to synchronize finger taps to isochronous metronome sequences. They were to start as soon as a sequence started and continue until the sequence ended. They were requested not to subdivide the beat in any way, for example, by covert counting or by moving their body. While it is not possible to guarantee that participants...
completely refrain from subdividing the beat, this instruction was given with the objective to keep subdivision to a minimum. A custom-built tapping board consisting of a piezoelectric sensor mounted on 5cm² corrugated fiberboard recorded the timing of the finger taps (see Bååth 2011 for details). Participants tapped with their index finger, their hand resting on a plastic foam cushion. The stimuli consisted of isochronous sequences of 440 Hz square wave tones of 20 ms, where each sequence was 45 seconds long. Sequences were presented at five tempi with ISIs of 600, 897, 1342, 2006, and 3000 ms, selected so as to be equidistant on a log scale. An Arduino microcontroller generated the sounds and registered the taps.

Covert response N-back task

Participants were asked to perform a visuospatial 2-back task. The visual stimuli was modeled after Jaeggi et al. (2007). It consisted of a white 3 × 3 grid on a black background, with a white fixation cross in the middle and a blue square in one of eight outer grid positions (see Fig. 1). The blue square changed position every 2150 ms, including 700 ms to fade in and 700 ms to fade out. These time intervals were chosen so that the presentation of the blue square would not regularly coincide with stimuli in the sensorimotor synchronization task.

A given stimulus presentation constituted a target if the blue square’s current position was the same as two positions back. The square’s position was randomized so that, on average, half the presentations were targets. Instead of responding overtly to each target, participants were instructed to count the number of targets silently and report the total at the conclusion of each trial. This variation of the n-back task was used as the the responses during the 2-back task could otherwise interfere with the motor part of the sensorimotor synchronization task. The 2-back difficulty level of the n-back task was used, rather than a 1-back or 3-back level, as the 2-back level was found to be subjectively difficult, but still possible to carry out, for all persons it was tested on during the development of the task. Furthermore, the 2-back is a common difficulty level of the n-back task in, for example, neuroimaging studies of working memory (Owen et al., 2005).

Trials were 47 seconds long: slightly longer than the sensorimotor synchronization task trials. The distractor task was implemented in the Java programming language using the Processing framework (Reas and Fry, 2007).

Procedure

Participants were tested individually in a quiet room. Sessions began with a number of practice trials. First a sensorimotor synchronization trial at 600ms ISI, then an n-back only trial, and finally a trial where the two tasks were presented simultaneously. After this the participant was given four n-back-only-trials to establish baseline performance. For the sensorimotor-synchronization task, audio was delivered through a pair of closed headphones. The n-back distractor task used a 27” monitor positioned 50 cm from the participant.

The experiment proper consisted of four blocks of five sensorimotor synchronization trials, one for each of the five ISI levels. The order of the trials within each block was randomized. Either the first and third or the second and fourth blocks included the n-back distractor task and whether or not a participant started with a distractor block was also randomized. Each participant performed 20 trials, four at each ISI level, where two included the distractor task and two were without the distractor task.

Analysis

The first three taps in every sensorimotor-synchronization trial were discarded to use only those taps where participants had time to synchronize to the sequence. For each tap, tone-to-tap asynchrony was calculated as the time difference between the tone and the tap, a negative asynchrony indicating that the tap preceded the tone and vice versa. Asynchrony SD was taken as a measure of timing variability and it was estimated for each participant and ISI level using the Bayesian hierarchical method described in Bååth (2015). Timing variability is here used as a measure of performance in the sensorimotor synchronization task with low variability taken to indicate high performance.

The Bayesian method was used, instead of the conventional sample SD, to mitigate the influence of taps that resulted from the participant reacting to the tones, rather than anticipating them. For ISIs shorter than 1500 ms participants tend to produce few reactive taps (Repp and Doggett, 2007) and the Bayesian estimates will be highly similar to sample SD estimates. At longer ISIs participants tend to
produce more reactive taps and the sample SD will underestimate the timing variability (Bååth, 2015). The Bayesian method corrects for this by discarding the information from the reactive taps and produces an estimate of asynchrony SD using only the information from the anticipatory taps. If no correction was made for reactive taps then the estimated variability at slow ISIs would be a combination of timing variability and auditory reaction time variability, and the rational for doing the correction is that the focus here is solely on timing performance.

As a second measure of timing performance we used the coefficient of variation, calculated for each participant and condition as the asynchrony SD divided by the ISI. The coefficient of variation is a measure of timing performance relative to the ISI. As such it is useful when comparing performance between different ISI levels, while the asynchrony SD is useful as an absolute measure of performance.

Statistical analysis was performed using the statistical computing environment R (R Core Team, 2012). Because timing variability was measured at five different ISIs for each participant, a linear mixed-effects model was used to assess how timing variability changed as a function of ISI and distractor condition. Mixed-effects model analyses were performed using the package lme4 (Bates et al., 2014).

### Results

The dependence of timing variability on ISI and distractor condition – control or n-back – was investigated by fitting a linear mixed-effects model, using loge asynchrony SD as the outcome variable and ISI, distractor condition, and the interaction between ISI and distractor condition as the predictor variables. The ISI was standardized prior to fitting the model and the asynchrony SD was loge transformed, as it was found to have a right skewed distribution. Table 1 reports the resulting parameter estimates. Figure 2 shows loge asynchrony SD as a function of ISI, with superimposed regression lines from the mixed-effects model.

![Fig. 2 Mean timing variability as measured by asynchrony SD for all participants and ISI levels. The regression lines show the results of the mixed-effects model analysis](image)

Table 1 Results of the mixed-effects model analysis of timing variability

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Estimate</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>4.25</td>
<td>[4.20, 4.36]</td>
<td>-</td>
</tr>
<tr>
<td>ISI</td>
<td>0.74</td>
<td>[0.69, 0.78]</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Distractor condition</td>
<td>0.35</td>
<td>[0.29, 0.40]</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>ISI  x  distractor condition</td>
<td>0.11</td>
<td>[0.060, 0.17]</td>
<td>-.001</td>
</tr>
</tbody>
</table>

The effect of both ISI and distractor condition on asynchrony SD was statistically significant, as was the interaction effect, where the difference between the control and the n-back condition increased with longer ISI. For example, the mean difference in loge asynchrony SD between the control and the n-back condition was more than three times as large at the 3000 ms ISI level compared to the 600 ms ISI level. This interaction effect can also be seen when looking at the difference between each participant’s asynchrony SD under the two conditions. Figure 3 shows how the difference increases as a function of ISI; a positive difference means that the timing variability was higher in the n-back than in the control condition. In this and all subsequent figures, error bars show 95% confidence intervals (CI) calculated as 1.96 × standard error.

The effect of the distractor condition can be seen in other measures of timing performance. Figure 4 shows the mean coefficient of variation as a function of ISI and distractor condition; the difference between the two distractor conditions increases with longer ISIs. Another measure of timing performance is the percentage of reactive responses (Miyake et al., 2004; Repp & Doggett, 2007), defined as the percentage of responses that overshot the target tones by more than 100 ms. Figure 5 shows very few reactive responses at 600 and 897 ms ISIs. For longer ISIs, the...
percentage of reactive responses was greater in the n-back condition.

Timing performance decreased under the n-back condition and so did performance in the n-back task at slower tempi. Baseline n-back performance was calculated for each participant as the mean number of errors made in the four n-back only trials. The difference between each participant’s baseline performance and performance during the experiment proper was then calculated for each ISI level. Figure 6 show the mean n-back error compared to baseline. The difference was statistically significantly different from zero at the 3000 ms ISI level (one sample t-test, M = 1.8, t(22) = 5.2, p < 0.001). For shorter ISIs, average n-back performance was less than one error above baseline.

**Discussion**

Many models of human timing and time perception have been proposed. One important way in which they differ is whether they posit a single, overreaching mechanism for timing or assume that timing recruits different mechanisms depending on the nature of the task. Regarding rhythmic timing, it has been proposed that different mechanisms are responsible depending on the tempo (Grondin, 2012). Relevant here is the notion of a slower limit of rhythm perception, a proposed temporal boundary where perceiving and synchronizing to a rhythmic sequence goes from being effortless and automatic to requiring attention and executive control (Repp, 2006). The present study used a dual-task
setup to investigate whether rhythmic timing requires more attentional resources at slow tempi compared to comfortable tempi. The main task was a sensorimotor synchronization task where participants tapped their finger in time with metronome sequences and the distractor task was a covert response n-back task.

The results point towards rhythmic timing requiring more attentional resources at slow tempi. At the slowest tempo – at an interstimulus interval (ISI) of 3000 ms – performance of the tapping task and n-back task simultaneously resulted in a significant performance degradation in both tasks. It is difficult to identify a particular tempo at which dual-task interference becomes significant. Looking at the different performance measures, the largest increase in interference occurs between an ISI of 897 and 1342 ms for the log asynchrony SD, and between an ISI of 2006 and 3000 ms for the coefficient of variation, percentage of reactive responses, and number of errors in the n-back task. The results reflect the authors’ own experience when piloting the experiment and participants’ informal verbal reports: keeping the beat with a fast metronome while doing a 2-back task is easy; keeping the beat with a metronome that strikes every third second while doing a 2-back task is hard.

The fastest tempo – at an ISI of 600 ms – also saw dual-task interference, however, the magnitude of interference was much lower in comparison to the slower ISI levels. Therefore, while the result of this study show that rhythmic timing requires more attentional resources at slow tempi, it does not support the notion that executive function resources are not employed to some degree also at short ISIs. This accords with studies of non-rhythmic timing where timing performance has been shown to correlate with measures of working memory capacity and intelligence (Engle et al., 1999; Broadway & Engle, 2011). Professional musicians have also been shown to make use of more attentional resources when involved in rhythmic timing compared to non-musicians (Fischinger, 2011). It is, therefore, possible that the dual-task interference at the fast ISI levels would be more pronounced would this study have employed professional musicians as participants rather than non-musicians. However, if all trials where participant’s n-back performance was worse than their baseline performance are removed, the statistical result presented in Table 1 and Fig. 2 still remains. That is, the effect of ISI, distractor condition and their interaction on asynchrony SD is still statistically significant with \( p < 0.001 \).

The results of the present study are consistent with those from the study by Miyake et al. (2004), who asked participants to perform a word-memory task and rhythmic tapping task. While Miyake et al. did not analyze timing variability, they found that participants produced more reactive responses when both tasks were performed simultaneously. As with the present study, the difference was not found at shorter ISIs but became pronounced at 1800 ms ISI.

The results are not consistent with a recent study by Holm et al. (2013), who asked participants to perform a rhythmic timing task under either a low or high cognitive load condition. They did not find an effect of cognitive load on timing performance, nor did they find an interaction between cognitive load and sequence tempo. The results may be due to the distractor task used. Under the low cognitive load condition, participants were asked to tap the rhythm in a random sequence. A possible reason for why no task interference was observed when participants synchronized at a slow tempo is because the distractor task is easier to perform at a slower compared to a faster tempo, i.e., the distractor task is not invariant to the sequence tempo. At 1000 ms a participant must make twice as many random decisions as at an ISI of 2000 ms. The cognitive load resulting from the timing task might indeed have been heavier at the slower tempi, but no interference effect was manifest, because the cognitive load resulting from the distractor task was lighter at the slower tempi.

In conclusion, the present study shows that, when the tempo is sufficiently slow, performing rhythmic timing demands attentional resources and involvement of executive control. These results accord with neural models of timing that suggest a dedicated, automatic timing mechanism for short intervals and a general, cognitive timing mechanism for longer intervals (Lewis and Miall, 2003). The results might also be explained though by a single timing mechanism that requires more cognitive resources at slower tempi. As shown in this study, rhythmic timing requires more cognitive resources the slower the tempo, and both attentional resources and executive control are presumably limited. Therefore, independent of whether rhythmic timing depends on one or several mechanisms, this study supports the view that rhythm perception and rhythmic timing have a slower limit.

Author Note  Rasmus Bååth, Lund University Cognitive Science, Department of Philosophy, Lund University, Sweden. This research was supported by the Linnaeus environment Thinking in Time: Cognition, Communication and Learning, Swedish Research Council grant number 349-2007-8695. T. Tjøstheim and M. Lingenblad contributed equally to this work. The authors would like to thank Frida Nelnans, David Sivén, Ryo Yamazaki, and Henrik Siljebråt for their help with the data collection, and Christian Balkenius and Joel Parthemore for helpful comments on this manuscript. Correspondence concerning this article should be addressed to Rasmus Bååth, Department of Philosophy, LUX, Lund University, Box 192, 221 00 Lund, Sweden. E-mail: rasmus.baath@gmail.com
References


