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.

Keywords: Temporal perception, rhythm perception, temporal expectation, attention, isochrony, prior entry

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, the second aim of the paper is to see if there is a distortion from veridical perception – that is – if isochronous stimuli in a sequence are perceived as being on time, or whether they are perceptually accelerated, or delayed. The existing accounts of temporal sensitivity in isochronous sequences can only account for this type of changes in perceived isochrony by appealing to a response bias (an imbalance in the probability of the two responses), which has no perceptual origin. Such a finding would open the road to models that are able to capture biases in perceived timing of stimuli in isochronous sequences.

1.1 Percept Averaging (PA) Model Description

Schulze (1989) proposed to frame the problem of detecting whether the final duration in a sequence of intervals is deviant as discrimination between the duration of the N^{th} interval from the average of the percept of the previous N-1 intervals. Here we term this approach Percept-Averaging (PA) model, which assumes that all intervals are stored in memory and the perceptual system is capable of averaging them in a statistically optimal fashion, thus increasing the precision of the average (Schulze, 1989).

First of all, we will consider a simple case, where all N intervals in the sequence are independently estimated. If each estimate of the duration of an interval E is affected by independent Gaussian noise with average $\mu=0$ and variance σ^2, then the average of N-1 estimates has variance equal to $V\left(\frac{1}{N-1}\sum_{i=1}^{N-1} E_i\right) = \frac{(N-1)\,\sigma^2}{(N-1)^2} = \frac{\sigma^2}{(N-1)}$. The predicted just-noticeable difference (JND') with a sequence of N intervals of which the last could be deviant is expressed by $JND_N{}' = \sqrt{\frac{\sigma^2}{(N-1)} + \sigma^2} = \sqrt{\frac{N\sigma^2}{(N-1)}}$. Using this formula we find that the JND' predicted with a sequence of 2 intervals is $JND_2{}' = \sqrt{2}\sigma$. We can then express the predicted $JND_N{}'$ of a sequence with N intervals where the change in tempo happens at the last interval as a function of the empirical JND_2 of a sequence with 2 stimuli by integrating the two formulas as such:

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

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

The results of Schulze (1989) suggest that the improvement in performance with interleaved presentation of different sequence durations in a block is *higher* than the one predicted by this formula. Schulze speculated about the possibility that participants learned the duration of intervals throughout the experiments rather than within a single sequence. He also investigated whether this discrepancy could be due to the correlation in the noise of the duration estimated of successive intervals. A correlation in this instance means that an error made on the estimate of one interval influences also the estimates of the neighbouring ones. With coefficient of correlation r between successive intervals (and 0 otherwise) the average of $N-1$ estimates has 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 sequence of N intervals where the last could be deviant can be, thus, expressed by $JND_N{}' = \sigma\sqrt{\frac{N}{N-1} - \frac{2r}{N^2}}$. The reader should note that this expression differs from the third equation on page 294 in Schulze (1989), as we believe that the mathematical derivation leads to a second term that should be negative, not positive. Since the JND' predicted with a sequence of 2 intervals is $JND_2{}' = \sigma\sqrt{(2-2r)}$, then (similarly to Eq. 1) we can express the $JND_N{}'$ as a function of the empirical JND_2 and r as such

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

The patterns that can be obtained with this formula as a function of r are shown in Figure 1.

Schulze proposed that the non-correlated formulation did not capture the results as well as the negatively correlated formulation, especially in the interleaved condition (Schulze, 1989). However, the value of coefficient of correlation, *r*, was not determined in the original manuscript. Also, Schulze did not analyse the case where noise in successive samples could be positively correlated (such cases could be due to protracted variation of attention whose duration spans multiple stimuli), giving rise to a lesser improvement in performance as a function of sequence duration. We instead perform this analysis and evaluate the predictions of the model with different correlation (Figure 1). With these quantitative predictions, we will be able to compare the predictions of all models with the empirical data.

1.2 Multiple Look (ML) Model Description

Drake and Botte (1993) investigated participants' ability to judge the difference in tempo that happened not at the end of the sequence as in Schulze (1989), but in the middle of the sequence. The change in tempo, thus, creates two isochronous sequences with different rhythms. The authors focused the analysis on the presence of multiple estimates of interval duration, and for this they coined the name Multiple-Look model (ML). The model posits that the precision of the estimate improves as the number of 'looks' at each sequence increases. The ML model has a formulation that is consistent to the model proposed by Schulze's (1989) with uncorrelated noise, where the multiple estimates of the intervals are stored in memory and their average is compared. Here, we will show how to derive the expression of the ML model following the logic of Schulze's (1989) demonstrating their mathematical equivalence. In the task of judging a tempo change in the middle of the sequence, participants perform the discrimination by comparing the average of the duration of the first $N/2$ intervals to the average of the second set of N/2 intervals. The noise in the estimate of 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* intervals, where the change in tempo happens in the middle of the sequence is $JND_N{}' = JND_2\sqrt{\frac{2}{N}\sigma^2}$ and by expressing it as a function of the empirical JND_2 we obtain

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

Miller and McAuley (2005) suggested a generalized ML model, whereby the two sequences (denoted n_1 and n_2, respectively, so that $N=n_1+n_2$) make independent contributions to the performance. Again, in Schulze's (1989) framework participants compare the average of the n_1 intervals to the average of the n_2 intervals, with a *JND'* 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 obtain:

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

It should be noted that this is a more general expression of the previous two formulations when noise is considered uncorrelated, so that with $n_2=1$ the formula is identical to Eq. 1 and with $n_1=n_2$ the formula is identical to Eq. 3.

The model of Miller and McAuley (2005) slightly departs from this formulation. Eq. 4, predicts that the JND_{n1+n2} should decrease as the number of 'looks' increases for either of the two intervals. For Miller and McAuley, instead, the contribution of the two sequences is allowed to vary depending on a weight parameter, *w* as such:

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

According to Miller and McAuley, the parameter *w* modulates the contribution of the two averaged estimates. If *w = 1* then the discrimination performance would be determined only by average of the first series of intervals, whereas if *w* = 0 then the *JND* would be determined by average of the second series of intervals. Such parameter cannot be reconciled with the functioning of the model proposed by Schulze (1989), as both averages are required to perform the discrimination and are, thus, influencing the performance.

If the general ML model expressed by Eq. 5 is instantiated for the case analysed by Schulze (1989) where the change in tempo happens at the last stimulus (*n1=N-1* and *n2=1*) the formula becomes

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

In the generalized ML model (Eq. 6), the weight parameter *w* ranges between 0 and 1 and describes how much reliance a participant has on the first of two sequences to be compared. The patterns of performance vary according to this value as shown in Figure 1. The model is based on the presence of a memory store to which future intervals are compared (Treisman, 1963). After comparison, the memory store is updated integrating every presentation of intervals, i.e., to form an internal reference (see Dyjas et al., 2012). In the formula, the weight *w* captures the proportion (across trials) in which the participant stores a combined memory trace of all previously presented intervals. With *w* = 1, the store is used in a statistically optimal fashion, combining information from all the preceding intervals. In this case, the JND_N' is determined by the limited precision of the comparison of the last interval with such a memory trace. With *w* = 0, instead, the store does not integrate information across intervals, thus it only contains a representation of the latest interval presented. Performance reflected by JND_N' with *w* = 0 is, thus, determined by the precision in comparing the last interval in a sequence with the previous one, regardless of how many preceding intervals there are.

The goal of the ML Model is to quantify the discrimination performance with two sequences of regular intervals. With this task, several studies have reported results consistent with the ML model (Grondin, 2001; Ivry & Hazeltine, 1995; McAuley & Jones, 2003; McAuley & Kidd, 1998; ten Hoopen, et al., 2011), although others have not found a close match with its predictions (Grondin, 2001; Hirsch et al., 1990; ten Hoopen et al., 2011). Furthermore, Grondin (2001) demonstrated a ML effect with visual stimuli only if tempo was compared in two separate sequences, whereas the effect was not present if a change in tempo happened within one sequence. Ivry and Hazeltine (1995) also compared one sequence performance with performance in two sequences, but with audio stimuli, finding a ML effect in both.

1.3 Internal Reference (IR) Model Description

The models examined so far are based on averaging the duration estimates of multiple intervals and comparing this value a final duration estimate. Such a process requires the storage in memory of all the estimates of all intervals to obtain a statistically optimal average. However, a more efficient alternative formulation is to compute the average iteratively each time a new estimate becomes available. As per the IR model, such a procedure can be performed using a recursive estimator, like the Kalman filter. The mean with *N=n+1* estimates is a weighted average of the mean μ_n of the previous *n* estimates and of the last estimate E_{n+1}, which can be expressed as

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

where $K = \frac{1}{(n+1)}$ is called the gain factor and indicates how the weight given to the single *E* value decreases with longer sequence. This idea is similar to the concept of a clock model in time perception (Gibbon et al., 1984; Treisman, 1963), where the representation of duration increases in precision by averaging the representation of successive estimates of intervals, thus leading to better performance (Dyjas et al., 2012; Schulze, 1979). If estimates are independent, this formula leads to the same variance decrease obtained by averaging all stimuli at once expressed by Eq. 1. On the positive side, however, this way of computing the average reduces the memory requirements to only a single estimate value at the time (plus the knowledge of how many stimuli have been averaged) albeit it increases the complexity of the computation, because a weighed average is required for each iteration. The iterative process, however, does not lead to statistically optimal variance reduction with positively correlated noise estimates.

An alternative to this scheme has been proposed by Dyjas et al. (2012), originally to account for serial effects in tasks requiring the comparison of two durations. The authors propose that weights are different from the statistically optimal *K* and do not depend on the sequence length. Instead, they propose a weight g for modulation of the current estimate and the contribution of the previous reference:

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

Such a scheme leads to a geometric moving average (Roberts 1959), where the weight g assigned to the historical list of estimates decreases as a geometric progression when time passes. The variance associated with such averaging method is (see Dyjas et al., 2012) $V(average) = \frac{s^2(g^{2n}+(1-g)^2(1-g)^{2n})}{1-g^2}$. As the participant would be comparing this average to the last interval, the predicted *JND'* for a sequence of *N* 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 sequence of only two intervals, the $JND_2{}'$ would be $JND_2{}' = \sqrt{s^2+s^2(g^2+(1-g)^2)}$. Performing the substitution of $JND_2{}'$ in $JND_N{}'$ gives

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

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

Predictions of the IR model expressed in Eq. 9 are shown in Figure 1 for different values of g. It is immediately evident that such a formulation cannot predict the same improvement and decrease in performance as the other proposals derived from Schulze (1989).

1.4 Diminishing Returns (DR) function

Ten Hoopen et al. (2011) investigated the issue of temporal sensitivity in a single sequence of audio stimuli where the change in tempo could happen at one of several positions. They found that performance changed more as a function of the number of intervals *before* the tempo change, rather than after. They adopted a reciprocal DR function to capture the performance increase:

$$\mathrm{JND}_{\mathrm{n_1:\,n_2}} = \mathrm{a} + \frac{\mathrm{b_1}}{\mathrm{n_1}} + \frac{\mathrm{b_2}}{\mathrm{n_2}} \qquad \text{Eq. (10).}$$

where *a* is the asymptotic performance and *b* are the amount of performance increase for each added interval before and after the tempo change. The parameters fitting the results of Ten Hoopen et al. highlight that performance increment is higher for changes before the tempo change are captured by $b_1>b_2$. It should be noted that the DR function expressed in Eq. 10 is not based on a process oriented model as the one proposed for example by Schulze (1989), because purpose was to fit the data. With this specification, in the rest of the manuscript we will refer to the DR as a model rather than a function. Eq. 10 can nevertheless be used to express the *JND* of a sequence of intervals where the last one is deviant as a function of the *JND* obtained in a sequence with two intervals. If we define *c* as the combined factor $c = a + b_2$

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

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

The ability of the DR model expressed in Eq. 11 to capture an improvement in performance in our empirical study can be analysed by looking at the range of possible fittings in Figure 1 (i.e., the change in the predictions of the DR as a function of the *c* parameter).

1.5 Experimental question

The models analysed so far (PA, ML, IR, DR) all make predictions that discrimination performance improves as the number of intervals to be examined increased. There are, however, quantitative differences in the predictions by Schulze's (1989) PA model (Eq. 1 and Eq. 2), the ML model (Eq. 6), the IR model (Eq. 9), and the DR model (Eq. 11). In this paper, we hope to be able to determine which model captures the data of two experimental conditions (interleaved and blocked presentation of duration) using the free parameter that each model has (respectively: correlation *r*, weight *w*, gain factor *g*, and combined factor *c*).

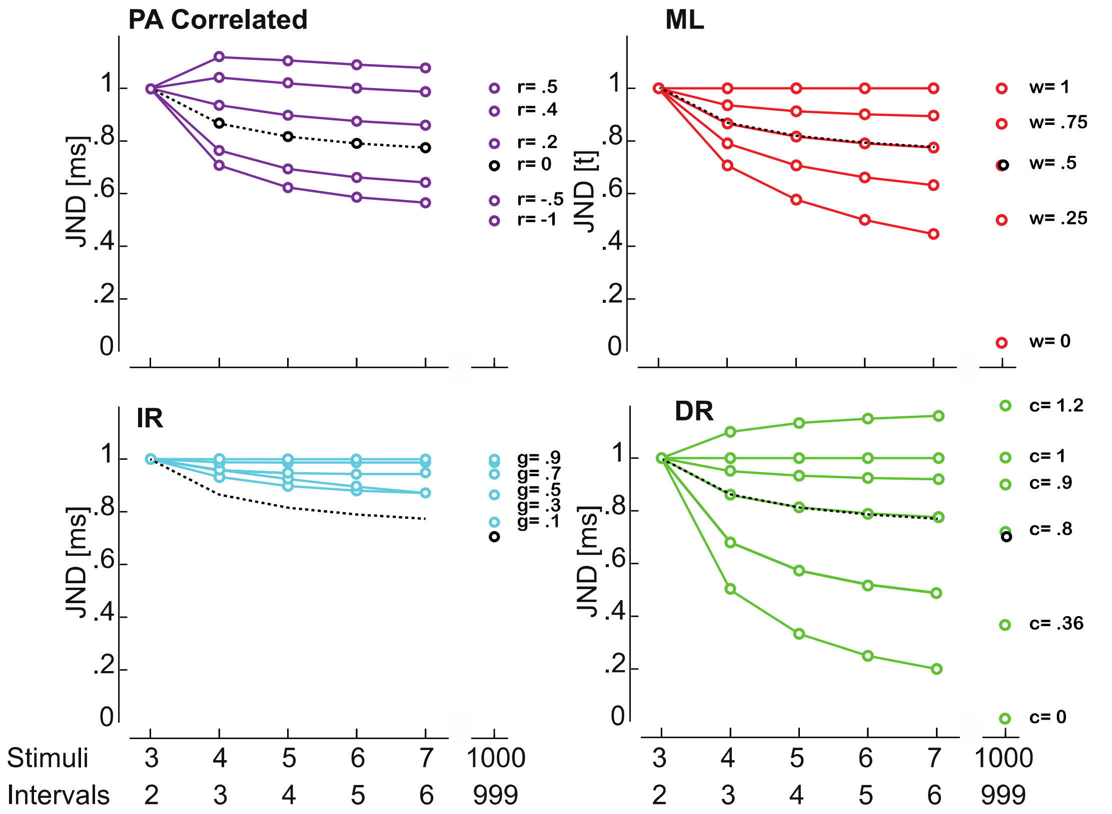

Figure 1. Predictions for the Percept Averaging (PA, Eq. 1 and 2), Multiple Look (ML, Eq. 6), Internal Reference (IR, Eq. 9), and Diminishing Return (DR, Eq. 11) models for JND_N with a sequence of N stimuli expressed as a function of JND_2=1ms. Each model has a single free parameter that has been varied 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 task is similar to ten Hoopen et al.'s (2011), as participants are asked to judge whether the last stimulus was presented 'earlier' or 'later' than isochrony (i.e., they reported whether the last interval was shorter or longer than the previous ones). The analysis of 'earlier' vs. 'later' judgments allows us to determine whether temporal expectations generated by the sequence of stimuli with identical interval can cause a consistent bias in perceived isochrony, an analysis that was possible but has not been performed by ten Hoopen et al. The motivation for this new analysis is to try to account for any consistent bias in responses with a perceptual mechanism. In particular, a bias in perceived isochrony can be explained by appealing to a modification of the perceived timing of the last stimulus in the sequence. This possibility requires a difference in the formulation of the problem of perceived isochrony as has been done so far: rather than considering the perceptive duration of the individual interval, here we propose to analyse the perceived timing of stimuli. In particular, we analyse the time at which the last stimulus in the sequence is perceived, which is presented right after the change in tempo. Perceived timing of stimuli can be affected by several factors in a way independent from perceived duration.

Titchener (1908) was the first to suggest that attention (among other factors) can modulate perceived timing of individual stimuli as a fully attended stimulus is processed faster than an unattended one. Summerfield and Egner (2009) investigated the contribution of attention in a recognition task supporting the idea of prioritized processing of attended stimuli. Such attentional facilitation speeds up perception, an effect termed *prior entry*, which has been highlighted in studies involving temporal judgments (Sternberg & Knoll, 1973; Shore et al., 2001; Vibell et al., 2007; Zampini et al., 2005; for a review see Spence & Parise, 2010) and at the neural level (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, 100, 150, and 200 ms. The trial types resulting from the combination of blocked/interleaved presentation (2), sequence length (4), and anisochrony of the last stimulus (15) were repeated 8 times in order to determine the parameters of eight psychometric functions (see results) for a total of 960 trials per participant.

2.1.3 Stimuli

Stimuli were identical tones produced by a speaker located on a desk approximately 50 cm from the participant (20 ms with 5 ms linear ramp, 1 kHz, 75.1 dBA). Trials were composed of a different number of stimuli within a sequence, where intervals between successive stimuli in the sequence remained the same (IOI = 700 ms) for all but the final stimulus, which could be presented at different anisochronies.

2.1.4 Procedure

Participants sat in a quiet testing cubicle. A sequence of auditory stimuli of different lengths were presented in which the participants had to respond whether the anisochrony of the final stimulus was ‘earlier’ or ‘later’ than the expected timing (Fig. 2). Sequence lengths were either presented blocked or interleaved and the order of the two presentations was counterbalanced across participants.

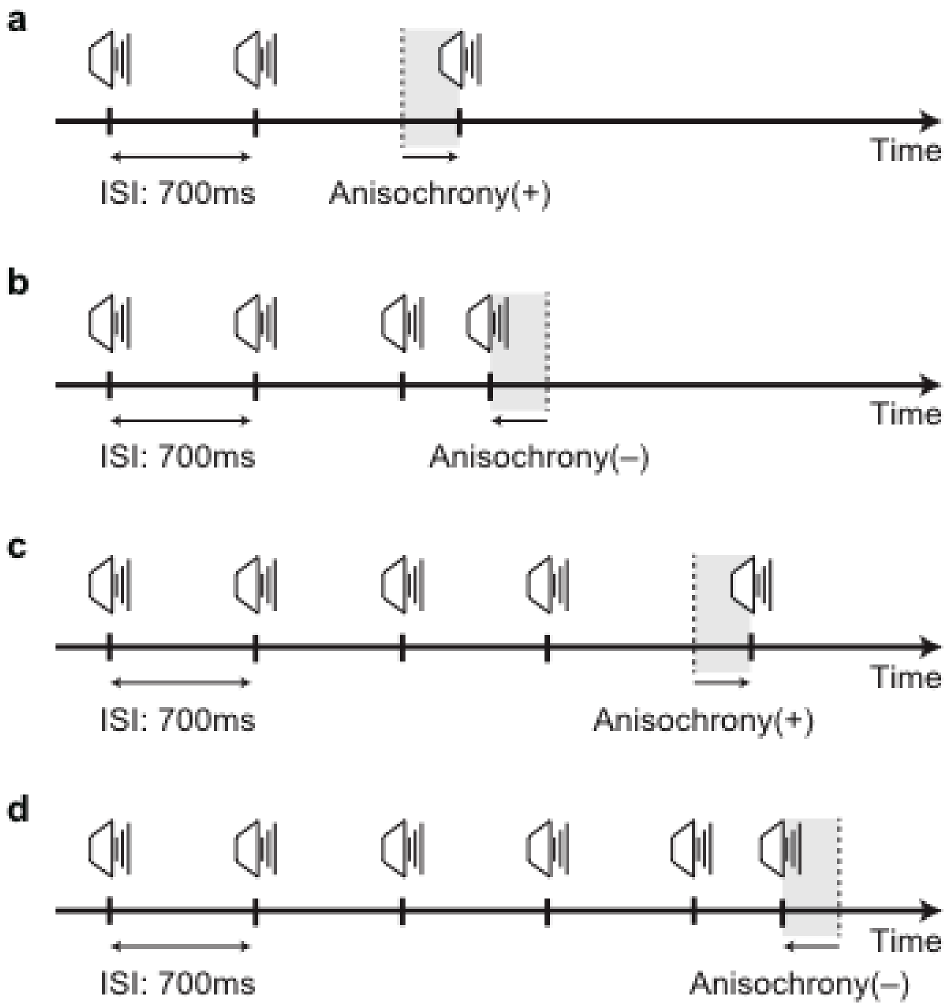

Figure 2. Examples of trials with different sequence length. (a) Sequence of three stimuli (two intervals) where the final stimulus is presented later than expected (+ Anisochrony). (b) Sequence of four stimuli (three intervals) where the final stimulus is presented earlier than expected (- Anisochrony). (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 and attention, we calculated the *Point of Subjective Equality* (*PSE*) as the anisochrony at which participants are most unsure about whether the final stimulus was presented early or late. Thus, the *PSE* is the time point the last stimulus needs to be presented in order for it to be perceived as being isochronous. The *PSE* is obtained by calculating the first order moment of the difference between successive proportions of responses using the Spearman-Kärber method (see Ulrich & Miller, 2004, for further details of this method). The second order moment is proportional to the inverse slope of the psychometric function, which here is termed *JND*.

To obtain *PSE* and *JND,* we employ the Spearman-Kärber method, which is a non-parametric estimate that avoids assumptions about the shape of the psychometric functions underlying the participants' responses. The formulae below are used to estimate the first and second moment of the psychometric function underlying the data. First we define the 15 anisochronies of the final stimulus, where ANI_i with i={1, ... 15} and p_i with i={1, ... 15} as the associated proportion of 'later' responses. 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 compute the intermediate *ANI* between two successive ones

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

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

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

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

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

and

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

2.2.2 Model Fitting

In order to find the best fit for the each of the model's parameter, for each participant we found the minimum sum of squares difference between the predicted $\mathrm{JND}_{\mathrm{N}}^{\prime}$ and the empirical $\mathrm{JND}_{\mathrm{N}}$. In Schulze's

PA model (Eq. 2) the minimisation is done with the correlation, in the generalized ML model (Eq. 6) with the weight, in the IR model (Eq. 9) with the gain factor, and the DR model (Eq. 11) with the combined factor. The fitting is done independently for the two conditions (blocked vs. interleaved).

3. Results

The average proportion of responses across participants for sequences of different lengths and type of presentation (interleaved and blocked) are shown in Fig. 3. A consistent difference in the shape of the response distributions with blocked and interleaved presentation is evident across the various sequence lengths.

Discrimination performance was characterised by *JND* values (Fig. 4), which are calculated according to the Spearman-Kärber method (see method section). The proportions of ‘late’ responses in each psychometric function were monotonized prior to analysis. To determine whether temporal sensitivity improves with sequence length and whether differences in sensitivity existed between blocked and interleaved presentations, *JND* values were submitted to a two-way repeated measure ANOVA with factors condition (blocked or interleaved) and number of intervals in the sequence (2, 3, 4 or 5). Results indicate better discrimination with blocked presentation of sequence length ($F(1,24)=20.3$, $p<0.001$, $\eta_p^2=0.46$, Fig. 3c), an 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 ($F(3,72)=4.1$, $p=0.009$, $\eta_p^2=0.38$). Such an interaction suggests that the improvement in temporal discrimination due to sequence length is present with the interleaved presentation of different sequence length (one-way repeated measure ANOVA with factor sequence length: $F(3,72)=5.1$, $p<0.003$, $\eta_p^2=0.18$) but performance is not affected with blocked presentation of one length ($F(3,72)=2.0$, $p=0.119$, $\eta_p^2=0.12$).

Biases in perceived isochrony are captured by *PSE* values (Fig. 5), which are also calculated according to the Spearman-Kärber method (see method section). In both conditions, we find that stimuli presented physically isochronous are actually reported more often to appear earlier than expected. Perceived isochrony is obtained when the last stimulus was presented later than it should – i.e., with a longer last interval (single sample t-test of *PSE* calculated on the data against 0: interleaved, $t(24)=6.1$, $p<0.001$, blocked: $t(24)=2.6$, $p=0.015$). In order to test whether there is a consistent difference of this effect with blocked or interleaved presentation of sequence lengths, we submitted *PSE* values a two-way repeated-measures ANOVA with factors presentation condition (interleaved or blocked) and number of interval in the sequence (2, 3, 4 or 5). Results indicate a change in *PSE* depending on the presentation condition ($F(1,24)=13.4$, $p=0.001$, $\eta_p^2=0.36$), as the final stimulus in the interleaved condition has to be presented 24.6 ms (4.0 ms SEM) after isochrony in order to be perceived isochronous, whereas the last stimulus in the blocked condition has to be presented 12.1 ms (4.6 ms SEM) after isochrony. The difference between both interleaved and blocked condition was 12.4 ms (4.5 ms SEM). We find no main effect of sequence length or an interaction (both $p > 0.11$).

In sum, the sensitivity of detecting anisochrony increases with longer sequences if different lengths are interleaved but is overall higher if only one sequence length is presented in a block. Perceived isochrony is consistently biased and the observed bias does not change due to sequence length, but it is affected by the presentation condition (interleaved and blocked). Not knowing the serial position of the interval to be judged leads to a higher bias, so that the sequence is perceived as being isochronous if the last stimulus is presented slightly later, i.e., after a longer interval compared to the previous ones.

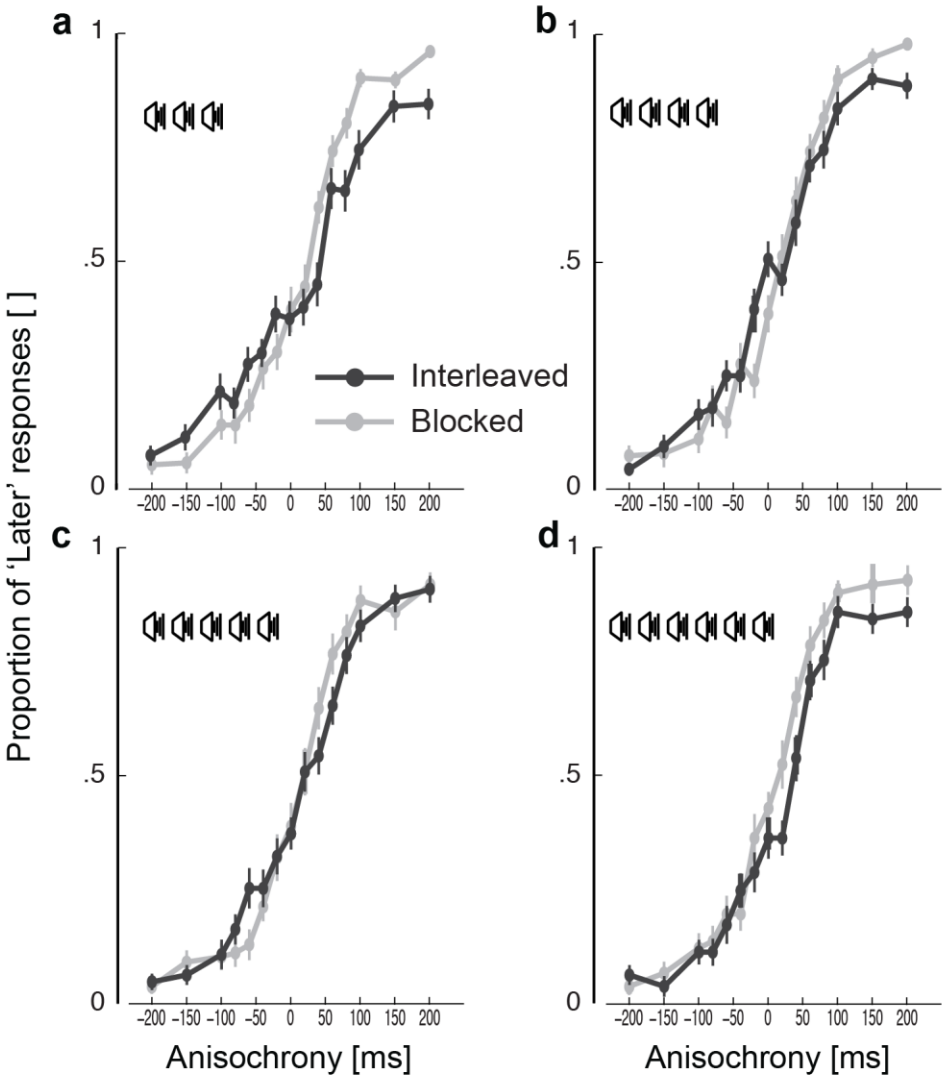

Figure 3. Proportion of 'later' responses as a function of the anisochrony of the final interval in the sequence for (a) 2, (b) 3, (c) 4, and (d) 5 intervals for interleaved and blocked presentation. Asterisks indicate significant difference between the two conditions according to the values in Table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

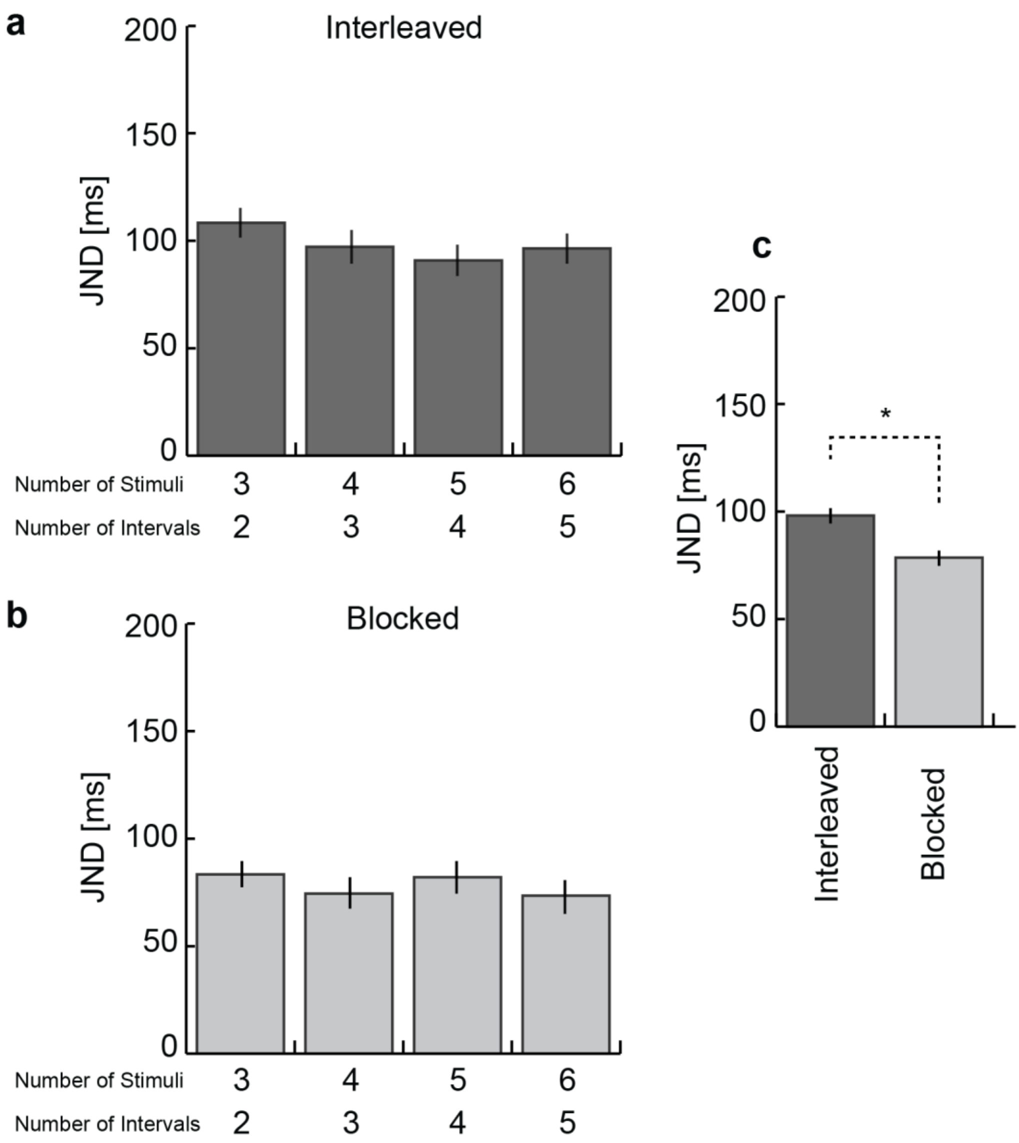

Figure 4. *JND* values as a function of sequence length for (a) interleaved and (b) blocked presentation. (c) *JND* values calculated on the proportion of ‘later’ responses across sequence lengths for blocked and interleaved conditions. The asterisk indicates a significant difference according to the ANOVA presented in the text. Error bars represent the standard error of the mean.

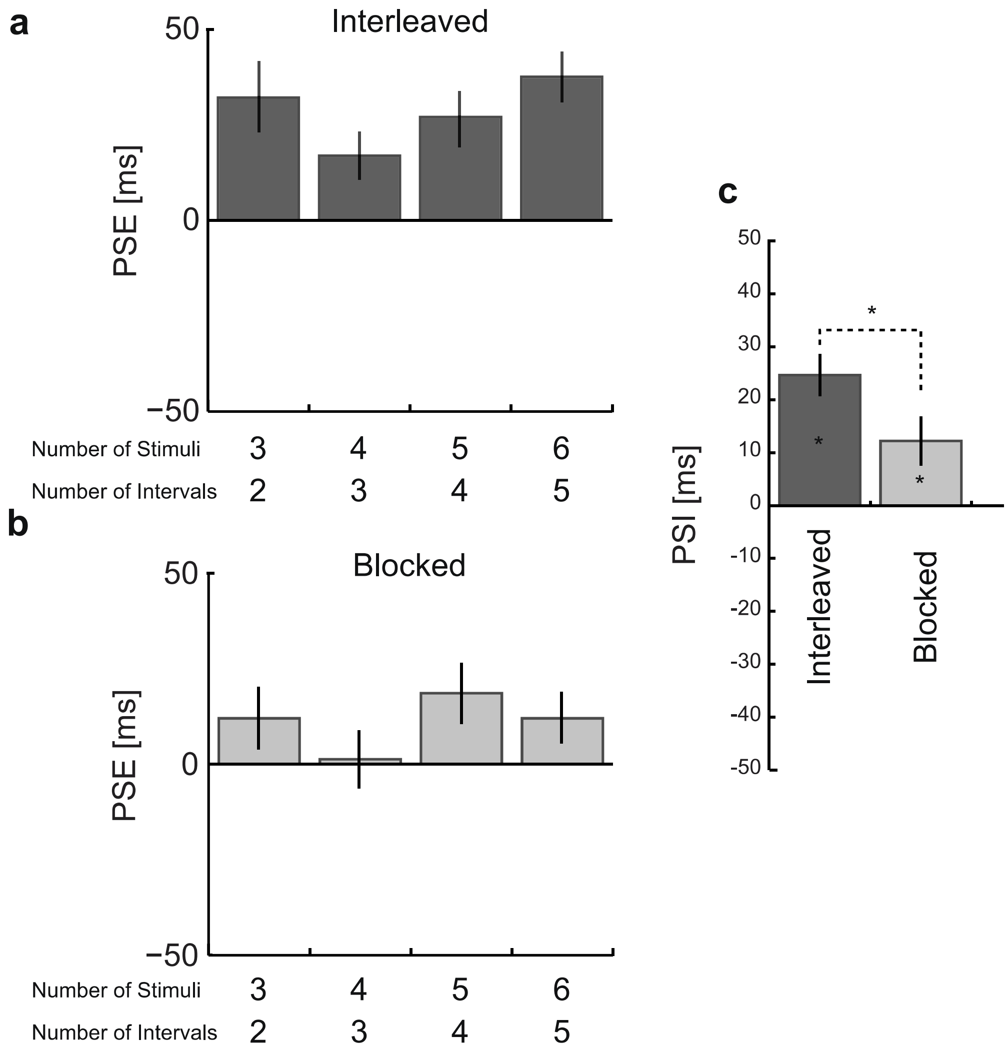

Figure 5. *PSE* values as a function of sequence length for (a) interleaved and (b) blocked presentation. (c) *PSE* values calculated on the proportion of 'later' responses across sequence length for interleaved and blocked presentation. The asterisk indicates a significant difference from 0 according 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 temporal sensitivity. We applied Eq. 1 to our data and (without any fitting procedure) it generally captures the decrease in the empirical JND_N in the interleaved condition and blocked condition (Fig. 6) with very similar sum of squares differences in the interleaved and blocked conditions, 1182±118 ms^2 and 1210±277 ms^2 respectively (t(24)=0.08, p=0.94; Fig. 7).

Extending the Schulze (1989) model to include correlated noise lead us to employ Eq. 2. We found the minimum sum of squared differences between the predicted $JND_N{}'$ and the empirical JND_N across the four durations for each participant through an exhaustive search of the value of correlation r. Predicted values that minimise such difference are shown in Figure 6. Such procedure will be employed for the following models and makes the models equivalent in terms of number of fitted parameters. The sums of squared differences for the PA Correlated model are 825±183 ms^2 and 587±115 ms^2 which, notably, are significantly lower than the values obtained with the unfitted PA Uncorrelated model (interleaved: t(24)=2.5, p=0.017; blocked: t(24)=5.3, p<0.001; Fig. 7). Despite this improvement, the average correlations that lead to the minimum sum of square difference for each participant in each condition are quite small -0.056±0.091 and -0.124±0.092 and do not differ from 0 (interleaved: t(24)=1.1, p=0.28; blocked: t(24)=1.4, p=0.18) nor differ from each other (t(24)=0.5, p=0.59).

3.2 ML Model Results

Like above, the ML model predicts that sensitivity to changes in tempo increases with longer sequences with a factor that limits performance compared to statistical optimality, the difference from 0.5 of the weight assigned to the two parts of the sequence (Drake & Botte, 1993; Miller & McAuley, 2005). Here we allowed individual participants' weights to span a range between -0.5 and 1.5 as noise between successive estimates can be correlated (see Schulze, 1989 and Oruç et al., 2003 for more detail). We performed the same sum of squared error minimization procedure as for the PA Correlated model. Predicted values of $JND_N{}'$ that minimise error are overlaid to the empirical values in Fig. 6. Average weights are 0.39±0.09 and 0.24±0.11 for the interleaved and blocked condition respectively, they differ from 0.5 (single sample t-test against 0.5, interleaved: $t(24)=2.6$, $p=0.014$; blocked: $t(24)=3.0$, $p=0.006$) but they do not differ significantly ($t(24)=1.1$, $p=0.26$). The model captures the increasing sensitivity in the interleaved condition slightly, but not significantly, worse than for the blocked condition – as the values of the average sum of squared differences for the ML model are 802±180 ms^2 and 579±119 ms^2 for the interleaved and blocked conditions respectively, do not differ significantly ($t(24)=1.0$, $p=0.32$; Fig. 7). The performance of the ML model in capturing the data is not significantly different than the PA Correlated model (t-test on average SSE across the two conditions between ML and PA $t(24)=1.0$, $p=0.30$).

3.3 IR Model Results

Slightly different from the averaging models stated above, the IR model proposed by Dyjas et al. (2012) can only capture a limited range of improvements in temporal discrimination (Fig. 4). The factor limiting performance is the weight of the current estimate *g*, which here was tuned with the same procedure followed above. The best-fitting weight *g* is 0.61±0.07 in the blocked and 0.66±0.05 in the interleaved condition, which do not differ significantly ($t(24)=0.5$, $p=0.65$). The sum of square difference for the IR model is 1000±180 for the interleaved condition and 778±162 for the blocked condition (Fig. 7). Such values are higher than the PA Correlated and MLM models (t-test on average SSE across the two conditions between IR and: PA $t(24)=3.7$, $p=0.0011$, 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 $p<0.001$).

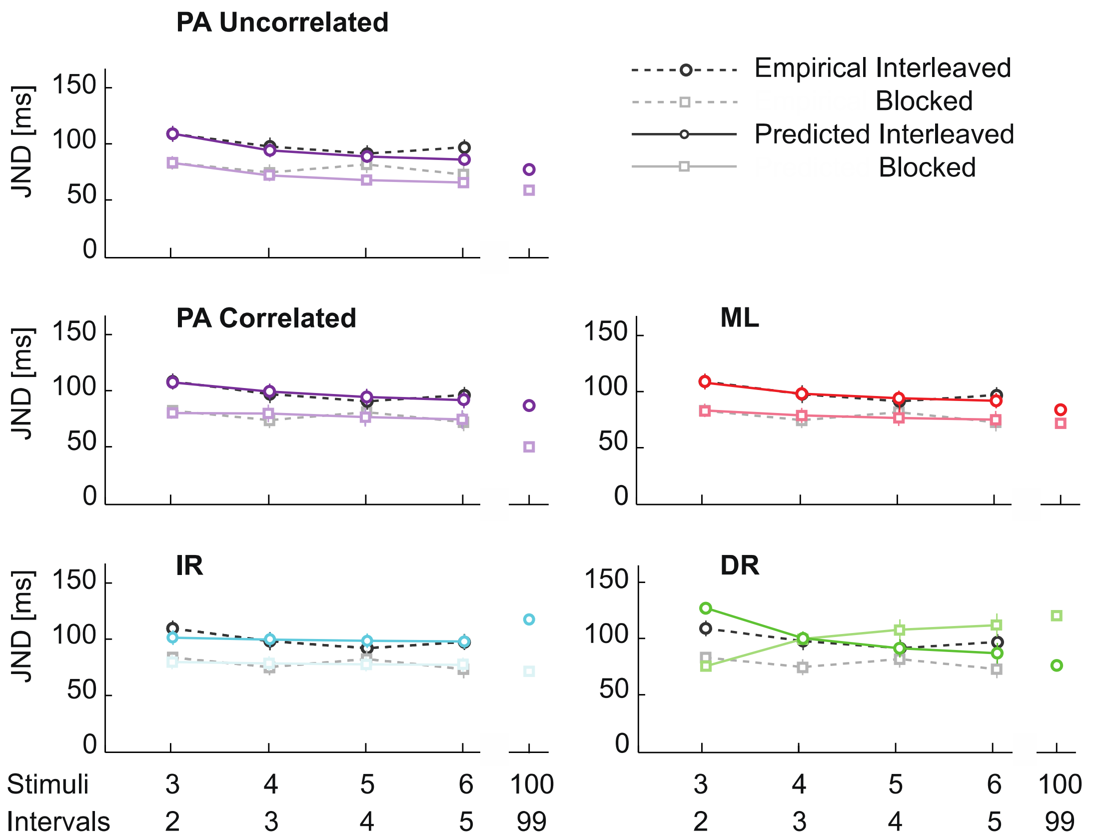

Figure 6. Predictions of the Percept Averaging (PA), Multiple Look (ML), Internal-Reference (IR), and Diminishing Returns (DR) model (see results section). The predictions of the PA (Schulze, 1978; 1989) and ML Models (Drake & Botte, 1993; Miller & McAuley, 2005) visually capture the increase in temporal sensitivity as a function of sequence length across the two conditions. The IR model (Dyjas et al., 2012) captures the flat course of JND for the blocked condition but cannot accurately capture the obvious increase in temporal sensitivity for the interleaved condition. The DR Model (ten Hoopen et al., 2011) captures the negatively accelerating course of the *JND* only for the interleaved condition but does not correctly account for flat course of *JND* in the blocked condition, as the fit for several participant predicts worse performance due to the presence of low-performance conditions.

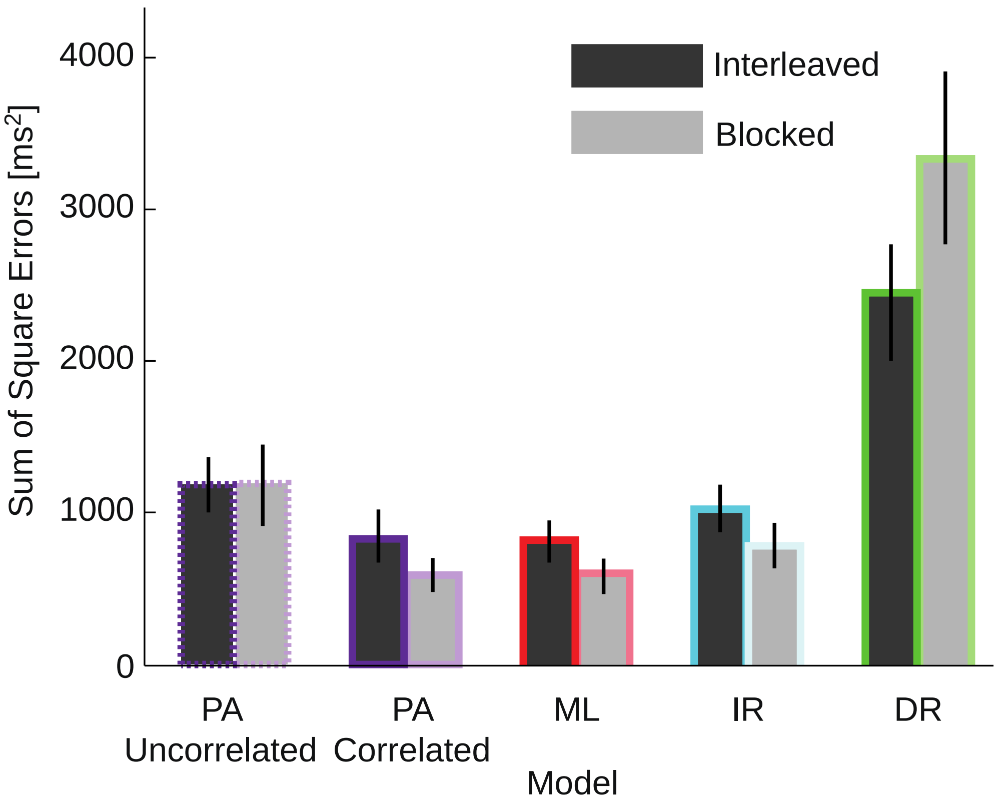

Figure 7. Comparison of the models fit to the empirical data captured by the sum of squared errors for the Percept Averaging (PA; Correlated and

Uncorrelated), Multiple Look (ML), Internal Reference (IR), and Diminishing Returns (DR) models. The dark grey bar represents the interleaved condition whilst the light grey indicates the blocked condition. A 2-way r.m. ANOVA on the data with factors models and interleaved/blocked is significant for the factor model ($F(4,96)=39.37$, $p<0.0001$, $\eta_p^2=0.62$) whereas the factor blocked/interleaved and interaction are not significant. Error bars represent the standard error of the mean across participants.

4. Discussion

In this paper, we aimed to compare the predictions of existing models of how the brain may deal with detecting deviations from isochrony in sequences of auditory tones. Second, we wanted to see if we could observe any distortions from veridical isochronous perception. To investigate this, similar to previous investigations (Halpern & Darwin, 1982; Hoopen et al., 2011; Schulze, 1978; 1989), we manipulated sequence length across trials (2, 3, 4 or 5 intervals in a sequence). The final interval in the sequence could be presented too early or too late, and participants needed to identify which of the two cases it was. By presenting the final stimulus either earlier or later as ten Hoopen et al. did, we could eliminate response biases that affected the measure of sensitivity. We also tested whether presenting the sequences either interleaved (difficult task as participants do not know the sequence length to be judged) or blocked (simpler task because participants know which interval could be deviant) has an impact on perception. Temporal discriminability (quantified by the *JND* calculated on the proportion of 'later' than expected responses) is found to be higher in the blocked condition than in the interleaved condition. Furthermore, we find that temporal sensitivity increases as a function of sequence length in the interleaved condition, but not in the blocked condition (Fig. 4a,b). This principal finding will now be reviewed in the context of the models of temporal deviation detection.

4.1 Model Comparison

The goal of the paper was to compare existing approaches to how the brain may deal with temporally deviant stimuli. As such, the finding that temporal sensitivity increases as a function of sequence length in the interleaved condition is consistent with the findings of Schulze (1989) and ten Hoopen et al. (2011). However, Schulze found a larger increase in performance with longer sequences than we report here and, thus, it is possible that such a difference could be due to the use of final intervals that could only be longer than the previous ones. The best fit of the predicted JND_N' to the empirical data JND_N was with the PA and MLmodels. The PA model without correlated noise predicted a too large improvement in performance in the blocked condition, but having the correlated noise included in the formulation, the PA model accurately captured the patterns of both conditions. The ML model finely captured the steeper slope of increased temporal sensitivity in the interleaved condition, and the limited improvement of blocked condition performances as well. On the other side, although the IR model was not able to capture the close-to statistically optimal improvement of temporal sensitivity in the interleaved condition, it instead accurately captured the flat course that was observed in the blocked condition. Of all the models we have implemented, the DR model was a relatively demanding fit, as it predicted an increased pattern of *JND* that we did not find in our averaged blocked condition results. The DR model also over-estimated the improvement of temporal sensitivity in the interleaved condition.

The parameters used to fit the models to the data are also interesting. Despite the increase in performance from the PA Correlated compared to the PA Uncorrelated, the correlation parameter *r* does not significantly vary across conditions nor statistically differs from 0, although there is a slight tendency to negativity as expected by Schulze (1989). Such results leads us to think that beyond the limiting performance increase due to the overall negative weight, the reason for better fit needs to be searched in inter-individual level, i.e., in the different pattern of performance increase for different sequence duration. The fit of the ML to the data is somewhat consistent with this view. Overall, the deviation of the weight from 0.5 suggests a limitation in the performance increase. However, the lack of a statistical difference in the weight depending on the conditions points at an inconsistency across participants.

The three interval-based models described here (PA, ML, IR) have a common explanation for the increase in sensitivity to temporal properties with longer sequences due to the increase in precision of the duration representation following exposure to multiple intervals (i.e., Dyjas et al., 2012; Schulze, 1979). Such improvement is consistent with internal clock models (Gibbon et al., 1984; Treisman, 1963), where duration is judged as the accumulation of 'ticks' from an internal pacemaker. The fact that the fit of the PA model fails to find a difference in correlation and that the ML model fails to find a difference in the weight assigned to the intervals with blocked and interleaved presentation suggest that the integration of information is not complete and, thus, sub-optimal. The result that there is no change in correlation and in weighting is logical, as sensory correlation and memory integration should not be affected by whether the sequence is presented interleaved with other sequence lengths.

To further compare the models, we generated predictions for a sequence of 100 stimuli (Fig. 1). We find that the models largely differ in their predicted performance. The ML expressed by Eq. 4 should lead to a progressive increase in performance as the sequence increases in length. A similar situation is present for the DR model. In comparison, the Correlated PA of Eq. 2 has a parameter that limits the integration of memory traces (Schulze, 1978, 1989). The IR model has also a hard stop in the performance and cannot go beyond statistical optimality with uncorrelated noise. Thus, the ML and DR models are unable to capture the asymptotic maximal performance with long sequences as they predict impossibly high performance.

4.2 Response Bias

A second aspect that our experiment allowed us to ascertain was the presence of a consistent bias in the reported isochrony, registered as consistent deviations of *PSE* from 0 in Fig. 5. Such bias changed depending on the interleaved/blocked presentation of durations. The PA model could, in principle, capture biases in perceived isochrony as an added constant in the comparison of durations (Schulze, 1989). What remains unclear is the need for such a bias in an otherwise quasi-statistically optimal performance and the reason why there should be a different bias in the two conditions presented here. The ML, IR, and DM models, on the other hand, do not make explicit predictions that can account for the registered biases in perceived isochrony. Such lack of an explanation calls for a novel model that can capture perceptual distortions or response biases in isochrony.

4.3 Temporal Uncertainty

We would like to speculate on the reasons why sensitivity to temporal deviations is lower in the interleaved condition, and we base our analysis on the observation that the uncertainty about which interval should be judged changes depending on condition and serial position. In the blocked condition, participants know exactly when the sequence will end, whereas in the interleaved condition they do not, but the uncertainty decreases as the sequence progresses. We can speculate that sensitivity to temporal deviations increases with longer sequences in the interleaved condition because later intervals have higher conditional probability to be the ones that need to be judged (see Table 1). The hazard conditional probability for each successive stimulus is related to temporal expectations (Nobre et al., 2007) and has been shown to lead to better discrimination and faster reactions (Coull, 2009).

Here, we speculate whether such probability could be connected to the consistent bias in response we find. In our results, isochrony is perceived when the final interval in the sequence is, on average, 17 ms longer than the previous ones. Such an effect is consistent with a positive time-order error (TOE; see Allan, 1979 and Woodrow, 1935 for a review) and a perceptual acceleration of the final stimulus, an effect compatible with prior entry (Spence & Parise, 2010) and a recent study that showed that intervals are perceptually shortened (accelerated) when below 3 seconds (Wackermann, 2014). The fact that the duration of the last interval was underestimated is particularly interesting if we consider that the intervals used in our experiment are lower than the commonly used indifference point of 700 ms (Woodrow, 1935). The effect size does not change across the sequence durations tested, but we find that the delay required for perceived isochrony is 12 ms larger in the interleaved condition than in the blocked presentation.

If this result is interpreted as an acceleration of the last stimulus, it should be considered that the difference in hazard probability would suggest greater expectation and, thus, more anticipation with longer sequences (Elithorn & Lawrence, 1955; Luce, 1986; Näätänen, 1970; Niemi & Näätänen, 1981;). Hazard probability alone, therefore, does not explain why there should be a perceptual acceleration of the last stimulus in the blocked condition, where no uncertainty about which stimulus to judge is present. Our data, in fact, show more anticipation for the interleaved condition, where intervals are actually more uncertain than in the blocked condition. Higher predictability in the blocked condition, instead, should have led to a stronger prior entry phenomenon.

Table 1. Probabilities associated with each of the interval in the sequences in the interleaved condition (see also Coull, 2009).

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

5. Conclusions

The present study first compared existing models of temporal sensitivity in isochronous sequences before demonstrating how the length of a sequence and interleaved presentation influence temporal judgments in isochronous sequences. Our results show that discrimination sensitivity increases for longer sequences in interleaved presentation and is overall better for blocked presentation. The pattern of performance increase is consistent with the averaging of successive estimate, but with a factor limiting performance. PA and ML models propose that either correlation between successive estimates or weighting of the representation are the key factors. Neither of the two exhaustively accounts for the pattern of performance increase found. The results also evidence that perceived isochrony is obtained if the last 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

We are grateful for the insightful comments of the anonymous reviewers. This research was funded by Marie Curie CIG 304235 ‘TICS’.

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.

Grondin, S., & McAuley, J. D. (2009). Duration discrimination in crossmodal sequences. *Percept, 38*(10), 1542.

Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. *Ann N Y Acad Sci*, *423*(1), 52-77.

Helson, H. (1947). Adaptation-level as frame of reference for prediction of psychophysical data. *Am J Psychol*, 1-29.

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

Horr, N. K., & Di Luca, M. (2015). Filling the blanks in temporal intervals: The type of filling influences perceived duration and discrimination performance. *Front Psychol*, *6*, 114.

Halpern, A. R., & Darwin, C. J. (1982). Duration discrimination in a series of rhythmic events. *Percept Psychophys*, *31*(1), 86-89.

Henry, M. J., & Herrmann, B. (2014). Low-frequency neural oscillations support dynamic attending in temporal context. *Timing Time Percept*, *2*(1), 62-86.

Hirsh, I. J., Monahan, C. B., Grant, K. W., & Singh, P. G. (1990). Studies in auditory timing: 1. Simple patterns. *Percept Psychophys*, *47*(3), 215-226.

Ivry, R. B., & Hazeltine, R. E. (1995). Perception and production of temporal intervals across a range of durations: Evidence for a common timing mechanism. *J Exp Psychol Hum Percept Perform*, *21*(1), 3.

Luce, P. A. (1986). A computational analysis of uniqueness points in auditory word recognition. *Percept Psychophys*, *39*(3), 155-158.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychol Rev*, *106*(1), 119.

Lapid, E., Ulrich, R., & Rammsayer, T. (2008). On estimating the difference limen in duration discrimination tasks: A comparison of the 2AFC and the reminder task. *Percept Psychophys*, *70*(2), 291-305.

McAuley, J. D., & Jones, M. R. (2003). Modeling effects of rhythmic context on perceived duration: A comparison of interval and entrainment approaches to short-interval timing. *J Exp Psychol Hum Percept Perform*, *29*(6), 1102.

McAuley, J. D., & Kidd, G. R. (1998). Effect of deviations from temporal expectations on tempo discrimination of isochronous tone sequences. *J Exp Psychol Hum Percept Perform*, *24*(6), 1786.

Miller, N. S., & McAuley, J. D. (2005). Tempo sensitivity in isochronous tone sequences: The multiple-look model revisited. *Percept Psychophys*, *67*(7), 1150-1160.

McDonald, J. J., Teder-Sälejärvi, W. A., Di Russo, F., & Hillyard, S. A. (2005). Neural basis of auditory-induced shifts in visual time-order perception. *Nat Neurosci*, *8*(9), 1197-1202.

Näätänen, R. (1970). The diminishing time-uncertainty with the lapse of time after the warning signal in reaction-time experiments with varying fore-periods. *Acta Psychol*, *34*, 399-419.

Niemi, P., & Näätänen, R. (1981). Foreperiod and simple reaction time. *Psychol Bull*, *89*(1), 133.

Nobre, A. C., Correa, A., & Coull, J. T. (2007). The hazards of time. *Curr Opin Neurobiol*, *17*(4), 465-470.

Oruç, I., Maloney, L. T., & Landy, M. S. (2003). Weighted linear cue combination with possibly correlated error. *Vision Res*, *43*(23), 2451-2468.

Pariyadath, V., & Eagleman, D. (2007). The effect of predictability on subjective duration. *PLoS One*, *2*(11), e1264.

Pecenka, N., & Keller, P. E. (2011). The role of temporal prediction abilities in interpersonal sensorimotor synchronization. *Exp Brain Res*, *211*(3-4), 505-515.

Roberts, S. W. (1959). Control chart tests based on geometric moving averages. *Technometrics*, *1*(3), 239-250.

Rohenkohl, G., & Nobre, A. C. (2011). Alpha oscillations related to anticipatory attention follow temporal expectations. *J Neurosci*, *31*(40), 14076-14084.

Rose, D., & Summers, J. (1995). Duration illusions in a train of visual stimuli. *Percept*, *24(10)*, 1177-1177.

Schulze, H. H. (1978). The detectability of local and global displacements in regular rhythmic patterns. *Psychol Res*, *40*(2), 173-181.

Schulze, H. H. (1989). The perception of temporal deviations in isochronic patterns. *Percept Psychophys*, *45*(4), 291-296.

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends Cogn Sci*, *13*(9), 403-409.

Sternberg, S., & Knoll, R. L. (1973). The perception of temporal order: Fundamental issues and a general model. In S. Kornblum (Ed.), *Attention and performance* (pp. 629-685). New York: Academic Press.

Spence, C., & Parise, C. (2010). Prior-entry: A review. *Conscious Cogn*, *19*(1), 364-379.

Shore, D. I., Spence, C., & Klein, R. M. (2001). Visual prior entry. *Psychol Sci*, *12*(3), 205-212.

Titchener, E. B. (1908). *Lectures on the elementary psychology of feeling and attention*. New York: Macmillan.

Treisman, M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the "internal clock". *Psychol Monogr Gen App*, *77*(13), 1.

ten Oever, S., Schroeder, C. E., Poeppel, D., van Atteveldt, N., & Zion-Golumbic, E. (2014). Rhythmicity and cross-modal temporal cues facilitate detection. *Neuropsychologia*, *63*, 43-50.

ten Hoopen, G., Van Den Berg, S., Memelink, J., Bocanegra, B., & Boon, R. (2011). Multiple-look effects on temporal discrimination within sound sequences. *Att Percept Psychophys*, *73*(7), 2249-2269.

Vibell, J., Klinge, C., Zampini, M., Spence, C., & Nobre, A. C. (2007). Temporal order is coded temporally in the brain: Early event-related potential latency shifts underlying prior entry in a cross-modal temporal order judgment task. *J Cogn Neurosci*, *19*(1), 109-120.

Woodrow, H. (1935). The effect of practice upon time-order errors in the comparison of temporal intervals. *Psychol Rev*, *42*(2), 127.

Wackermann, J., Pacer, J., & Wittmann, M (2014). Perception of acoustically presented time series with varied intervals. *Acta Psychol*, *147*, 105-110.

Wolpert, D. M., & Ghahramani. (2000). Computational principles of movement neuroscience. *Nat Neuroscie*, 3(Suppl),1212–1217.

Zampini, M., Shore, D. I., & Spence, C. (2005). Audiovisual prior entry. *Neurosci lett*, *381*(3), 217-222.