Psychological Research (2002) 66: 201–207
DOI 10.1007/s00426-001-0085-y

ORIGINAL ARTICLE

Guy Madison · Björn Merker

On the limits of anisochrony in pulse attribution

Received: 17 July 2001 / Accepted: 22 November 2001 / Published online: 15 February 2002
© Springer-Verlag 2002

Abstract Pulse is the subjective experience of isochrony, which is typically elicited by series of sensory events with close to isochronous spacing, as is common in music and poetry. We measured the amount of anisochrony in a 10-event sequence with 570- to 630-ms nominal inter-onset intervals (IOI) that corresponded to the threshold for pulse attribution. This threshold was 8.6% of the IOI across 28 participants with a wide range of musical training, as compared with 3.5% for detection of anisochrony in the same kind of sequence. Musical training led to lower thresholds for detection of irregularity but had no effect on pulse attribution. The relatively larger amount of anisochrony in pulse attribution may reflect the limit for predicting and synchronising with future events. We suggest that this limit reflects a compromise between tolerance for naturally occurring deviations and the need for precision in timing.

Keywords Pulse · Music · Musicians · Timing · Temporal discrimination

Introduction

The human experience of pulse is typically elicited by a series of sensory events with isochronous or close to isochronous spacing. In such stimulus trains the occurrence of the next stimulus in the series is predictable, making isochrony within a range from approximately 200-ms to 1800-ms inter-onset interval (IOI) a cardinal device for synchronising the behaviour of several individuals (Fraisse, 1982, pp. 154–155), as in ensemble music and other group activities. Since naturally produced sequences, including human serial time production, contain a range of deviations from strict isochrony (for a review see Madison, 2000), a margin of tolerance for such deviations is necessary for synchronisation to occur. We assume that the experience of pulse is closely related to the ability to synchronise, and that it should, therefore, not be rigidly dependent on physical isochrony. On the other hand, too large a tolerance will be inefficient by leading to predictive imprecision and wastage of resources on seemingly useful information which turns out not to allow any prediction at all. Thus, we would expect our limit for attributing pulse to sensory events to represent a compromise between tolerance and the need for precision. Whatever the underlying mechanism, the purpose of the study reported here was to determine the magnitude of deviation from isochrony compatible with the attribution of pulse.

Whereas the ability to detect single deviations from isochrony has been studied in a variety of tasks (e.g., Fraisse, 1967; Friberg & Sundberg, 1995; Halpern & Darwin, 1982; Hibi, 1983; Jones & Yee, 1997), considerably less is known about sensory sequences characterised by irregular or systematic deviations from strict isochrony. Such deviations are nevertheless more ecologically relevant than are single deviations. They occur not only in music, dance, and speech, but also in acoustic and bioluminescent co-operative and competitive sexual communication in other species (for a review see Greenfield, 1994).

One kind of anisochrony that has been carefully studied is the duple pattern obtained by displacing every other event by a fixed amount of time. The detection threshold for anisochrony in duple patterns with 18 clicking sounds was 6.2% (18.5 ms) for 300-ms IOI, but decreased to 4.5% (13.5 ms) when each duple pattern was compared with an isochronous sequence according to the constant method (ten Hoopen, Boelaarts, Gruisen, Apon, Donders, Mul, & Akerboom, 1994). In another experiment with 10 clicks and 400-ms IOI, the threshold was only 3% (ten Hoopen, personal communication).

G. Madison (✉)
Department of Psychology, Uppsala University,
Box 1225, 751 42 Uppsala, Sweden
E-mail: guy.madison@psyk.uu.se
Fax: +46-18-4712123

B. Merker
Royal University College of Music, Box 27711,
115 91 Stockholm, Sweden

However, humans' experience of pulse elicited by anisochronous sequences, rather than their detection threshold, is even less studied. The only source is, to our knowledge, Wallin (1911), who divided ranges of deviation in rhythmic patterns (iambic and trochee) according to qualitative perceptual grades. His most relevant data for the present context refer to sequences with 18 alternating strong and weak sounds with 570-ms IOI, in which every sixth weak sound was displaced. In effect, the rhythm was, by a group of listeners, considered "excellent" when the displacement of the sixth sound was 6.4%, "good" when it was 8.0%, "medium" corresponded with 12.1%, 15.2% was considered "very jerky", and 17.1% was "disrupted" (Wallin, 1911, p. 107 and Table VI, p. 125).

Another example of supra-threshold deviations in nominally isochronous sound sequences is the so-called performance expression in music (e.g., Clarke, 1995). It is common within certain musical styles for performers to generate substantial temporal and other deviations from the nominal values given by the score. The amount of such timing variability may be related to the size of deviations compatible with pulse attribution, because it can be assumed not to interrupt the perceived pulse. However, these deviations are typically related to the structure of the music, that is, the pattern of nominal durations and pitches given by the score (e.g., Drake & Palmer, 1993; Gabrielsson, 1974, 1987; Repp, 1992). This is probably one reason why it has not been considered relevant to determine a general threshold for pulse attribution in musical patterns. As an example of deviations in real music, average timing profiles of nine skilled pianists playing Chopin's Etude in E major demonstrated some IOIs to be in the order of 100–150 ms longer than others, which was in excess of 30% of the typical IOI (Repp, 1998d). Also the coefficient of variation in 115 commercially available performances of this piece ranged from 10% to 30% (Repp, 1998a).

The present study was designed as a first attempt to explore the limit of anisochrony compatible with experiencing a sequence as harbouring pulse. The psychophysical thresholds for both the attribution of pulse and the detection of irregularity were furthermore measured, to assess if they were related.

Method

Participants

Participants in the experiment, 14 women and 14 men, were recruited at two colleges, one of which is a music conservatory, and were paid for their participation. There was a wide range of musical training among participants, on the basis of which they were divided into two equally sized groups, each with 7 men and 7 women. The members of the non-musician group had between 0 and 3 years of formal music education ($M = 0.85$ years) and had devoted between 0 and 15 years to playing a musical instrument, including the voice. The musician group had between 1 and 17 years of education ($M = 8.42$ years) and between 14 and 21 years of playing ($M = 16.7$ years). The non-musicians were between 20 and 56 years of age ($M = 31.3$ years) and the musicians were between 21 and 39 years of age ($M = 25.5$ years).

Stimuli and materials

The stimuli were sequences of ten identical sound events with more or less deviation from isochrony. A methodological requirement was that the pattern of deviations should vary, so as not to allow learning. Also, this variation should not affect the amount of subjective irregularity. Furthermore, this subjective irregularity should be closely related to the objective amount of deviation controlled in the experiment, to make it possible to use an adaptive psychophysical method and to minimise the amount of noise in the data.

A series of pilot experiments with different types of deviations[1] indicated that these requirements were difficult to fulfil using various kinds of random deviations. We therefore used the first 50 elements in the self-describing binary Kolakoski sequence (Kolakoski, 1966), which is typically implemented with the numbers 1 and 2. The largest number of successive identical elements in this sequence is two, whereas a sequence based on a random distribution can have an arbitrary number of successive identical elements. Intervals whose position in the series corresponded with '1' in the Kolakoski sequence were shortened, and intervals corresponding with '2' were lengthened by the same amount of time.

This is exemplified in Fig. 1, which depicts the IOIs corresponding with the first (1–10) and second ten (11–20) elements of the employed sequence 1, 2, 2, 1, 1, 2, 1, 2, 2, 1, 2, 2, 1, 1, 2, 1, 1, 2, 2, 1, 2, 1, 1, 2, 1, 2, 2, 1, 1, 2, 1, 1, 2, 1, 2, 2, 1, 2, 2, 1, 1, 2, 1, 2, 2, 1, 2, 1, 1, 2[2]. This sequence has the property that all samples of 10 successive elements from any of the 40 possible starting positions are different, and that subpatterns within samples seldom repeat themselves. This means that learning is unlikely, due to the large number of different patterns. It also means that the amount of deviation should be closely related to the amount of subjective irregularity, which should in turn be minimally affected by different patterns. This is because of the frequent switching between shortenings and lengthenings in the stimulus patterns, and because the deviations could be made with one single amount of time throughout the sequence, as opposed to some random distribution. In effect, the standard deviation equals the mean deviation

$$d = \sum \frac{|X - M|}{n}$$

All aspects of the experiment were controlled by a specially designed software running in DOS on a PC. A MPU-401-compatible MIDI interface triggered the sampled sound Prc/66 Hi Claves in an Alesis D4 drum sound module, which was presented through Panasonic RP-HT500 sound-attenuated headphones. This sound has a brief attack and a fast decay, resulting in a click quality and a supra-threshold duration of approximately 40 ms, and it has no ambience which might constitute an inter-stimulus sound event.

---

[1] We tried several varieties of two different types of deviations. The first type was based on successive samples from a rectangular distribution with a specified SD, but the combination of size of deviation and the place of deviations in the sequence turned out to have strong perceptual effects. The second type, which is similar to the one we used in the final experiment, consisted of either adding or subtracting the same duration in successive intervals according to a binary random distribution. In this case, however, it could happen that three or more intervals in succession were changed in the same direction, which resulted in isochrony for that segment of the sequence. We assessed these effects in pilot experiments with a limited population of 40 different random sequences, each generated on basis of a certain initial value for the pseudo-random algorithm, and found them too large.

[2] As can readily be inferred, the Kolakoski sequence is serially anticorrelated, and we estimated its box-counting dimension (Beran, 1994) D to 1.8.

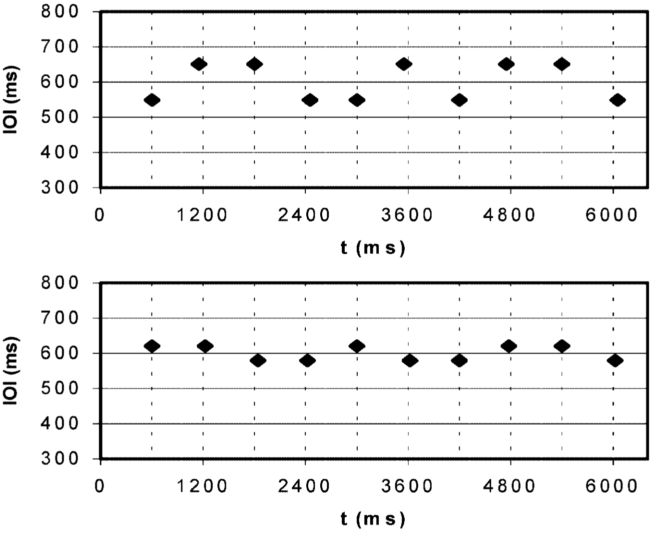

Fig. 1. Graphical representation of a typical stimulus sequence, with IOI on the *ordinate* and (cumulative) time on the *abscissa*. The *upper panel* exemplifies the amount of deviation at the threshold of pulse attribution, implemented with the first 10 (1–10) elements in the Kolakoski sequence. The *lower panel* exemplifies the amount of deviation at the threshold for detection of irregularity, implemented with the second 10 (11–20) elements in the same sequence (*IOI* inter-onset interval)

Design

The thresholds were obtained with an adaptive signal detection procedure called Parameter Estimation by Sequential Testing (PEST) (Gescheider, 1997). PEST is particularly suitable when large individual differences in the dependent variable can be expected, because it quickly adapts to the participant's criterion and moves about the threshold in smaller and smaller steps. This yes-no, forced choice paradigm was started with 10% deviation and a step size of 1.6%, which was eight times the minimal step size 0.2%. Step size is the amount of increase or decrease between judgements, in this case the amount of deviation. The deviation in the sequence following a judgement was increased with the current step size if the response indicated that the present deviation was too small, and vice versa. The step size was governed by five rules: (1) following each change in judgement (YES followed by NO, or vice versa), the step size was reduced by half until (2) it reached 0.2%. However, (3) following no change in judgement, the step size remained the same for two judgements but (4) doubled for each additional judgement up to eight times the minimum step size. When the ratio between the maximum and minimum step size is a power of two, the step size conveniently traverses the same values in both directions. Finally, (5) when a change in judgement followed a doubling of step size, three steps instead of two in the same direction were required before a doubling of step size was made. The trial ended when either eight judgements had been made with the minimal step size (not necessarily consecutive) or a maximum number of 46 judgements had been made. The threshold was defined as the mean of the eight judgements with the smallest step sizes.

The starting point should not have any effect with an adaptive procedure, and the 10% deviation was therefore chosen on the basis that it would be more motivating to start with a situation requiring action than one in which there was already a pulse or no detectable deviation from isochrony. Another reason for always starting from a large amount of deviation was that the pilot experiments suggested that the perception of pulse could be persistent, thus creating an hysteresis effect. Although this should not affect the thresholds, it might nevertheless result in longer trials and greater variability among the eight valid judgements. Another precaution against possible carry-over of pulse attribution from judgements based on smaller deviation to those based on larger deviation was to use five audibly different levels of mean IOI (the reciprocal of mean tempo), which were rotated in a sequence that maximised the change in IOI between successive judgements, namely 600, 570, 615, 585, and 630 ms. The starting point in the Kolakoski sequence for each judgement and the starting point of the IOI for each trial were both randomly chosen from a rectangular distribution.

Procedure

The entire experiment consisted of three blocks, which are described in detail below. Participants were seated in a quiet room, and instructed in Swedish. The part of this instruction pertaining to judgements translates as follows: "For every sequence of sounds your task is to answer the question before you by either a 'yes' or a 'no' using the designated computer keys. You should judge the sequence as a whole. Your judgement of pulse should be based on one pulse beat for each sound. By pulse is meant an even, regular pulse, not a rhythmic pattern".

In addition to judging whether they could attribute pulse to the stimulus sequence (henceforth called the 'pulse' task), participants were asked to detect deviations from isochrony (henceforth called the 'irregularity' task) in a separate block of trials. A total of ten pulse threshold trials were divided in two blocks separated by a block with the irregularity task. The purpose was to assess how stable the pulse thresholds would be when interrupted by a contradictory task. Six replications of the pulse threshold determination were obtained in the first block, eight replications of the irregularity threshold were obtained in the second block, and the third block comprised another four replications of the pulse threshold.

Most participants preferred to divide the experiment into two sessions run on different days. In their case the initial session consisted of instructions, the first block, and filling out the questionnaire. Remaining participants ran all three blocks consecutively in one session. The question specifying the task was mounted in large type on the computer screen facing the participant. For the first and third blocks the question was "Is there a pulse in this sound sequence, such that I would be able to beat along with it?", and for the second block "Can I hear any irregularity in this sound sequence?". These framings of the questions (in Swedish) were carefully chosen on the basis of the pilot experiments so as to be simple and unambiguous. Because the first asks if there is not irregularity (pulse) and the second if there is irregularity, the effect of pressing the yes – no buttons was reversed between the pulse and irregularity task. Whereas this might constitute a semantic response bias we deemed its possible effect to be negligible in comparison with the fundamental difference between the tasks.

Particular care was taken to make sure participants understood the difference between the pulse and irregularity tasks, and a practice session was run during which questions regarding the nature of the task or the experimental procedure were answered by the experimenter. The stimuli were presented with an individually chosen, comfortable loudness level through headphones.

After each block, participants also rated, on a five point scale, the appropriateness of a number of statements about his or her experience of the task just performed, pertaining to the extent to which the task was interesting, participant's confidence in his or her judgements, task difficulty, loss of concentration, perceived changes in the difficulty of the task over time, the meaningfulness of the task, and whether the sound stimulus was disturbing. Some of these were merely included to check that no problems had been experienced, while others might help to elucidate the results. Finally, after each experimental condition, participants answered a set of four questions in which they could express in their own words their impression of the task and how they performed. One of these questions asked whether they had used a strategy in performing the experimental task.

Results

The two first trials in the first pulse block were excluded and the irregularity block was divided in two halves to obtain equal block sizes for a four-way mixed (2 musical training × 2 task × 2 sessions × 4 replications) analysis of variance (ANOVA). With threshold as dependent variable and musical training as between-subjects variable, a significant effect was obtained for task [$F(1, 26) = 72.72$, $p < 0.00001$], but not for musical training [$F(1, 26) = 0.070$, $p = 0.793$], session [$F(1, 26) = 2.39$, $p = 0.134$], or replications [$F(1, 26) = 0.1497$, $p = 0.929$]. Task × musical training was the only significant two- or three-way interaction [$F(1, 26) = 7.735$, $p < 0.01$], with the next closest being task × session ($p = 0.176$). These results indicate that there were no substantial order-effects in the experiment.

There is no unambiguous error term for post hoc comparisons involving between-group by within-participant interactions, so the task × musical training interaction was examined with separate ANOVAs (2 musical training × 2 sessions × 4 replications) for each task.

First, the threshold for irregularity was 1.86% higher for less-trained participants ($M = 4.39$) than for those with more training ($M = 2.53$), which was significant on the Bonferoni-adjusted 0.025 (0.05/2) level, $F(1, 26) = 8.02$, $p = 0.0088$ (unadjusted p values are presented throughout). Second, using the same data as in the first ANOVA (with the first two trials excluded), the pulse threshold difference between the lower ($M = 7.86$) and higher musical training group ($M = 9.36$) was 1.5%, [$F(1, 26) = 1.83$, $p = 0.187$]. There were no significant effects of sessions or replications for any task. Whereas the pulse threshold difference was non-significant, we noted that it was even smaller (1.0%) when all ten replications were included. This is due to a tendency for less-trained participants to decrease their pulse threshold over the first few trials (9.56, 8.53, 8.60, 8.32, 8.14 and 7.62 for trials 1–6), while the threshold for more-trained participants showed a slight tendency to increase (from 8.52 to 9.40 for trials 1–6). In other words, the difference between the two groups increased with additional trials. All subsequent data presentation will be based on all eight irregularity trials and the last eight pulse trials.

Figure 2 describes the thresholds in a manner that takes the individual differences into account. Each participant is represented by a point. The points' position along the ordinate expresses the threshold in percent of the IOI (which varied from 570 to 630 ms), and their position along the abscissa expresses the consistency among the eight pulse and irregularity scores that make up that point, in terms of the coefficient of variation (CV), which is scale independent. The figure shows that the thresholds for pulse and irregularity overlap to some extent, although it must be stressed that pulse was always larger than irregularity within each participant, and that there was only a small difference in consistency between pulse (mean CV = 0.217) and irregularity (mean CV = 0.319). The meaning of these CV values can be stated in terms of a three-way mixed (2 musical training × 14 participant × 2 task) ANOVA, with the eight replications as error term, which demonstrated that all effects were highly significant ($p < 0.00001$) except musical training ($p > 0.10$). The ranges of participant means that can be gleaned from Fig. 2 were for irregularity 2.31–6.87 for less-trained participants and 0.55–7.78 for more-trained participants. For pulse, the mean thresholds were 3.78–10.10 for less-trained and 3.98–16.83 for more-trained participants.

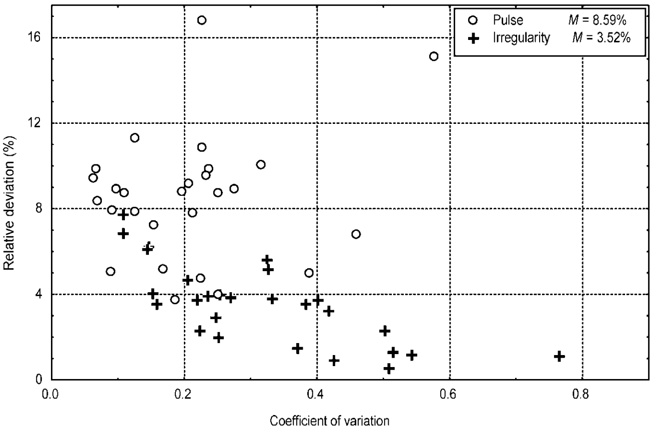

Fig. 2. Relative deviation at threshold as a function of task (pulse or irregularity). *Each point* represents one participant, whose consistency is expressed by its position along the *abscissa* (in terms of the coefficient of variation; $n_{\text{pulse}} = 8$, $n_{\text{irregularity}} = 8$)

These differences between the tasks suggest that they were uncorrelated across participants, which was indeed the case ($r = 0.03$, $n = 28$), whereas the correlation between the two pulse sessions was 650 ($p < 0.0005$, $n = 28$). Separate correlations for the two groups did not reveal any differences in these respects as a result of musical training.

The grand mean for pulse was 8.59%, and based on only between-participants variance the 95% confidence interval for pulse attribution was 7.50–9.70, and the 99% interval was 7.21–9.97.

No consistent pattern of variability could be found for the seven participants with the highest CV for pulse threshold. Two of them had run all blocks in succession, but whereas the pulse thresholds showed a tendency to increase for one of them, they decreased for the other. This ambiguity was evident also for those five who attended separate sessions, in that the thresholds tended to increase for one of them and decrease for two, while the remaining two participants showed no trend. However, the thresholds were lower in the third block whenever there was a clear break between the first and third block. As was mentioned before, this was also the general, albeit not significant, effect.

The mean number of judgements was 20.7 per trial, with a slight but nonsignificant difference between the tasks, such that irregularity required on average 1.5 fewer judgements than did pulse. The number of judgements varied from 11 to 46, which was the maximum number of judgements. Twenty-one participants had no trials with as much as 46 judgements, whereas 16 trials (3.2%) distributed among seven participants reached the maximum number of judgements. Of these, 13 trials (2.6%) distributed among five participants did not reach the prescribed eight judgements with the minimum step size. These occurrences were essentially not more common for either task, nor for certain threshold values. There was a tendency for the number of judgements and the thresholds to be correlated, but there were large individual differences in this respect. The mean r across participants was 0.10 for pulse, and –0.19 for irregularity ($n=8$), whereas individual correlations ranged from –0.83 to 0.70. A scatterplot suggested that outliers among the 448 trials were mainly responsible for these correlations, which in turn indicates that their prime cause was the participants' uncertainty about their own criterion. For example, the records for three participants with extremely small irregularity threshold values showed instances where they repeatedly pressed the "yes (I can hear an irregularity)" key, although the deviation was already zero. Thus, part of the negative correlation between the irregularity threshold and the corresponding CV, seen in Fig. 2, is likely related to the floor effect caused by the denominator (SD/M) approaching zero. Likewise, it is conceivable that pulse threshold values high above the grand mean were associated with a more variable threshold criterion, resulting both in many judgements and large variability among trials.

The participants found both tasks equally interesting (3.8 on the 5-point scale) but felt that pulse was more difficult (3.0) than irregularity (2.6). Furthermore, the 14 participants with more musical training found the irregularity task considerably less difficult (1.9) than pulse (2.5), whereas this difference was much smaller for the less-trained participants (irregularity 3.2 and pulse 3.5). Observations during the instruction phase and answers to the questionnaire indicated that some part of the body, typically finger, hand, foot, or head, was moved in concord with the pulse in 75.4% of the pulse task trials, and in 57.0% of the irregularity trials, and that this was somewhat more common among those with more musical training. No point-biserial correlation between using the body and the mean individual thresholds for irregularity ($r_{\text{pb}}=0.22$, $n=28$, $p=0.259$) or pulse ($r_{\text{pb}}=0.02$) was significant, however.

Discussion

Our study was designed as a first step in exploring the extent to which the subjective experience of pulse tolerates anisochrony in the inducing stimulus train. The difference in thresholds estimates obtained for the two tasks indicates that there is in fact a margin of tolerance for irregularity in assigning a pulse-like quality to a stimulus train, as the mean threshold for pulse attribution was more than twice as high as the mean threshold for irregularity detection in the same type of stimulus material. It should however be emphasised that these tasks are profoundly different, and required partly different procedures, all of which imposes limitations on the conclusions that can be drawn from comparing them.

It is noteworthy that the CV was no greater for pulse, although one might expect the criterion for pulse to be more variable. We hesitate to attach any significance to the small difference in CV between pulse and irregularity, since, as pointed out under results, a "floor effect" associated with the lowest thresholds for irregularity may inflate the CV in this case. The comparable amount of variability obtained for the two tasks indicates that participants were able to set a subjective criterion for the extent of irregularity compatible with classifying a stimulus train as exhibiting pulse, and were able to exercise a degree of consistency in this regard comparable to that for irregularity detection in the same type of stimulus sequence. Similar thresholds in two blocks of judgements separated by the irregularity task is one demonstration of this consistency. Since the grand mean threshold obtained was 8.59% and the maximum extent of irregularity that sequences of this sort can exhibit amounts to an average deviation of 25%, the margin of tolerance is not arbitrarily wide, but leaves ample room for being exceeded (i.e. for rejecting more irregular sequences as non-pulse-like). This is in good agreement with our suggestion in the introduction that tolerance for irregularity in pulse estimates should not be arbitrarily wide if pulse is to serve its unique synchronising function based on making the time of the next stimulus in the sequence predictable with some amount of precision. This finding is all the more remarkable in that it was obtained in the absence of any objective check on participants' actual ability to keep the pace of the pulse they claimed to perceive in the irregular stimulus trains to which they were exposed. Nevertheless, they chose a criterion of acceptability to which they were able to adhere with a measure of consistency not altogether unlike that for irregularity detection, a far less subjective task which participants also tended to rate as easier than the pulse task.

One may ask why the extent of musical training affected the threshold for irregularity but not for pulse, and why these were not correlated ($r=0.03$). In a musical context isochrony is the primary structural variable critical for synchronising the performance of several individuals, and experienced musicians might develop special sensitivity to anisochrony. For example, Jones and Yee (1997) found that musicians' lower thresholds were related to a regular underlying spacing of events in time, whereas effects of musical skill were very small for irregular patterns. Musicians might simply have less occasion to encounter deviations on the brink of failing to elicit a pulse in their musical interactions with fellow musicians, and might therefore be relatively "naive" with respect to setting such a criterion. Also, musicians have reason to develop their ability to entrain to the kind of dynamic pulse that occurs in some kinds of music, and may therefore be more apt to attribute a pulse even to very irregular sequences. Ultimately the two tasks are very different in their demand on perceptual and decision mechanisms. One might expect the pulse task to involve attitudinal and personality factors to a greater extent than the irregularity task. In this light the lack of correlation between the two sets of judgements is hardly surprising.

Regarding strategies used by participants, two aspects of the results are to be noted. A majority of participants (57% in the irregularity and 75% in the pulse tasks) used rhythmic body movements to assist their task performance. This high incidence of spontaneous use of rhythmic body movements might bear on the nature of human timing, specifically to the relationship between locomotor rhythms and time keeping (Fraisse, 1982; Merker, 2000; Todd, 1999). Since, if anything, the tendency was more pronounced among participants with more musical training, the use of the body in time keeping should not be regarded as a primitive expedient, but may be an intrinsic part of human entrainment to isochronous stimulus trains. Also, we interpret the instances of very low thresholds for irregularity in our study as an effect of subjective rhythmisation, as reviewed for example by Fraisse (1982, pp. 155–156).

No thresholds have, to our knowledge, been reported for detection of deviations similar to those used in this study. Friberg and Sundberg (1995) summarised a number of temporal discrimination studies, which for comparable IOIs and numbers of intervals indicated that thresholds for the detection of displacement and lengthening/shortening of single intervals was the highest (~6%), followed by cyclic displacement (~3–5%) and discrimination of sequences with different IOIs (~1–2%). This seems to be in good agreement with the 3.52% threshold found for the irregularity task, whose recurrent deviations from the underlying isochrony are more akin with cyclic displacement than with deviation of a single interval or with comparing perfectly isochronous sequences with different IOIs. For example, the detection threshold for a cyclic, duple pattern in a nominally 480-ms IOI sequence was 3.3%, but increased to 4.9% when the discrimination was made between a duple pattern and a standard, isochronous sequence (ten Hoopen et al., 1994).

With regard to the pulse threshold, the only comparable data seem to be Wallin's (1911), as mentioned in the introduction. While he devised qualitative perceptual categories in which to order rhythmic patterns as a function of the size of a recurring deviation, it is not clear how these categories relate to the criterion used in the present study, namely "Is there a pulse in this sound sequence, such that I would be able to beat along with it?". It seems likely that the limit for this ability should correspond with "very jerky" (15.2% of the nominal IOI), because this label suggests there is still some structure in relation to which something can be jerky. However, a close correspondence between Wallin's data with ours should not be expected, because the sequences he used were very different. Not only was the temporal deviation imposed only with every sixth sound, but the duple pattern was also created by differences in loudness rather than timing.

We mentioned that the CV for performance expression timing in a piece of solo piano music was 10–30%, in other words much larger than the obtained limit for pulse attribution. It should be emphasised that these values reflect local deviations, whereas the continually alternating deviation in our stimulus materials is global in that it introduces irregularity throughout a sequence as a whole. In the case of expressive timing, musical context provides information that allows people to anticipate certain deviations. For example, some performance timing is characterised by gradually decreasing and increasing IOIs (among sound events belonging to the same metrical level). In addition, deviations at boundaries between musically coherent sections, for example phrases, tend to be larger than at non-boundaries, and to be proportional to the (hierarchical) level of the boundary (e.g. Todd, 1989). Thus, listeners' understanding of the musical structure may create a pulse with dynamically changing tempo, which is in agreement with musicians' intuitions (e.g. Howat, 1995). Furthermore, even when the experience of pulse continuity is actually broken, listeners can quickly recover: experimentally, it has been found that two to three sounds on the pulse level suffice (Fraisse, 1946). In other words, the CV in music performances is likely to include deviations both below and above the pulse attribution threshold.

Another aspect of timing in music is temporal distortions that arise from the musical structure, so-called obligatory expectations (Repp, 1998b). Although the present kind of stimulus did not include any structure that may evoke such expectations, a comparison of the magnitude of deviations may be relevant. Obligatory expectations result either in perceived deviations in a physically isochronous rendition of music, or in an inability to detect temporal deviations that would have been clearly detectable in a simple, non-musical stimulus. Repp (1998c) investigated the detection threshold for single deviations as a function of their position in the music, and varied the 'amplitude' of a timing profile obtained by averaging across performances by nine skilled pianists. He found that the detection bias as a function of position in the structure was indeed related to the timing profile, but that the detection bias profile was flattened or inverted when only between 25% and 10% of the original timing profile was imposed. With a CV of 10–30% in those original music performances, this corresponds to 1–7.5%, to be compared with 8.59% in the present experiment.

We turn finally to the functional significance of the margin of tolerance for anisochrony in pulse attribution found in the present study. As already mentioned, the threshold derived so far is a subjective one. To relate this estimate to functional issues of synchronisation and entrainment, one would need a 'performance' measure of the pulse that participants claim to perceive. There are several problems with such measures, however, most obviously that people are generally proficient at producing pulse. It would therefore be difficult to devise a sensitive measure that actually reflects the properties of the stimulus sequence. Future studies might nevertheless profitably be directed at developing such a performance measure to explore the extent to which subjective perception of pulse in anisochronous stimulus sequences can be put to practical use in synchronising performance to anisochronous stimulus trains.

Acknowledgements Part of Dr. Madison's work in preparing this paper was supported by a grant to Alf Gabrielsson from the Bank of Sweden Tercentenary Foundation. Dr. Merker's work in this study was also supported by a grant from the Bank of Sweden Tercentenary Foundation. We wish to express our sincere gratitude to Augustin Mannerheim for the inspiration provided by his work on pulse in poetic performance, and for our stimulating discussions with him. We further thank three anonymous reviewers for valuable comments that helped improve the presentation of this study.

References

Beran, J. (1994). *Statistics for long-memory processes*. New York: Chapman and Hall.

Clarke, E. (1995). Expression in performance: generativity, perception, and semiosis. In J. Rink (Ed.), *The practice of performance* (pp. 21–54). Cambridge, UK: Cambridge University Press.

Drake, C., & Palmer, C. (1993). Accent structures in music performance. *Music Perception, 10*, 343–378.

Fraisse, P. (1946). Contribution à l'étude du rhytme en tant que forme temporelle. *Journal de Psychologie Normale et Pathologique, 39*, 283–304.

Fraisse, P. (1967). Le seuil différentiel de durée dans une suite régulière d'intervalles. *Année Psychologique, 67*, 43–49.

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

Friberg, A., & Sundberg, J. (1995). Time discrimination in a monotonic, isochronous sequence. *Journal of the Acoustical Society of America, 98*, 2524–2531.

Gabrielsson, A. (1974). Performance of rhythm patterns. *Scandinavian Journal of Psychology, 15*, 63–72.

Gabrielsson, A. (1987). Once again: The theme from Mozart's piano sonata in A minor (K.331). In A. Gabrielsson (Ed.), *Action and perception in Rhythm and Music*. (pp. 81–103). Stockholm: Publications issued by the Royal Swedish Academy of Music.

Gescheider, G. A. (1997). *Psychophysics: The fundamentals*. Hillsdale, NJ: Erlbaum.

Greenfield, M. D. (1994). Cooperation and conflict in the evolution of signal interactions. *Annual Review of Ecological Systems, 25*, 97–126.

Halpern, A. R., & Darwin, C. L. (1982). Duration discrimination in a series of rhythmic events. *Perception & Psychophysics, 31*, 86–89.

Hibi, S. (1983). Rhythm perception in repetitive sound sequence. *Acoustical Society of Japan(E), 4*, 83–95.

Howat, R. (1995). What do we perform? In J. Rink (Ed.), *The practice of performance* (pp. 3–20). Cambridge, UK: Cambridge University Press.

Jones, M. R., & Yee, W. (1997). Sensitivity to time change: The role of context and skill. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 693–709.

Kolakoski, W. (1966). Problem 5304. *American Mathematics Monthly, 73*, 681–682.

Madison, G. (2000). On the nature of variability in isochronous serial interval production. In P. Desain & L. Windsor (Eds.), *Rhythm perception and production* (pp. 95–113). Lisse, The Netherlands: Swets and Zeitlinger.

Merker, B. (2000). Synchronous chorusing and the origins of music. *Musicae scientiae, Special issue 1999–2000*, 59–73.

Repp, B. H. (1992). Diversity and commonality in music performance: an analysis of timing micro-structure in Schumann's Träumererei. *Journal of the Acoustical Society of America, 92*, 2546–2568.

Repp, B. H. (1998a). A microcosm of musical expression. I. Quantitative analysis of pianists' timing in the initial measures of Chopin's Etude in E major. *Journal of the Acoustical Society of America, 104*, 1085–1100.

Repp, B. H. (1998b). Obligatory "expectations" of expressive timing induced by perception of musical structure. *Psychological Research, 61*, 33–43.

Repp, B. H. (1998c). The detectability of local deviations from a typical expressive timing pattern. *Music Perception, 15*, 265–289.

Repp, B. H. (1998d). Variations on a theme by Chopin: relations between perception and production of timing in music. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 791–811.

ten Hoopen, G., Boelaarts, L., Gruisen, A., Apon, I., Donders, K., Mul, N., & Akerboom, S. (1994). The detection of anisynchrony in monaural and interaural sound sequences. *Perception & Psychophysics, 56*, 110–120.

Todd, N. P. M. (1989). A computational model of rubato. *Contemporary Music Review, 3*, 69–88.

Todd, N. P. M. (1999). Motion in music: a neurobiological perspective. *Music Perception, 17*, 115–126.

Wallin, J. E. W. (1911). Experimental studies of rhythm and time. I. Qualitative limens or grades of rhythm, and the difference limen in the perception of time. *Psychological Review, 18*, 100–131.