For The Last Time: Temporal Sensitivity and Perceived Timing of the Final Stimulus in an Isochronous Sequence

Min Susan Li[1], Darren Rhodes[1], and Massimiliano Di Luca[1*]

[1]Centre for Computational Neuroscience and Cognitive Robotics, School of Psychology, University of Birmingham.

Short Title: Temporal judgments in regular sequences

Word Count: 7508 words

*Corresponding Author: m.diluca@bham.ac.uk

Centre for Computational Neuroscience and Cognitive Robotics, School of Psychology, University of Birmingham, Edgbaston, Birmingham, UK, B152TT.

Phone: +44 121 414 5526

Abstract

An isochronous sequence is a series of repeating events with the same inter-onset-interval. A common finding, is that as a the length of a sequence increases, so does temporal sensitivity to irregularities – that is, the detection of deviations from isochrony is better with a longer sequence. Several theoretical accounts exist in the literature as to how the brain processes sequences for the detection of irregularities, yet there remains to be a systematic comparison of the predictions that such accounts make. To compare the predictions of these accounts, we asked participants to report whether the last stimulus of a regularly-timed sequence appeared 'earlier' or 'later' than expected. Such task allowed us to separately analyse bias and performance. Sequences lengths (3, 4, 5 or 6 beeps) were either randomly interleaved or presented in separate blocks. We replicate previous findings showing that temporal sensitivity increases with longer sequence in the interleaved condition but not in the blocked condition (where performance is higher overall). Results also indicate that there is a consistent bias in reporting whether the last stimulus is isochronous (irrespectively of how many stimuli the sequence is composed of). Such result is consistent with a perceptual acceleration of stimuli embedded in isochronous sequences. From the comparison of the models' predictions we determine that the improvement in sensitivity is best captured by an averaging of successive estimates, but with an element that limits performance improvement below statistical optimality. None of the models considered, however, provides an exhaustive explanation for the pattern of results found.

1. Introduction

Psychological time is subject to several types of distortions (e.g., Allan, 1979). For instance, temporal structure (Horr & Di Luca, 2015), violations of regularity (Pariyadath & Eagleman, 2007; Rose & Summers, 1995), and musical context (Pecenka & Keller, 2011) can all influence the perceived duration of events. Here, we investigate the effect of temporal regularity on time perception. The simplest form of regularity in time is created by an isochronous sequence, that is, the repetition of identical stimuli after equal temporal intervals. Isochronous sequences create temporal expectations based on their regular rhythm and repeated pattern (Arnal & Giraud, 2012; Large & Jones, 1999) and can influence perceptual judgments and behaviour (Brochard et al., 2013; Coull, 2009; Cravo et al., 2013; Escoffier et al., 2010; ten Oever et al., 2014). The sensitivity of judgments about the temporal properties of sequences is also improved by temporal regularities (Drake & Botte, 1993; Grondin, 2001; Hirsch et al., 1990; McAuley & Kidd, 1998).

The aim of this paper is twofold: first, we analyse existing models of how the brain deals with detecting temporal deviations in isochronous sequences (sequences of stimuli spaced by identical intervals). To do this, we utilize stimuli and conditions taken from previous investigations (Halpern & Darwin, 1982; Hoopen et al., 2011; Schulze, 1978; 1989) whereby observers are presented a sequence of isochronous tones except for the last interval. In concert with the methodology of Halpern and Darwin (1982) and ten Hoopen et al. (2011), the last interval could be shorter or longer than expected, whereas in Schulze's (1989) study the last interval could only be equal or longer than the preceding intervals. Using such a methodology allows us to measure the temporal sensitivity to temporal deviations as well as finding the point at which participants subjectively report a single stimulus was isochronous. As such,

26   the second aim of the paper is to see if there is a distortion from veridical perception –

27   that is – if isochronous stimuli in a sequence are perceived as being on time, or

28   whether they are perceptually accelerated, or delayed. The existing accounts of

29   temporal sensitivity in isochronous sequences can only account for this type of

30   changes in perceived isochrony by appealing to a response bias (an imbalance in the

31   probability of the two responses), which has no perceptual origin. Such a finding

32   would open the road to models that are able to capture biases in perceived timing of

33   stimuli in isochronous sequences.

34   **1.1 Percept Averaging (PA) Model Description**

35   Schulze (1989) proposed to frame the problem of detecting whether the final duration

36   in a sequence of intervals is deviant as discrimination between the duration of the Nth

37   interval from the average of the percept of the previous *N-1* intervals. Here we term

38   this approach Percept-Averaging (PA) model, which assumes that all intervals are

39   stored in memory and the perceptual system is capable of averaging them in a

40   statistically optimal fashion, thus increasing the precision of the average (Schulze,

41   1989).

42   First of all, we will consider a simple case, where all *N* intervals in the

43   sequence are independently estimated. If each estimate of the duration of an interval *E*

44   is affected by independent Gaussian noise with average *μ=0* and variance *σ²*, then the

45   average of *N-1* estimates has variance equal to $V\left(\frac{1}{N-1}\sum_{i=1}^{N-1}E_i\right) = \frac{(N-1)\,\sigma^2}{(N-1)^2} = \frac{\sigma^2}{(N-1)}$.

46   The predicted just-noticeable difference (*JND'*) with a sequence of *N* intervals of

47   which the last could be deviant is expressed by $JND_N' = \sqrt{\frac{\sigma^2}{(N-1)} + \sigma^2} = \sqrt{\frac{N\sigma^2}{(N-1)}}$.

48   Using this formula we find that the *JND'* predicted with a sequence of 2 intervals is

49    $JND_2' = \sqrt{2}\sigma$. We can then express the predicted JND_N' of a sequence with *N*

5

50   intervals where the change in tempo happens at the last interval as a function of the

51   empirical JND_2 of a sequence with 2 stimuli by integrating the two formulas as such:

52   $$JND_N' = JND_2 \sqrt{\frac{N}{2(N-1)}}$$ Eq. (1).

53   The pattern generated by this formula is shown in Figure 1.

54   The results of Schulze (1989) suggest that the improvement in performance

55   with interleaved presentation of different sequence durations in a block is *higher* than

56   the one predicted by this formula. Schulze speculated about the possibility that

57   participants learned the duration of intervals throughout the experiments rather than

58   within a single sequence. He also investigated whether this discrepancy could be due

59   to the correlation in the noise of the duration estimated of successive intervals. A

60   correlation in this instance means that an error made on the estimate of one interval

61   influences also the estimates of the neighbouring ones. With coefficient of correlation

62   *r* between successive intervals (and 0 otherwise) the average of *N-1* estimates has

63   variance equal to $V\left(\frac{1}{N-1}\sum_{i=1}^{N-1} E_i\right) = \frac{\sigma^2}{(N-1)^2} + \frac{2r(N-2)\sigma^2}{(N-1)^2}$. The *JND'* predicted with a

64   sequence of *N* intervals where the last could be deviant can be, thus, expressed by

65   $JND_N' = \sigma \sqrt{\frac{N}{N-1} - \frac{2r}{N^2}}$. The reader should note that this expression differs from the

66   third equation on page 294 in Schulze (1989), as we believe that the mathematical

67   derivation leads to a second term that should be negative, not positive. Since the *JND'*

68   predicted with a sequence of 2 intervals is $JND_2' = \sigma\sqrt{(2-2r)}$, then (similarly to

69   Eq. 1) we can express the *JND_N'* as a function of the empirical *JND_2* and *r* as such

70   $$JND_N' = JND_2 \sqrt{\frac{1}{2-2r}\left(\frac{N}{N-1} - \frac{2r}{(N-1)^2}\right)}$$ Eq. (2).

71   The patterns that can be obtained with this formula as a function of *r* are shown in

72   Figure 1.

73   Schulze proposed that the non-correlated formulation did not capture the

74   results as well as the negatively correlated formulation, especially in the interleaved

75   condition (Schulze, 1989). However, the value of coefficient of correlation, *r*, was not

76   determined in the original manuscript. Also, Schulze did not analyse the case where

77   noise in successive samples could be positively correlated (such cases could be due to

78   protracted variation of attention whose duration spans multiple stimuli), giving rise to

79   a lesser improvement in performance as a function of sequence duration. We instead

80   perform this analysis and evaluate the predictions of the model with different

81   correlation (Figure 1). With these quantitative predictions, we will be able to compare

82   the predictions of all models with the empirical data.

83   **1.2 Multiple Look (ML) Model Description**

84   Drake and Botte (1993) investigated participants' ability to judge the difference in

85   tempo that happened not at the end of the sequence as in Schulze (1989), but in the

86   middle of the sequence. The change in tempo, thus, creates two isochronous

87   sequences with different rhythms. The authors focused the analysis on the presence of

88   multiple estimates of interval duration, and for this they coined the name Multiple-

89   Look model (ML). The model posits that the precision of the estimate improves as the

90   number of 'looks' at each sequence increases. The ML model has a formulation that is

91   consistent to the model proposed by Schulze's (1989) with uncorrelated noise, where

92   the multiple estimates of the intervals are stored in memory and their average is

93   compared. Here, we will show how to derive the expression of the ML model

94   following the logic of Schulze's (1989) demonstrating their mathematical equivalence.

95   In the task of judging a tempo change in the middle of the sequence, participants

96   perform the discrimination by comparing the average of the duration of the first *N/2*

97   intervals to the average of the second set of N/2 intervals. The noise in the estimate of

98 half the sequence is $V\left(\frac{1}{N/2}\sum_{i=1}^{N/2}E_i\right) = \frac{\frac{N}{2}\sigma^2}{\frac{N^2}{4}} = \frac{2}{N}\sigma^2$. So, the *JND* for a sequence of *N*

99 intervals, where the change in tempo happens in the middle of the sequence is

100 $JND_N' = JND_2\sqrt{\frac{2}{N}\sigma^2}$ and by expressing it as a function of the empirical JND_2 we

101 obtain

102 $$JND_N' = \frac{JND_2}{\sqrt{N}}$$ Eq. (3).

103 Miller and McAuley (2005) suggested a generalized ML model, whereby the

104 two sequences (denoted n_1 and n_2, respectively, so that $N=n_1+n_2$) make independent

105 contributions to the performance. Again, in Schulze's (1989) framework participants

106 compare the average of the n_1 intervals to the average of the n_2 intervals, with a *JND'*

107 that is $JND_{n_1+n_2}' = \sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}$, or expressed as a function of the empirical JND_2 we

108 obtain:

109 $$JND_{n_1+n_2}' = \sqrt{\frac{1}{2}\frac{JND_2^2}{n_1} + \frac{1}{2}\frac{JND_2^2}{n_2}}$$ Eq. (4).

110 It should be noted that this is a more general expression of the previous two

111 formulations when noise is considered uncorrelated, so that with $n_2=1$ the formula is

112 identical to Eq. 1 and with $n_1=n_2$ the formula is identical to Eq. 3.

113 The model of Miller and McAuley (2005) slightly departs from this

114 formulation. Eq. 4, predicts that the JND_{n1+n2} should decrease as the number of 'looks'

115 increases for either of the two intervals. For Miller and McAuley, instead, the

116 contribution of the two sequences is allowed to vary depending on a weight parameter,

117 *w* as such:

118 $$JND_{n_1+n_2}' = \sqrt{w\frac{JND_2^2}{n_1} + (1-w)\frac{JND_2^2}{n_2}}$$ Eq. (5).

119 According to Miller and McAuley, the parameter *w* modulates the contribution of the

120   two averaged estimates. If $w = 1$ then the discrimination performance would be

121   determined only by average of the first series of intervals, whereas if $w = 0$ then the

122   *JND* would be determined by average of the second series of intervals. Such

123   parameter cannot be reconciled with the functioning of the model proposed by

124   Schulze (1989), as both averages are required to perform the discrimination and are,

125   thus, influencing the performance.

126   If the general ML model expressed by Eq. 5 is instantiated for the case

127   analysed by Schulze (1989) where the change in tempo happens at the last stimulus

128   (*n1=N-1* and *n2=1*) the formula becomes

129
$$JND_N' = \sqrt{\frac{w(JND_2)^2}{N-1} + \frac{(1-w)(JND_2)^2}{1}} = JND_2\sqrt{1 + w\frac{2-N}{N-1}} \qquad \text{Eq. (6).}$$

130   In the generalized ML model (Eq. 6), the weight parameter w ranges between

131   0 and 1 and describes how much reliance a participant has on the first of two

132   sequences to be compared. The patterns of performance vary according to this value

133   as shown in Figure 1. The model is based on the presence of a memory store to which

134   future intervals are compared (Treisman, 1963). After comparison, the memory store

135   is updated integrating every presentation of intervals, i.e., to form an internal

136   reference (see Dyjas et al., 2012). In the formula, the weight w captures the proportion

137   (across trials) in which the participant stores a combined memory trace of all

138   previously presented intervals. With $w = 1$, the store is used in a statistically optimal

139   fashion, combining information from all the preceding intervals. In this case, the

140   JND_N' is determined by the limited precision of the comparison of the last interval

141   with such a memory trace. With $w = 0$, instead, the store does not integrate

142   information across intervals, thus it only contains a representation of the latest interval

143   presented. Performance reflected by JND_N' with $w = 0$ is, thus, determined by the

144   precision in comparing the last interval in a sequence with the previous one,

145   regardless of how many preceding intervals there are.

146   The goal of the ML Model is to quantify the discrimination performance with

147   two sequences of regular intervals. With this task, several studies have reported

148   results consistent with the ML model (Grondin, 2001; Ivry & Hazeltine, 1995;

149   McAuley & Jones, 2003; McAuley & Kidd, 1998; ten Hoopen, et al., 2011), although

150   others have not found a close match with its predictions (Grondin, 2001; Hirsch et al.,

151   1990; ten Hoopen et al., 2011). Furthermore, Grondin (2001) demonstrated a ML

152   effect with visual stimuli only if tempo was compared in two separate sequences,

153   whereas the effect was not present if a change in tempo happened within one

154   sequence. Ivry and Hazeltine (1995) also compared one sequence performance with

155   performance in two sequences, but with audio stimuli, finding a ML effect in both.

156   **1.3 Internal Reference (IR) Model Description**

157   The models examined so far are based on averaging the duration estimates of multiple

158   intervals and comparing this value a final duration estimate. Such a process requires

159   the storage in memory of all the estimates of all intervals to obtain a statistically

160   optimal average. However, a more efficient alternative formulation is to compute the

161   average iteratively each time a new estimate becomes available. As per the IR model,

162   such a procedure can be performed using a recursive estimator, like the Kalman filter.

163   The mean with $N=n+1$ estimates is a weighted average of the mean μ_n of the

164   previous n estimates and of the last estimate E_{n+1}, which can be expressed as

165     $$\mu_{n+1} = \frac{n}{n+1}\sum_{i=1}^{n+1} E_i = \frac{n}{n+1}\mu_n + \frac{1}{n+1}E_{n+1} \qquad \text{Eq. (7).}$$

166   where $K = \frac{1}{(n+1)}$ is called the gain factor and indicates how the weight given to the

167   single E value decreases with longer sequence. This idea is similar to the concept of a

168   clock model in time perception (Gibbon et al., 1984; Treisman, 1963), where the

169 representation of duration increases in precision by averaging the representation of

170 successive estimates of intervals, thus leading to better performance (Dyjas et al.,

171 2012; Schulze, 1979). If estimates are independent, this formula leads to the same

172 variance decrease obtained by averaging all stimuli at once expressed by Eq. 1. On

173 the positive side, however, this way of computing the average reduces the memory

174 requirements to only a single estimate value at the time (plus the knowledge of how

175 many stimuli have been averaged) albeit it increases the complexity of the

176 computation, because a weighed average is required for each iteration. The iterative

177 process, however, does not lead to statistically optimal variance reduction with

178 positively correlated noise estimates.

179   An alternative to this scheme has been proposed by Dyjas et al. (2012),

180 originally to account for serial effects in tasks requiring the comparison of two

181 durations. The authors propose that weights are different from the statistically optimal

182 *K* and do not depend on the sequence length. Instead, they propose a weight *g* for

183 modulation of the current estimate and the contribution of the previous reference:

$$\mu_N = \mu_{n+1} = (1 - g)\mu_n + gE_{n+1} \qquad \text{Eq. (8).}$$

185 Such a scheme leads to a geometric moving average (Roberts 1959), where the weight

186 *g* assigned to the historical list of estimates decreases as a geometric progression

187 when time passes. The variance associated with such averaging method is (see Dyjas

188 et al., 2012) $V(average) = \frac{s^2(g^{2n}+(1-g)^2(1-g)^{2n})}{1-g^2}$. As the participant would be

189 comparing this average to the last interval, the predicted *JND'* for a sequence of *N*

190 interval can be calculated as $JND_N' = \sqrt{\frac{s^2+s^2(g^{2n}+(1-g)^2(1-g)^{2n})}{1-g^2}}$, whereas for a

191 sequence of only two intervals, the *JND₂'* would be $JND_2' =$

192 $\sqrt{s^2 + s^2(g^2 + (1 - g)^2)}$. Performing the substitution of *JND₂'* in *JNDₙ'* gives

193 $$JND_N' = JND_2 \sqrt{\frac{(1+(g^{2n}+(1-g)^2(1-g^{2n})))}{\frac{1-g^2}{1+(g^2+(1-g)^2)}}}$$ that simplifies to:

194 $$JND_N' = JND_2 \sqrt{\frac{g^{(1+2n)}+1}{g^3+1}}$$ Eq. (9).

195 Predictions of the IR model expressed in Eq. 9 are shown in Figure 1 for different

196 values of g. It is immediately evident that such a formulation cannot predict the same

197 improvement and decrease in performance as the other proposals derived from

198 Schulze (1989).

199 **1.4 Diminishing Returns (DR) function**

200 Ten Hoopen et al. (2011) investigated the issue of temporal sensitivity in a single

201 sequence of audio stimuli where the change in tempo could happen at one of several

202 positions. They found that performance changed more as a function of the number of

203 intervals *before* the tempo change, rather than after. They adopted a reciprocal DR

204 function to capture the performance increase:

205 $$JND_{n_1: n_2} = a + \frac{b_1}{n_1} + \frac{b_2}{n_2}$$ Eq. (10).

206 where a is the asymptotic performance and b are the amount of performance increase

207 for each added interval before and after the tempo change. The parameters fitting the

208 results of Ten Hoopen et al. highlight that performance increment is higher for

209 changes before the tempo change are captured by $b_1 > b_2$. It should be noted that the

210 DR function expressed in Eq. 10 is not based on a process oriented model as the one

211 proposed for example by Schulze (1989), because purpose was to fit the data. With

212 this specification, in the rest of the manuscript we will refer to the DR as a model

213 rather than a function. Eq. 10 can nevertheless be used to express the *JND* of a

214 sequence of intervals where the last one is deviant as a function of the *JND* obtained

215 in a sequence with two intervals. If we define c as the combined factor $c = a + b_2$

12

216 and we simplify JND_2 to be $JND_2 = c + b_1$ then JND_N cab be expressed as a function

217 of JND_2 and c as such:

218 $$JND_N = c + \frac{JND_2 - c}{n-1}$$ Eq. (11).

219 The ability of the DR model expressed in Eq. 11 to capture an improvement in

220 performance in our empirical study can be analysed by looking at the range of

221 possible fittings in Figure 1 (i.e., the change in the predictions of the DR as a function

222 of the c parameter).

223 **1.5 Experimental question**

224 The models analysed so far (PA, ML, IR, DR) all make predictions that

225 discrimination performance improves as the number of intervals to be examined

226 increased. There are, however, quantitative differences in the predictions by Schulze's

227 (1989) PA model (Eq. 1 and Eq. 2), the ML model (Eq. 6), the IR model (Eq. 9), and

228 the DR model (Eq. 11). In this paper, we hope to be able to determine which model

229 captures the data of two experimental conditions (interleaved and blocked

230 presentation of duration) using the free parameter that each model has (respectively:

231 correlation r, weight w, gain factor g, and combined factor c).

232

233

234

235
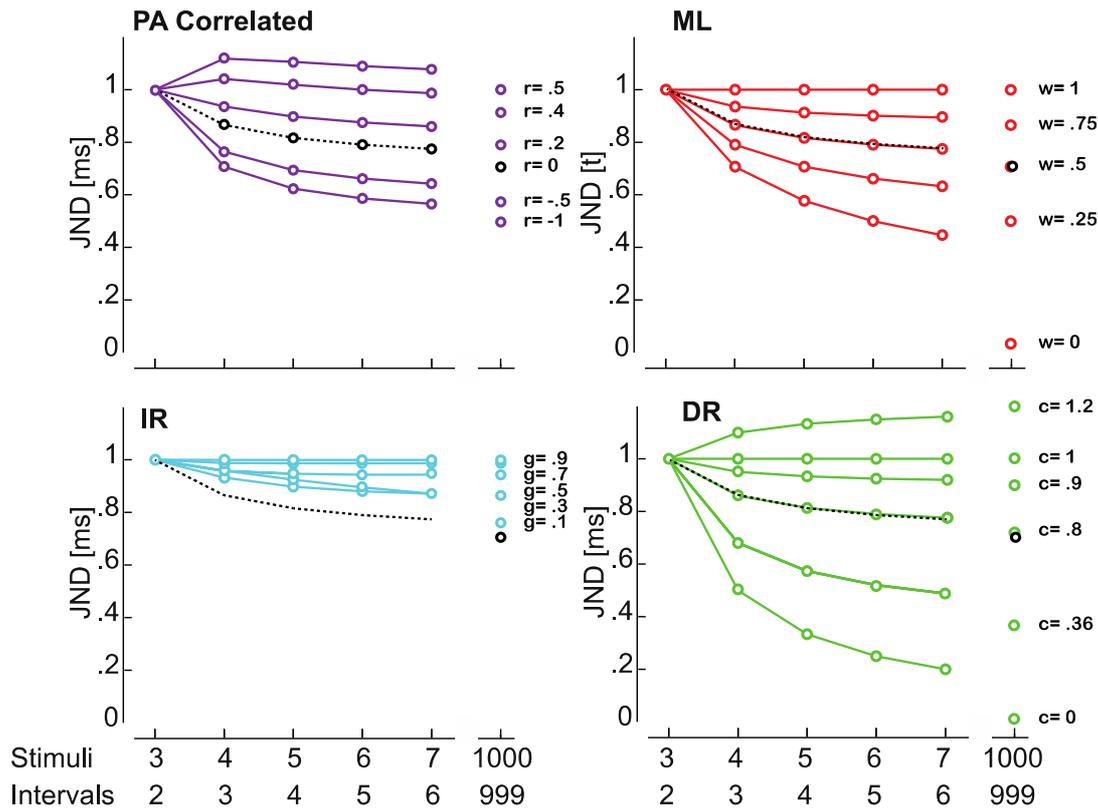
Figure 1. Predictions for the Percept Averaging (PA, Eq. 1 and 2), Multiple Look (ML, Eq. 6), Internal Reference (IR, Eq. 9), and Diminishing Return (DR, Eq. 11) models for JND_N with a sequence of N stimuli expressed as a function of JND_2=1ms. Each model has a single free parameter that has been varied to show the range of patterns that can be captured by the models. The value of the parameters for the DR model has been tuned (as discussed in the results section) to capture statistical optimality obtaining a value of $c=0.8$.

As in Schulze's (1989) study, we investigate the case where sequence lengths are presented either interleaved or blocked. Schulze found that only in the case of the interleaved presentation there was an increase in performance with longer sequences. In contrast to Schulze's studies (1978; 1989), we allow the last interval to be either longer or shorter than the previous ones. That is, the last stimulus could be presented anisochronously compared to the previous sequence, either too early or too late. The

14

251   task is similar to ten Hoopen et al.'s (2011), as participants are asked to judge whether

252   the last stimulus was presented 'earlier' or 'later' than isochrony (i.e., they reported

253   whether the last interval was shorter or longer than the previous ones). The analysis of

254   'earlier' vs. 'later' judgments allows us to determine whether temporal expectations

255   generated by the sequence of stimuli with identical interval can cause a consistent bias

256   in perceived isochrony, an analysis that was possible but has not been performed by

257   ten Hoopen et al. The motivation for this new analysis is to try to account for any

258   consistent bias in responses with a perceptual mechanism. In particular, a bias in

259   perceived isochrony can be explained by appealing to a modification of the perceived

260   timing of the last stimulus in the sequence. This possibility requires a difference in the

261   formulation of the problem of perceived isochrony as has been done so far: rather

262   than considering the perceptive duration of the individual interval, here we propose to

263   analyse the perceived timing of stimuli. In particular, we analyse the time at which the

264   last stimulus in the sequence is perceived, which is presented right after the change in

265   tempo. Perceived timing of stimuli can be affected by several factors in a way

266   independent from perceived duration.

267   Titchener (1908) was the first to suggest that attention (among other factors)

268   can modulate perceived timing of individual stimuli as a fully attended stimulus is

269   processed faster than an unattended one. Summerfield and Egner (2009) investigated

270   the contribution of attention in a recognition task supporting the idea of prioritized

271   processing of attended stimuli. Such attentional facilitation speeds up perception, an

272   effect termed *prior entry*, which has been highlighted in studies involving temporal

273   judgments (Sternberg & Knoll, 1973; Shore et al., 2001; Vibell et al., 2007; Zampini

274   et al., 2005; for a review see Spence & Parise, 2010) and at the neural level

275   (McDonald et al., 2005). According to a time-frequency analysis of

electroencephalographic (EEG) recordings by Rohenkohl and Nobre (2011), decreased brain activity in the alpha band for expected stimuli is correlated with faster responses, tentatively suggesting a neural basis for the prior entry hypothesis.

To evidence the relationship between attention and perceptual acceleration we manipulated task demand by presenting stimulus sequences of different length either in an interleaved or blocked presentation. This condition was also present in the original study by Schulze (1989). We posit that in the interleaved condition, participants do not know when the sequence will end and thus will have to pay closer attention. Such uncertainty will increase the reliance on sensory predictions, which should result in a stronger prior entry effect. The perceived timing of stimuli in the interleaved condition should be accelerated and, consequently, perceived isochrony should be obtained with slightly delayed stimuli (and thus slightly longer intervals) rather than stimuli presented at the expected time point.

2. Methods and Materials

2.1.1 Participants

Twenty-five undergraduate students (age range from 18 to 25 years and mean age of 21.3 years) with self-reported normal hearing were recruited by the research participation system of the University of Birmingham. They gave informed consent before taking part in the experiment and were rewarded with course credits or a payment of six pounds per hour. Ethical guidelines have been followed in all the experiments and were approved by the STEM Ethics Committee of the University of Birmingham.

2.1.2 Design

There were two sessions, one with interleaved presentation and one with blocked presentation of trials with different sequence lengths: 3, 4, 5 or 6 stimuli (2, 3, 4 or 5 intervals). For every sequence length, the timing of the last stimulus was selected among 15 possible anisochronies: ±0, 20, 40, 60, 80,

301   100, 150, and 200 ms. The trial types resulting from the combination of blocked/interleaved

302   presentation (2), sequence length (4), and anisochrony of the last stimulus (15) were repeated 8 times in

303   order to determine the parameters of eight psychometric functions (see results) for a total of 960 trials

304   per participant.

305   2.1.3 Stimuli

306   Stimuli were identical tones produced by a speaker located on a desk approximately 50 cm

307   from the participant (20 ms with 5 ms linear ramp, 1 kHz, 75.1 dBA). Trials were composed of a

308   different number of stimuli within a sequence, where intervals between successive stimuli in the

309   sequence remained the same (IOI = 700 ms) for all but the final stimulus, which could be presented at

310   different anisochronies.

311   2.1.4 Procedure

312   Participants sat in a quiet testing cubicle. A sequence of auditory stimuli of different lengths

313   were presented in which the participants had to respond whether the anisochrony of the final stimulus

314   was 'earlier' or 'later' than the expected timing (Fig. 2). Sequence lengths were either presented

315   blocked or interleaved and the order of the two presentations was counterbalanced across participants.

316
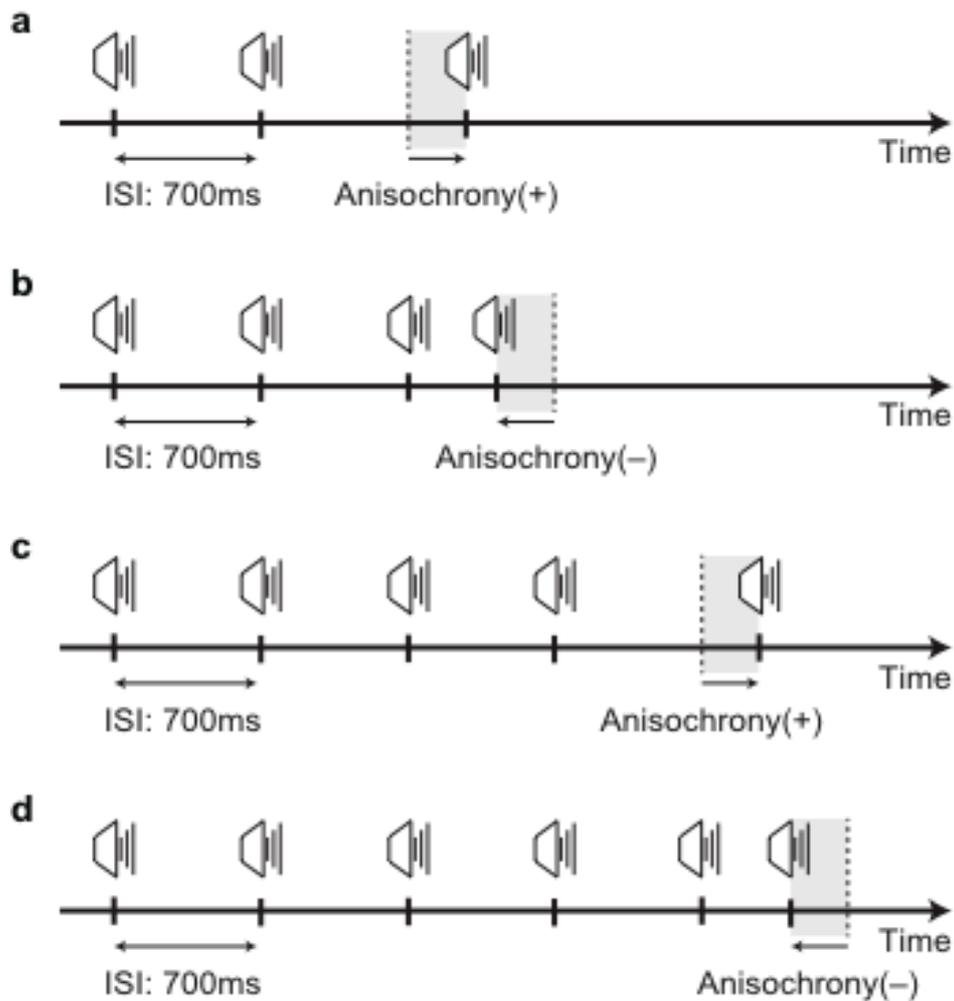
Figure 2. Examples of trials with different sequence length. (a) Sequence of three stimuli (two intervals) where the final stimulus is presented later than expected (+ Anisochrony). (b) Sequence of four stimuli (three intervals) where the final stimulus is presented earlier than expected (- Anisochrony). (c) Sequence of five stimuli (four intervals) where the final stimulus is presented later than expected (+ Anisochrony). (d) Sequence of six stimuli (five intervals) where the final stimulus is presented earlier than expected (- Anisochrony).

2.2.1 Data Analysis

We analyzed the proportion of 'later' responses for each anisochrony of the last stimulus, to obtain a distribution for each sequence length with interleaved and with blocked presentation. In order to determine if a change in the perceived isochrony of stimuli changes due to temporal expectations

330   and attention, we calculated the *Point of Subjective Equality* (*PSE*) as the anisochrony at which

331   participants are most unsure about whether the final stimulus was presented early or late. Thus, the *PSE*

332   is the time point the last stimulus needs to be presented in order for it to be perceived as being

333   isochronous. The *PSE* is obtained by calculating the first order moment of the difference between

334   successive proportions of responses using the Spearman-Kärber method (see Ulrich & Miller, 2004, for

335   further details of this method). The second order moment is proportional to the inverse slope of the

336   psychometric function, which here is termed *JND*.

337   To obtain *PSE* and *JND,* we employ the Spearman-Kärber method, which is a non-parametric

338   estimate that avoids assumptions about the shape of the psychometric functions underlying the

339   participants' responses. The formulae below are used to estimate the first and second moment of the

340   psychometric function underlying the data. First we define the 15 anisochronies of the final stimulus,

341   where ANI_i with $i=\{1, ... 15\}$ and p_i with $i=\{1, ... 15\}$ as the associated proportion of 'later' responses.

342   We further define $ANI_0=-250$ ms, $ANI_{16}=+250$ ms and we assume $p_0=0$ and $p_{16}=1$, to be able to

343   compute the intermediate *ANI* between two successive ones

344
$$s^i = \frac{ANI_{i+1} + ANI_i}{2}, \text{ with } i=\{0, ... 15\} \qquad \text{Eq. (12).}$$

345   and the associated values of the difference in proportion of responses, taken at and above 0 to

346   monotonize the proportion of responses

347
$$dp_i = max(0, \ p_{i+1} - p_i), \text{ with } i=\{0, ... 15\} \qquad \text{Eq. (13).}$$

348   With these indexes we can express *PSE* and *JND* analytically as such:

349
$$PSE = \frac{1}{\sum_{i=0}^{15} dp_i} \sum_{i=0}^{15} s_i \ dp_i \qquad \text{Eq. (14).}$$

350   and

351
$$JND = \sqrt{\frac{1}{\sum_{i=0}^{15} dp_i} \sum_{i=0}^{15} dp_i \ (s_i - PSE)^2} \qquad \text{Eq. (15).}$$

352

353   2.2.2 Model Fitting

354   In order to find the best fit for the each of the model's parameter, for each participant we found the

355   minimum sum of squares difference between the predicted JND_N' and the empirical JND_N. In Schulze's

356 PA model (Eq. 2) the minimisation is done with the correlation, in the generalized ML model (Eq. 6)

357 with the weight, in the IR model (Eq. 9) with the gain factor, and the DR model (Eq. 11) with the

358 combined factor. The fitting is done independently for the two conditions (blocked vs. interleaved).

3. Results

360 The average proportion of responses across participants for sequences of different

361 lengths and type of presentation (interleaved and blocked) are shown in Fig. 3. A

362 consistent difference in the shape of the response distributions with blocked and

363 interleaved presentation is evident across the various sequence lengths.

364 Discrimination performance was characterised by *JND* values (Fig. 4), which

365 are calculated according to the Spearman-Kärber method (see method section). The

366 proportions of 'late' responses in each psychometric function were monotonized prior

367 to analysis. To determine whether temporal sensitivity improves with sequence length

368 and whether differences in sensitivity existed between blocked and interleaved

369 presentations, *JND* values were submitted to a two-way repeated measure ANOVA

370 with factors condition (blocked or interleaved) and number of intervals in the

371 sequence (2, 3, 4 or 5). Results indicate better discrimination with blocked

372 presentation of sequence length ($F(1,24)=20.3$, $p<0.001$, $\eta_p^2=0.46$, Fig. 3c), an

373 improvement in performance due to sequence length ($F(3,72)=3.4$, $p=0.022$, $\eta_p^2=0.12$), and a significant interaction between the two factors ($F(3,72)=4.1$, $p=0.009$,

374 $\eta_p^2=0.38$). Such an interaction suggests that the improvement in temporal

375

376 discrimination due to sequence length is present with the interleaved presentation of

377 different sequence length (one-way repeated measure ANOVA with factor sequence

378 length: $F(3,72)=5.1$, $p<0.003$, $\eta_p^2=0.18$) but performance is not affected with

379 blocked presentation of one length ($F(3,72)=2.0$, $p=0.119$, $\eta_p^2=0.12$).

380   Biases in perceived isochrony are captured by *PSE* values (Fig. 5), which are

381   also calculated according to the Spearman-Kärber method (see method section). In

382   both conditions, we find that stimuli presented physically isochronous are actually

383   reported more often to appear earlier than expected. Perceived isochrony is obtained

384   when the last stimulus was presented later than it should – i.e., with a longer last

385   interval (single sample t-test of *PSE* calculated on the data against 0: interleaved,

386     $t(24)=6.1$, $p<0.001$, blocked: $t(24)=2.6$, $p=0.015$). In order to test whether there is a

387   consistent difference of this effect with blocked or interleaved presentation of

388   sequence lengths, we submitted *PSE* values a two-way repeated-measures ANOVA

389   with factors presentation condition (interleaved or blocked) and number of interval in

390   the sequence (2, 3, 4 or 5). Results indicate a change in *PSE* depending on the

391   presentation condition ($F(1,24)=13.4$, $p=0.001$, $\eta_p^2=0.36$), as the final stimulus in the

392   interleaved condition has to be presented 24.6 ms (4.0 ms SEM) after isochrony in

393   order to be perceived isochronous, whereas the last stimulus in the blocked condition

394   has to be presented 12.1 ms (4.6 ms SEM) after isochrony. The difference between

395   both interleaved and blocked condition was 12.4 ms (4.5 ms SEM). We find no main

396   effect of sequence length or an interaction (both $p > 0.11$).

397   In sum, the sensitivity of detecting anisochrony increases with longer

398   sequences if different lengths are interleaved but is overall higher if only one

399   sequence length is presented in a block. Perceived isochrony is consistently biased

400   and the observed bias does not change due to sequence length, but it is affected by the

401   presentation condition (interleaved and blocked). Not knowing the serial position of

402   the interval to be judged leads to a higher bias, so that the sequence is perceived as

403   being isochronous if the last stimulus is presented slightly later, i.e., after a longer

404   interval compared to the previous ones.
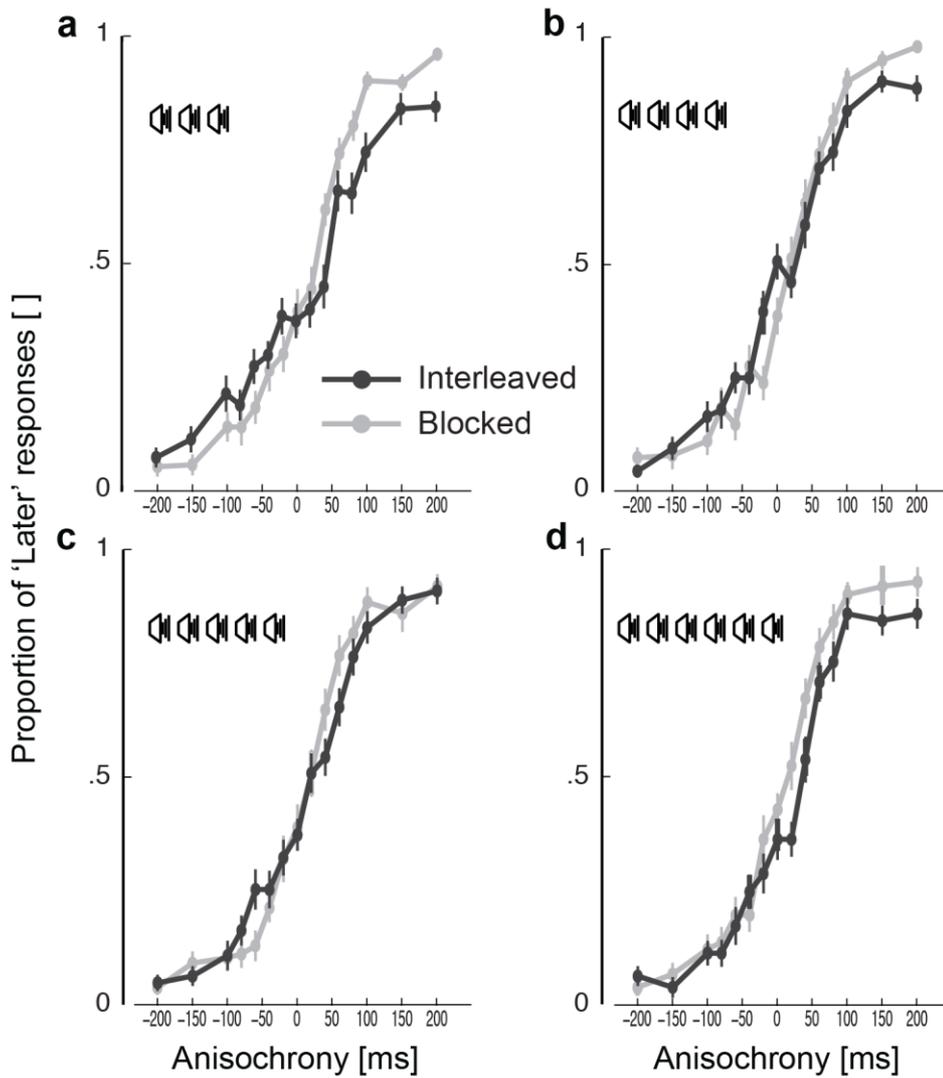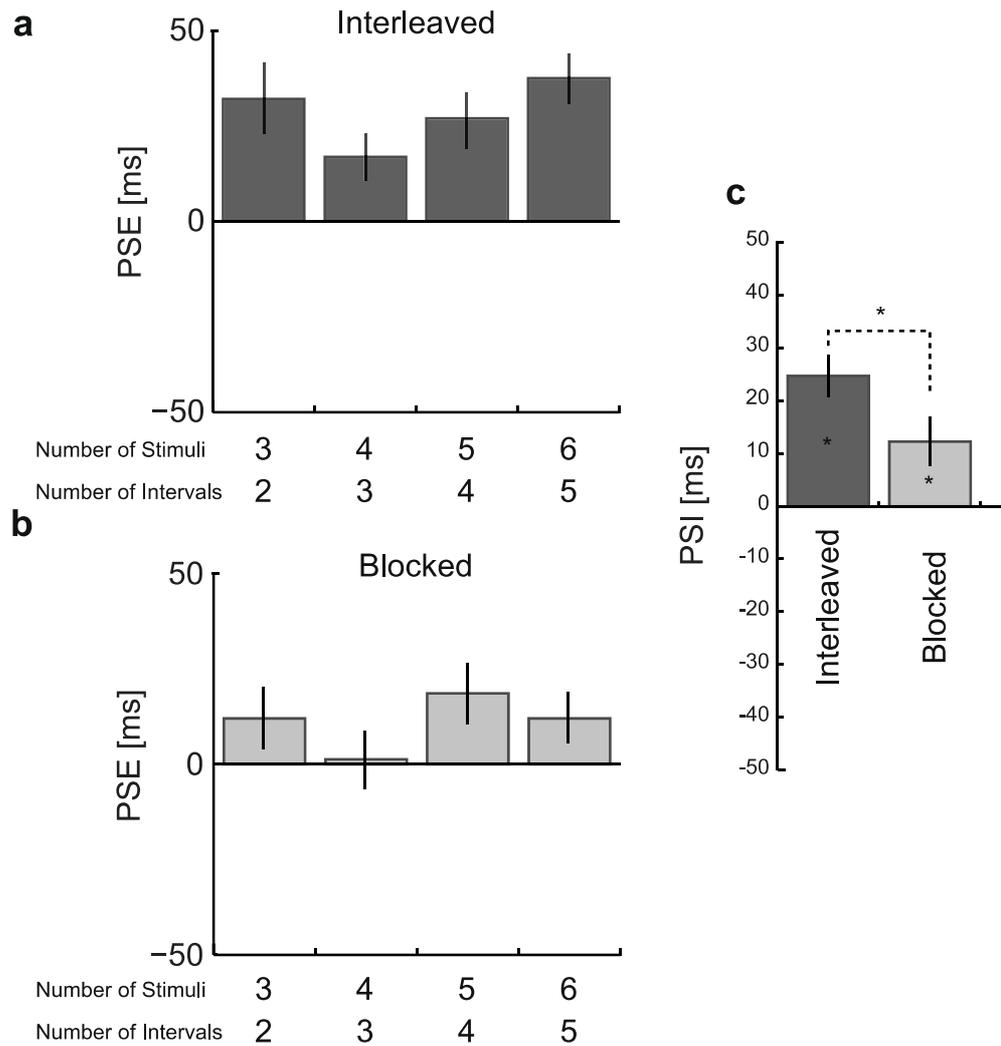
Figure 3. Proportion of 'later' responses as a function of the anisochrony of the final interval in the sequence for (a) 2, (b) 3, (c) 4, and (d) 5 intervals for interleaved and blocked presentation. Asterisks indicate significant difference between the two conditions according to the values in Table 1. Error bars represent the standard error of the mean.
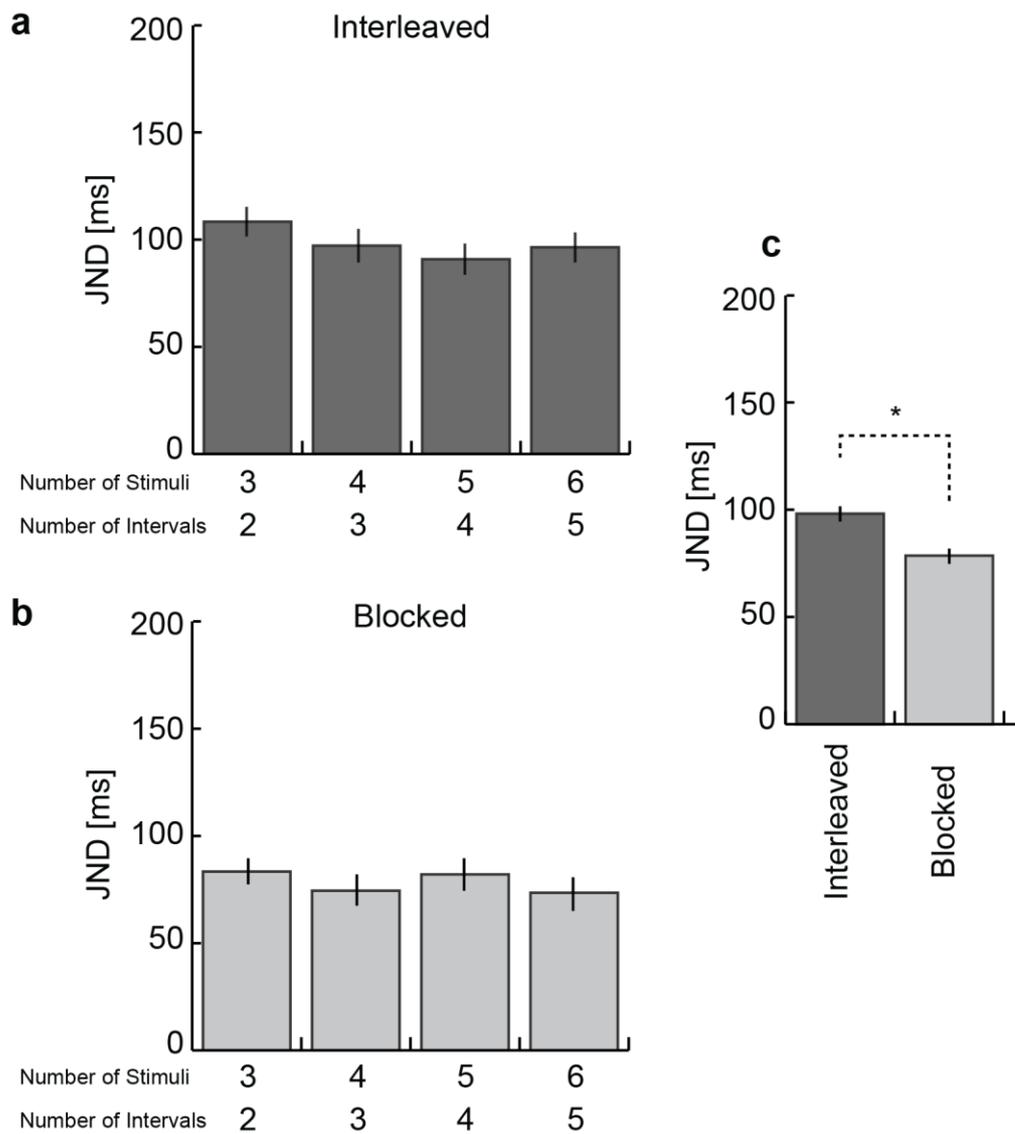
412

413　Figure 4. *JND* values as a function of sequence length for (a)

414　interleaved and (b) blocked presentation. (c) *JND* values

415　calculated on the proportion of 'later' responses across sequence

416　lengths for blocked and interleaved conditions. The asterisk

417　indicates a significant difference according to the ANOVA

418　presented in the text. Error bars represent the standard error of

419　the mean.

a

PSE [ms] — Interleaved

50 ... 0 ... −50

| Number of Stimuli | 3 | 4 | 5 | 6 |
| Number of Intervals | 2 | 3 | 4 | 5 |

b

PSE [ms] — Blocked

50 ... 0 ... −50

| Number of Stimuli | 3 | 4 | 5 | 6 |
| Number of Intervals | 2 | 3 | 4 | 5 |

c

PSI [ms] — Interleaved / Blocked

50 40 30 20 10 0 -10 -20 -30 -40 -50

Figure 5. *PSE* values as a function of sequence length for (a) interleaved and (b) blocked presentation. (c) *PSE* values calculated on the proportion of 'later' responses across sequence length for interleaved and blocked presentation. The asterisk indicates a significant difference from 0 according to single-sample t-tests and between conditions according to the ANOVA (details presented in the text). Error bars represent the standard error of the mean.

3.1 PA Model Results

The Schulze (1978; 1989) PA model predicts that as the representation of previous duration becomes more accurate with longer sequences, and as such, increases

432   temporal sensitivity. We applied Eq. 1 to our data and (without any fitting procedure)

433   it generally captures the decrease in the empirical JND_N in the interleaved condition

434   and blocked condition (Fig. 6) with very similar sum of squares differences in the

435   interleaved and blocked conditions, 1182±118 ms^2 and 1210±277 ms^2 respectively

436   ($t(24)$=0.08, p=0.94; Fig. 7).

437   Extending the Schulze (1989) model to include correlated noise lead us to

438   employ Eq. 2. We found the minimum sum of squared differences between the

439   predicted $JND_N{'}$ and the empirical JND_N across the four durations for each participant

440   through an exhaustive search of the value of correlation r. Predicted values that

441   minimise such difference are shown in Figure 6. Such procedure will be employed for

442   the following models and makes the models equivalent in terms of number of fitted

443   parameters. The sums of squared differences for the PA Correlated model are

444   825±183 ms^2 and 587±115 ms^2 which, notably, are significantly lower than the values

445   obtained with the unfitted PA Uncorrelated model (interleaved: $t(24)$=2.5, p=0.017;

446   blocked: $t(24)$=5.3, p<0.001; Fig. 7). Despite this improvement, the average

447   correlations that lead to the minimum sum of square difference for each participant in

448   each condition are quite small -0.056±0.091 and -0.124±0.092 and do not differ from

449   0 (interleaved: $t(24)$=1.1, p=0.28; blocked: $t(24)$=1.4, p=0.18) nor differ from each

450   other ($t(24)$=0.5, p=0.59).

451   **3.2 ML Model Results**

452   Like above, the ML model predicts that sensitivity to changes in tempo increases with

453   longer sequences with a factor that limits performance compared to statistical

454   optimality, the difference from 0.5 of the weight assigned to the two parts of the

455   sequence (Drake & Botte, 1993; Miller & McAuley, 2005). Here we allowed

456   individual participants' weights to span a range between -0.5 and 1.5 as noise between

457 successive estimates can be correlated (see Schulze, 1989 and Oruç et al., 2003 for

458 more detail). We performed the same sum of squared error minimization procedure as

459 for the PA Correlated model. Predicted values of JND_N' that minimise error are

460 overlaid to the empirical values in Fig. 6. Average weights are 0.39±0.09 and

461 0.24±0.11 for the interleaved and blocked condition respectively, they differ from 0.5

462 (single sample t-test against 0.5, interleaved: $t(24)=2.6$, $p=0.014$; blocked: $t(24)=3.0$,

463 $p=0.006$) but they do not differ significantly ($t(24)=1.1$, $p=0.26$). The model captures

464 the increasing sensitivity in the interleaved condition slightly, but not significantly,

465 worse than for the blocked condition – as the values of the average sum of squared

466 differences for the ML model are 802±180 ms^2 and 579±119 ms^2 for the interleaved

467 and blocked conditions respectively, do not differ significantly ($t(24)=1.0$, $p=0.32$;

468 Fig. 7). The performance of the ML model in capturing the data is not significantly

469 different than the PA Correlated model (t-test on average SSE across the two

470 conditions between ML and PA $t(24)=1.0$, $p=0.30$).

471 **3.3 IR Model Results**

472 Slightly different from the averaging models stated above, the IR model proposed by

473 Dyjas et al. (2012) can only capture a limited range of improvements in temporal

474 discrimination (Fig. 4). The factor limiting performance is the weight of the current

475 estimate g, which here was tuned with the same procedure followed above. The best-

476 fitting weight g is 0.61±0.07 in the blocked and 0.66±0.05 in the interleaved condition,

477 which do not differ significantly ($t(24)=0.5$, $p=0.65$). The sum of square difference for

478 the IR model is 1000±180 for the interleaved condition and 778±162 for the blocked

479 condition (Fig. 7). Such values are higher than the PA Correlated and MLM models

480 (t-test on average SSE across the two conditions between IR and: PA $t(24)=3.7$,

481 $p=0.0011$, MLM: $t(24)=4.3$, $p<0.001$).

3.4 DR Model Results

We also fitted the results using the DR model proposed by ten Hoopen et al. (2011).

Akin to the previous models, the DR model predicts that temporal sensitivity to

irregularities increases with the amount of intervals presented. However, with each

additional interval, the increase in sensitivity is less and less. We applied Eq. 10 to

our data and found the best fit for the combined parameter c. Predicted average values

of JND_N' with such individually tuned parameters are presented in Fig. 6. We find

that the values that best fit the empirical data for the combined factor c in the

interleaved condition are 78.8 ± 10.2 and 105.6 ± 10.2 which differ significantly

($t(24)=336.3$, $p<0.001$). With such values, the average sum of squared error is

2500 ± 524 ms^2 and 3332 ± 574 ms^2 in the interleaved and blocked conditions

respectively which do not differ significantly from each other ($t(24)=0.3$, $p=0.77$), but

it is obviously much higher than all three other models (Figure 7, all $p<0.001$).
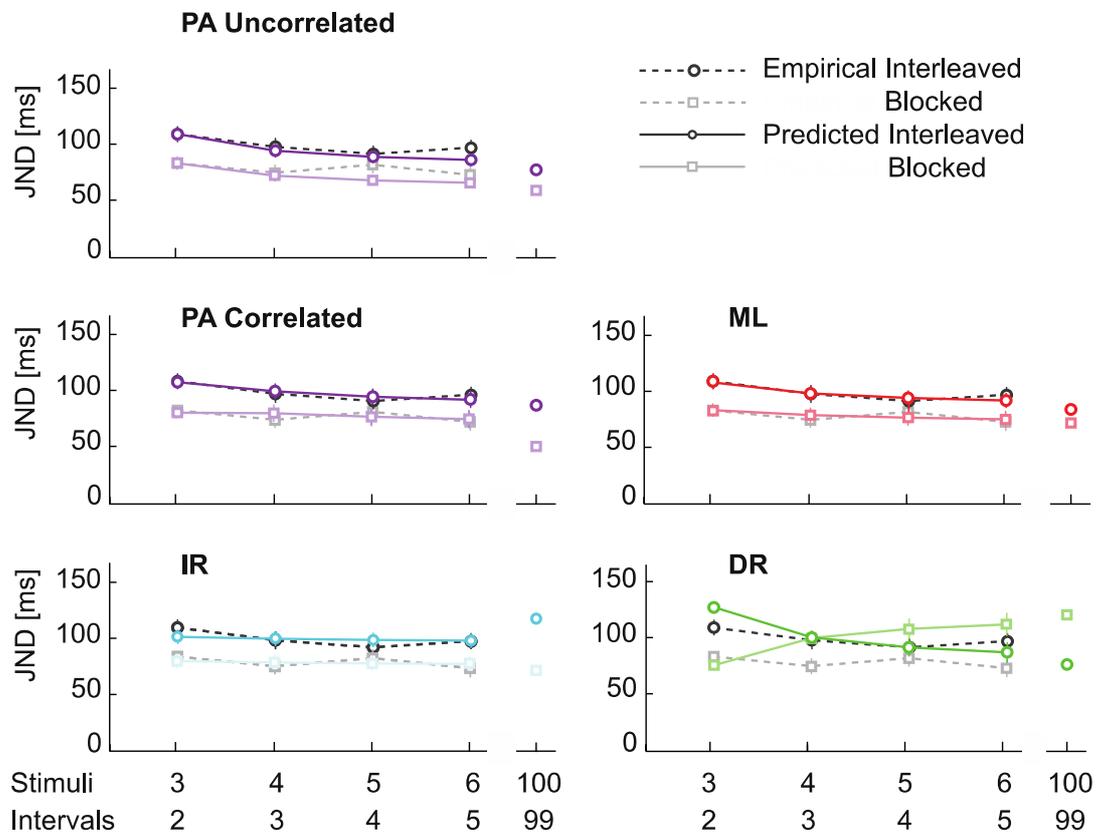
496 Figure 6. Predictions of the Percept Averaging (PA), Multiple Look (ML),

497 Internal-Reference (IR), and Diminishing Returns (DR) model (see results

498 section). The predictions of the PA (Schulze, 1978; 1989) and ML Models

499 (Drake & Botte, 1993; Miller & McAuley, 2005) visually capture the increase in

500 temporal sensitivity as a function of sequence length across the two

501 conditions. The IR model (Dyjas et al., 2012) captures the flat course of JND

502 for the blocked condition but cannot accurately capture the obvious increase

503 in temporal sensitivity for the interleaved condition. The DR Model (ten

504 Hoopen et al., 2011) captures the negatively accelerating course of the *JND*

505 only for the interleaved condition but does not correctly account for flat course

506 of *JND* in the blocked condition, as the fit for several participant predicts

507 worse performance due to the presence of low-performance conditions.
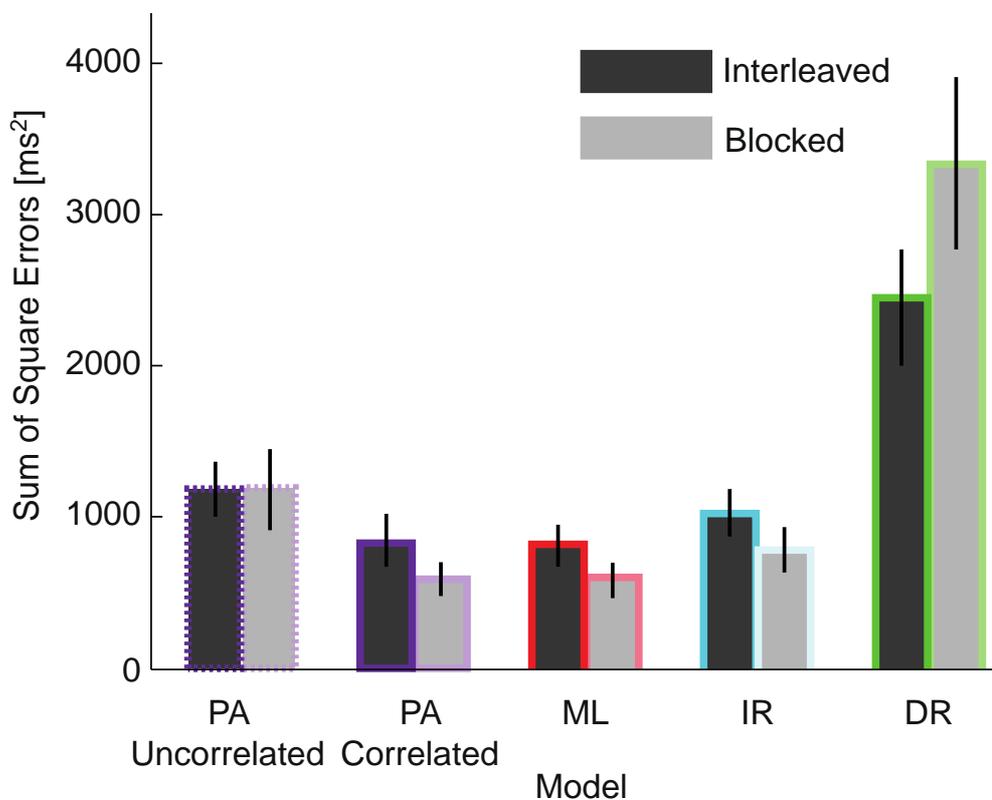
508

509 Figure 7. Comparison of the models fit to the empirical data captured by the

510 sum of squared errors for the Percept Averaging (PA; Correlated and

511   Uncorrelated), Multiple Look (ML), Internal Reference (IR), and Diminishing

512   Returns (DR) models. The dark grey bar represents the interleaved condition

513   whilst the light grey indicates the blocked condition. A 2-way r.m. ANOVA on

514   the data with factors models and interleaved/blocked is significant for the

515   factor model ($F(4,96)=39.37$, $p<0.0001$, $\eta_p^2=0.62$) whereas the factor

516   blocked/interleaved and interaction are not significant. Error bars represent

517   the standard error of the mean across participants.

518

4. Discussion

520   In this paper, we aimed to compare the predictions of existing models of how the

521   brain may deal with detecting deviations from isochrony in sequences of auditory

522   tones. Second, we wanted to see if we could observe any distortions from veridical

523   isochronous perception. To investigate this, similar to previous investigations

524   (Halpern & Darwin, 1982; Hoopen et al., 2011; Schulze, 1978; 1989), we

525   manipulated sequence length across trials (2, 3, 4 or 5 intervals in a sequence). The

526   final interval in the sequence could be presented too early or too late, and participants

527   needed to identify which of the two cases it was. By presenting the final stimulus

528   either earlier or later as ten Hoopen et al. did, we could eliminate response biases that

529   affected the measure of sensitivity. We also tested whether presenting the sequences

530   either interleaved (difficult task as participants do not know the sequence length to be

531   judged) or blocked (simpler task because participants know which interval could be

532   deviant) has an impact on perception. Temporal discriminability (quantified by the

533   *JND* calculated on the proportion of 'later' than expected responses) is found to be

534   higher in the blocked condition than in the interleaved condition. Furthermore, we

535   find that temporal sensitivity increases as a function of sequence length in the

536   interleaved condition, but not in the blocked condition (Fig. 4a,b). This principal

537   finding will now be reviewed in the context of the models of temporal deviation

538   detection.

539   **4.1 Model Comparison**

540   The goal of the paper was to compare existing approaches to how the brain may deal

541   with temporally deviant stimuli. As such, the finding that temporal sensitivity

542   increases as a function of sequence length in the interleaved condition is consistent

543   with the findings of Schulze (1989) and ten Hoopen et al. (2011). However, Schulze

544   found a larger increase in performance with longer sequences than we report here and,

545   thus, it is possible that such a difference could be due to the use of final intervals that

546   could only be longer than the previous ones. The best fit of the predicted JND_N' to the

547   empirical data JND_N was with the PA and MLmodels. The PA model without

548   correlated noise predicted a too large improvement in performance in the blocked

549   condition, but having the correlated noise included in the formulation, the PA model

550   accurately captured the patterns of both conditions. The ML model finely captured the

551   steeper slope of increased temporal sensitivity in the interleaved condition, and the

552   limited improvement of blocked condition performances as well. On the other side,

553   although the IR model was not able to capture the close-to statistically optimal

554   improvement of temporal sensitivity in the interleaved condition, it instead accurately

555   captured the flat course that was observed in the blocked condition. Of all the models

556   we have implemented, the DR model was a relatively demanding fit, as it predicted an

557   increased pattern of JND that we did not find in our averaged blocked condition

558   results. The DR model also over-estimated the improvement of temporal sensitivity in

559   the interleaved condition.

560   The parameters used to fit the models to the data are also interesting. Despite

561   the increase in performance from the PA Correlated compared to the PA Uncorrelated,

562   the correlation parameter r does not significantly vary across conditions nor

563   statistically differs from 0, although there is a slight tendency to negativity as

564   expected by Schulze (1989). Such results leads us to think that beyond the limiting

565   performance increase due to the overall negative weight, the reason for better fit

566   needs to be searched in inter-individual level, i.e., in the different pattern of

567   performance increase for different sequence duration. The fit of the ML to the data is

568   somewhat consistent with this view. Overall, the deviation of the weight from 0.5

569   suggests a limitation in the performance increase. However, the lack of a statistical

570   difference in the weight depending on the conditions points at an inconsistency across

571   participants.

572   The three interval-based models described here (PA, ML, IR) have a common

573   explanation for the increase in sensitivity to temporal properties with longer

574   sequences due to the increase in precision of the duration representation following

575   exposure to multiple intervals (i.e., Dyjas et al., 2012; Schulze, 1979). Such

576   improvement is consistent with internal clock models (Gibbon et al., 1984; Treisman,

577   1963), where duration is judged as the accumulation of 'ticks' from an internal

578   pacemaker. The fact that the fit of the PA model fails to find a difference in

579   correlation and that the ML model fails to find a difference in the weight assigned to

580   the intervals with blocked and interleaved presentation suggest that the integration of

581   information is not complete and, thus, sub-optimal. The result that there is no change

582   in correlation and in weighting is logical, as sensory correlation and memory

583   integration should not be affected by whether the sequence is presented interleaved

584   with other sequence lengths.

585   To further compare the models, we generated predictions for a sequence of

586   100 stimuli (Fig. 1). We find that the models largely differ in their predicted

587   performance. The ML expressed by Eq. 4 should lead to a progressive increase in

588   performance as the sequence increases in length. A similar situation is present for the

589   DR model. In comparison, the Correlated PA of Eq. 2 has a parameter that limits the

590   integration of memory traces (Schulze, 1978, 1989). The IR model has also a hard

591   stop in the performance and cannot go beyond statistical optimality with uncorrelated

592   noise. Thus, the ML and DR models are unable to capture the asymptotic maximal

593   performance with long sequences as they predict impossibly high performance.

594   **4.2 Response Bias**

595   A second aspect that our experiment allowed us to ascertain was the presence of a

596   consistent bias in the reported isochrony, registered as consistent deviations of *PSE*

597   from 0 in Fig. 5. Such bias changed depending on the interleaved/blocked

598   presentation of durations. The PA model could, in principle, capture biases in

599   perceived isochrony as an added constant in the comparison of durations (Schulze,

600   1989). What remains unclear is the need for such a bias in an otherwise quasi-

601   statistically optimal performance and the reason why there should be a different bias

602   in the two conditions presented here. The ML, IR, and DM models, on the other hand,

603   do not make explicit predictions that can account for the registered biases in perceived

604   isochrony. Such lack of an explanation calls for a novel model that can capture

605   perceptual distortions or response biases in isochrony.

606   **4.3 Temporal Uncertainty**

607   We would like to speculate on the reasons why sensitivity to temporal deviations is

608   lower in the interleaved condition, and we base our analysis on the observation that

609   the uncertainty about which interval should be judged changes depending on

610 condition and serial position. In the blocked condition, participants know exactly

611 when the sequence will end, whereas in the interleaved condition they do not, but the

612 uncertainty decreases as the sequence progresses. We can speculate that sensitivity to

613 temporal deviations increases with longer sequences in the interleaved condition

614 because later intervals have higher conditional probability to be the ones that need to

615 be judged (see Table 1). The hazard conditional probability for each successive

616 stimulus is related to temporal expectations (Nobre et al., 2007) and has been shown

617 to lead to better discrimination and faster reactions (Coull, 2009).

618   Here, we speculate whether such probability could be connected to the

619 consistent bias in response we find. In our results, isochrony is perceived when the

620 final interval in the sequence is, on average, 17 ms longer than the previous ones.

621 Such an effect is consistent with a positive time-order error (TOE; see Allan, 1979

622 and Woodrow, 1935 for a review) and a perceptual acceleration of the final stimulus,

623 an effect compatible with prior entry (Spence & Parise, 2010) and a recent study that

624 showed that intervals are perceptually shortened (accelerated) when below 3 seconds

625 (Wackermann, 2014). The fact that the duration of the last interval was

626 underestimated is particularly interesting if we consider that the intervals used in our

627 experiment are lower than the commonly used indifference point of 700 ms

628 (Woodrow, 1935). The effect size does not change across the sequence durations

629 tested, but we find that the delay required for perceived isochrony is 12 ms larger in

630 the interleaved condition than in the blocked presentation.

631   If this result is interpreted as an acceleration of the last stimulus, it should be

632 considered that the difference in hazard probability would suggest greater expectation

633 and, thus, more anticipation with longer sequences (Elithorn & Lawrence, 1955; Luce,

634 1986; Näätänen, 1970; Niemi & Näätänen, 1981;). Hazard probability alone, therefore,

635 does not explain why there should be a perceptual acceleration of the last stimulus in

636 the blocked condition, where no uncertainty about which stimulus to judge is present.

637 Our data, in fact, show more anticipation for the interleaved condition, where

638 intervals are actually more uncertain than in the blocked condition. Higher

639 predictability in the blocked condition, instead, should have led to a stronger prior

640 entry phenomenon.

641   Table 1. Probabilities associated with each of the interval in

642   the sequences in the interleaved condition (see also Coull,

643   2009).

644

	2nd	3rd	4th	5th
Probability of interval	1	3/4	2/4	1/4
Conditional probability of judgment	1/4	1/3	1/2	1

645 # 5. Conclusions

646 The present study first compared existing models of temporal sensitivity in

647 isochronous sequences before demonstrating how the length of a sequence and

648 interleaved presentation influence temporal judgments in isochronous sequences. Our

649 results show that discrimination sensitivity increases for longer sequences in

650 interleaved presentation and is overall better for blocked presentation. The pattern of

651 performance increase is consistent with the averaging of successive estimate, but with

652 a factor limiting performance. PA and ML models propose that either correlation

653 between successive estimates or weighting of the representation are the key factors.

654 Neither of the two exhaustively accounts for the pattern of performance increase

655 found. The results also evidence that perceived isochrony is obtained if the last

656 interval is longer than the previous one – i.e., with the last stimulus presented with a

delay between 10-20 ms – a finding that is consistent with a perceptual acceleration of the last stimulus in a sequence. The models analysed do not make explicit predictions for such a bias. Explanations based on stimulus probability could prove fruitful in counting for the difference in performance between the two conditions and the anticipation effect with blocked presentation of a sequence length as a higher task demand in the interleaved condition increases attentional deployment leading to stronger anticipation of the last stimulus.

7. Acknowledgments

8. References

Allan, L. G. (1979). The perception of time. *Percept Psychophys*, *26*(5), 340-354.

Arnal, L. H., & Giraud, A. L. (2012). Cortical oscillations and sensory predictions. *Trends Cogn Sci*, *16*(7), 390-398.

Brochard, R., Tassin, M., & Zagar, D. (2013). Got rhythm... for better and for worse. Cross-modal effects of auditory rhythm on visual word recognition. *Cognition*, *127*(2), 214-219.

Coull, J. T. (2009). Neural substrates of mounting temporal expectation. *PLoS Biol*, *7*(8), e1000166.

Cravo, A. M., Rohenkohl, G., Wyart, V., & Nobre, A. C. (2013). Temporal expectation enhances contrast sensitivity by phase entrainment of low-frequency oscillations in visual cortex. *J Neurosci*, *33*(9), 4002-4010.

Drake, C., & Botte, M. C. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. *Percept Psychophys*, *54*(3), 277-286.

Durlach, N. I., & Braida, L. D. (1969). Intensity perception. I. Preliminary theory of intensity resolution. *J Acoust Soc Am, 46*(2B), 372-383.

Dyjas, O., Bausenhart, K. M., & Ulrich, R. (2012). Trial-by-trial updating of an internal reference in discrimination tasks: Evidence from effects of stimulus order and trial sequence. *Atten Percept Psychophys*, *74*(8), 1819-1841.

Elithorn, A., & Lawrence, C. (1955). Central inhibition-some refractory observations. *Q J Exp Psychol*, *7*(3), 116-127.

Escoffier, N., Sheng, D. Y. J., & Schirmer, A. (2010). Unattended musical beats enhance visual processing. *Acta Psychol*, *135*(1), 12-16.

Grondin, S. (2001). From physical time to the first and second moments of psychological time. *Psychol Bull*, *127*(1), 22.

692   Grondin, S., & McAuley, J. D. (2009). Duration discrimination in crossmodal

693   sequences. *Percept, 38*(10), 1542.

694   Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. *Ann N Y*

695   *Acad Sci*, *423*(1), 52-77.

696   Helson, H. (1947). Adaptation-level as frame of reference for prediction of

697   psychophysical data. *Am J Psychol*, 1-29.

698   Helson, H. (1964). *Adaptation-level theory*. Oxford, England: Harper & Row.

699   Horr, N. K., & Di Luca, M. (2015). Filling the blanks in temporal intervals: The type

700   of filling influences perceived duration and discrimination performance. *Front*

701   *Psychol*, *6*, 114.

702   Halpern, A. R., & Darwin, C. J. (1982). Duration discrimination in a series of

703   rhythmic events. *Percept Psychophys*, *31*(1), 86-89.

704   Henry, M. J., & Herrmann, B. (2014). Low-frequency neural oscillations support

705   dynamic attending in temporal context. *Timing Time Percept*, *2*(1), 62-86.

706   Hirsh, I. J., Monahan, C. B., Grant, K. W., & Singh, P. G. (1990). Studies in auditory

707   timing: 1. Simple patterns. *Percept Psychophys*, *47*(3), 215-226.

708   Ivry, R. B., & Hazeltine, R. E. (1995). Perception and production of temporal

709   intervals across a range of durations: Evidence for a common timing

710   mechanism. *J Exp Psychol Hum Percept Perform*, *21*(1), 3.

711   Luce, P. A. (1986). A computational analysis of uniqueness points in auditory word

712   recognition. *Percept Psychophys*, *39*(3), 155-158.

713   Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track

714   time-varying events. *Psychol Rev*, *106*(1), 119.

715   Lapid, E., Ulrich, R., & Rammsayer, T. (2008). On estimating the difference limen in

716   duration discrimination tasks: A comparison of the 2AFC and the reminder

717   task. *Percept Psychophys*, *70*(2), 291-305.

718   McAuley, J. D., & Jones, M. R. (2003). Modeling effects of rhythmic context on

719   perceived duration: A comparison of interval and entrainment approaches to

720   short-interval timing. *J Exp Psychol Hum Percept Perform*, *29*(6), 1102.

721   McAuley, J. D., & Kidd, G. R. (1998). Effect of deviations from temporal

722   expectations on tempo discrimination of isochronous tone sequences. *J Exp*

723   *Psychol Hum Percept Perform*, *24*(6), 1786.

724   Miller, N. S., & McAuley, J. D. (2005). Tempo sensitivity in isochronous tone

725   sequences: The multiple-look model revisited. *Percept Psychophys*, *67*(7),

726   1150-1160.

727   McDonald, J. J., Teder-Sälejärvi, W. A., Di Russo, F., & Hillyard, S. A. (2005).

728   Neural basis of auditory-induced shifts in visual time-order perception. *Nat*

729   *Neurosci*, *8*(9), 1197-1202.

730   Näätänen, R. (1970). The diminishing time-uncertainty with the lapse of time after the

731   warning signal in reaction-time experiments with varying fore-periods. *Acta*

732   *Psychol*, *34*, 399-419.

733   Niemi, P., & Näätänen, R. (1981). Foreperiod and simple reaction time. *Psychol Bull*,

734   *89*(1), 133.

735   Nobre, A. C., Correa, A., & Coull, J. T. (2007). The hazards of time. *Curr Opin*

736   *Neurobiol*, *17*(4), 465-470.

737   Oruç, I., Maloney, L. T., & Landy, M. S. (2003). Weighted linear cue combination

738   with possibly correlated error. *Vision Res*, *43*(23), 2451-2468.

739   Pariyadath, V., & Eagleman, D. (2007). The effect of predictability on subjective

740   duration. *PLoS One*, *2*(11), e1264.

741   Pecenka, N., & Keller, P. E. (2011). The role of temporal prediction abilities in

742   interpersonal sensorimotor synchronization. *Exp Brain Res, 211*(3-4), 505-

743   515.

744   Roberts, S. W. (1959). Control chart tests based on geometric moving averages.

745   *Technometrics*, *1*(3), 239-250.

746   Rohenkohl, G., & Nobre, A. C. (2011). Alpha oscillations related to anticipatory

747   attention follow temporal expectations. *J Neurosci*, *31*(40), 14076-14084.

748   Rose, D., & Summers, J. (1995). Duration illusions in a train of visual stimuli.

749   *Percept, 24(10)*, 1177-1177.

750   Schulze, H. H. (1978). The detectability of local and global displacements in regular

751   rhythmic patterns. *Psychol Res*, *40*(2), 173-181.

752   Schulze, H. H. (1989). The perception of temporal deviations in isochronic patterns.

753   *Percept Psychophys*, *45*(4), 291-296.

754   Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition.

755   *Trends Cogn Sci*, *13*(9), 403-409.

756   Sternberg, S., & Knoll, R. L. (1973). The perception of temporal order: Fundamental

757   issues and a general model. In S. Kornblum (Ed.), *Attention and performance*

758   (pp. 629-685). New York: Academic Press.

759   Spence, C., & Parise, C. (2010). Prior-entry: A review. *Conscious Cogn*, *19*(1), 364-

760   379.

761   Shore, D. I., Spence, C., & Klein, R. M. (2001). Visual prior entry. *Psychol Sci*, *12*(3),

762   205-212.

763   Titchener, E. B. (1908). *Lectures on the elementary psychology of feeling and*

764   *attention*. New York: Macmillan.

765   Treisman, M. (1963). Temporal discrimination and the indifference interval:

766   Implications for a model of the "internal clock". *Psychol Monogr Gen App*,

767   *77*(13), 1.

768   ten Oever, S., Schroeder, C. E., Poeppel, D., van Atteveldt, N., & Zion-Golumbic, E.

769   (2014). Rhythmicity and cross-modal temporal cues facilitate detection.

770   *Neuropsychologia*, *63*, 43-50.

771   ten Hoopen, G., Van Den Berg, S., Memelink, J., Bocanegra, B., & Boon, R. (2011).

772   Multiple-look effects on temporal discrimination within sound sequences. *Att

773   Percept Psychophys*, *73*(7), 2249-2269.

774   Vibell, J., Klinge, C., Zampini, M., Spence, C., & Nobre, A. C. (2007). Temporal

775   order is coded temporally in the brain: Early event-related potential latency

776   shifts underlying prior entry in a cross-modal temporal order judgment task. *J

777   Cogn Neurosci*, *19*(1), 109-120.

778   Woodrow, H. (1935). The effect of practice upon time-order errors in the comparison

779   of temporal intervals. *Psychol Rev*, *42*(2), 127.

780   Wackermann, J., Pacer, J., & Wittmann, M (2014). Perception of acoustically

781   presented time series with varied intervals. *Acta Psychol*, *147*, 105-110.

782   Wolpert, D. M., & Ghahramani. (2000). Computational principles of movement

783   neuroscience. *Nat Neuroscie*, 3(Suppl),1212–1217.

784   Zampini, M., Shore, D. I., & Spence, C. (2005). Audiovisual prior entry. *Neurosci lett*,

785   *381*(3), 217-222.