Perception, 2006, volume 35, pages 267 – 280

DOI:10.1068/p5353

The perception of tempo in music

Sandra Quinn, Roger Watt
Department of Psychology, University of Stirling, Stirling FK9 4LA, Scotland, UK;
e-mail: s.c.m.quinn@stir.ac.uk
Received 20 September 2004, in revised form 12 March 2005; published online 18 January 2006

Abstract. Tempo is one factor that is frequently associated with the expressive nature of a piece of music. Composers often indicate the tempo of a piece of music through the use of numerical markings (beats min $^{-1}$) and subjective terms (adagio, allegro). Three studies were conducted to assess whether listeners were able to make consistent judgments about tempo that varied from piece to piece. Listeners heard short extracts of Scottish music played at a range of tempi and were asked to make a two-alternative forced choice of "too fast" or "too slow" for each extract. The responses for each study were plotted as proportion of too fast responses as a function of tempo for each piece, and cumulative normal curves were fitted to each data set. The point where these curves cross 0.5 is the tempo at which the music sounds right to the listeners, referred to as the optimal tempo. The results from each study show that listeners are capable of making consistent tempo judgments and that the optimal tempo varies across extracts. The results also revealed that rhythm plays a role, but not the only role in making temporal judgments.

1 Introduction

There have been a number of attempts to investigate tempo in music. Some studies have concentrated on the variations in expressiveness of performances of pieces of music (Timmers et al 2000). Others have focused on the features of expressive performances such as rhythm (Repp et al 2002), meter (London 2002), dynamics, and mode (Kamenetsky et al 1997). Some have shown that infants as young as 7 to 9 months are capable of discriminating between simple tonal rhythmic patterns that varied in tempo (Trehub and Thorpe 1989). Whilst these investigations offer an important insight into musical abilities and the temporal deviations within performance, little is known about the processes involved in the perception of global tempo, and whether there are optimal tempi which listeners will perceive to be the most appropriate and perhaps expressive for pieces of music. The purpose of this paper is to investigate this more fully.

The main focus of psychological investigation into tempo has involved interpreting the expressive timing deviations of musicians' performances. As these deviations often mark moments in the music that heighten the expressive nature of a piece, such evidence is clearly important to our understanding of music. It has been argued that detecting deviations from global tempi in real-time performances is highly developed in musicians because of their exposure to listening and playing music (Repp 1992). Subsequent evidence indicated that the perceptual experience of temporal deviations is influenced by musical structure and not necessarily through specialised musical experience (Repp 1999). Therefore, individuals with no formal training should be able to perceive deviations in tempo.

Temporal variability between successive note onsets is due to the musician's interpretation of the expressive nuances of the music (Repp et al 2002). Large and Palmer (2002) point out that listeners are remarkable in their ability to hear musical events in terms of discrete durational categories (such as whole notes or eighth notes), even in the presence of changes in continuously varying timings that are present in human performance. It is supposed that there is an internal system that allows listeners to keep track of the durational values of notes in a piece of music and to perceive deviations in these during human performance.

Dynamic attending theory (Jones 1976, 1987, 1990; Jones and Boltz 1989) suggests that attention to an auditory stimulus is directed by an oscillator (Drake et al 2000). In the presence of musical stimuli, the oscillator will synchronise itself to the periodicity of physical characteristics of the music, such as accents and the steady beat of the music. This theory also allows for music containing periodicity at more than one level, such as melodic and harmonic changes, by suggesting that the listener makes use of multiple oscillators. These multiple oscillators are thought to produce an expectancy schema, which anticipates the succeeding temporal pattern. Specifically, the expectancy schema is organised at three levels of attending: referent level, future-oriented level, and an analytic level. The last two levels are used for attending to key and harmonic changes (future-oriented attending) or dynamics and tone onsets (analytic attending). Listeners are thought to switch between levels, through a process of focal attending, to the strongest expectancy within the stimulus. The particular oscillator that is used will be determined by the most salient features of the music. Although this theory can explain the way listeners deploy attention over time whilst listening to music, it is not intended as an explanation for the perception of global tempo in music.

Schulze (1978) attempted to assess the way listeners perceive temporal regularity in auditory sequences by testing three models of tempo perception. The first model assumes that a listener perceives temporal regularity by comparing neighbouring intervals to decipher whether the sequence is regular. The second model supposes that listeners produce their own rhythmic pattern and use that to judge the regularity of the sequence. This possibility is based upon the assumption that an internal mechanism is involved, called a 'time keeper', that synchronises with the input sequence, but in order to do so has an identical base duration to the input sequence. Lastly, the third model suggests that once the listeners have heard the first few intervals of a sequence they produce their own internal representation of the interval. This is used as a reference to judge all other incoming intervals and the stability of the sequence. The first model relates to local tempo from note onset to note onset, whilst the other two relate to global temporal effects.

Each listener heard a regular sequence and one of three comparison sequences where the regularity of the tones had been displaced. Listeners were asked to judge the comparison sequences to determine whether they were either regular or irregular. The results suggested that listeners were capable of judging the regularity of the sequences. Particularly, the results supported the notion that an internal time keeper was involved and that it was not based upon comparisons with further intervals. This is essentially similar to the assumptions made by the dynamic attending theory and supports the global internal mechanism put forward by Schulze.

The stimuli in the Schulze study were sequences of tones where the pitch was held constant. Given this, the extent to which those results can be generalised to music listening is unclear. It would be of interest to assess musical sequences that contain variations in pitch as well as the rhythmic structure where there was a musical context for assessing the perception of tempo. Furthermore, the fact that listeners were told where to expect changes in these tonal sequences seems to suggest that listeners had an expectation of what might follow. Had a control group been introduced to assess listeners' capabilities where they had no prior information, the results may have suggested that the other models were more appropriate in this context. Whilst this is not clear from Schulze's results, later studies provided support for these findings by using a psychophysical methodology.

Using the assumptions made by Schulze's internal time keeper, Vos et al (1997) attempted to assess listeners' detection of tempo changes. They investigated the role of the internal time keeper suggested by Schulze (1978) in order to predict whether listeners could detect accelerations and decelerations in the tempo of sequences of music. It was thought that if the base tempo was around 100 beats min^{-1}, the results would not be biased. If the tempo was faster than this, the listeners were assumed to become biased towards responding that they perceived an acceleration in tempo. Where the base tempo was slower than this, they were thought to be biased towards judging the sequence as decelerating. In order to assess this, listeners were required to make a forced directional-change response of "tempo acceleration" or "tempo deceleration" for sequences of tones. They were told that the stimulus would change in tempo at some point during the listening phase, and they should make a decision once the sequence had finished. Even when they were unsure of their response, they were asked to use their best guess. In essence, this methodology is more useful, because it allows the possibility of assessing the listeners' sensitivity to changes in the stimulus as well as their detection of such changes directly.

The psychometric functions of the proportion of acceleration responses were plotted against tempo change. The psychometric functions clearly showed that listeners were capable of detecting temporal changes in both directions and indicated their sensitivity to those changes. The results provided support for Schulze's internal system of tempo perception and the assumptions made by the authors. Indeed, they showed that listeners responded as often on both choices and were unbiased when global tempo was around 100 beats min^{-1}. They also showed that the listeners would become biased towards acceleration responses when the global tempo was faster than the unbiased tempi and deceleration responses when the sequence was slower than the unbiased tempi.

Few researchers have evaluated whether there is a global tempo that listeners perceive as optimal for the expressiveness of music. The amount of expressiveness relies not only on temporal deviations from global tempi, but for such deviations to occur global tempi must play a crucial role. If a piece of music is played at a tempo that is too slow for a piece, then this would affect any subsequent temporal deviations that performers use to heighten its expressiveness.

Levitin and Cook (1996) investigated whether memory for global tempo was absolute. They assessed the ability to match a tempo accurately to pieces of music that were familiar to the participants. They suggest that, like absolute pitch, where some individuals have an ability to recognise the exact pitch of a note, people may hold a similar ability to hear tempo in an absolute manner, where they retain information about the exact tempo for a piece of music. They refer to this as absolute tempo. They used data from a previous study evaluating whether people could sing a familiar song at the same pitch as the original composition. They found that the participants' reproductions were at similar tempi to the original compositions. From this, they conclude that long-term memory for tempo in music is accurate at least for music that is very familiar to the listener. As memory for tempo in music seems to be highly accurate, it could be argued that presenting a stimulus more than once allows a listener the opportunity to pull upon information they have stored in memory. It follows that, in studies measuring perception (rather than memory) musical stimuli should not be presented more than once to participants.

Lapidaki (2000) attempted to assess whether listeners could set the tempi of pieces of music in a consistent manner. The assumption was that, if preferred tempi exist, different listeners would produce similar settings. Using the method of adjustment, subjects set each piece of music to their own preferred tempo. Judgments were relatively consistent across most listeners. The initial tempo dominated the settings of preferred tempo: a slow initial tempo seemed to provoke slower tempo selections, and so on.

The results showed a bias effect, and do not necessarily show the optimal tempi for pieces of music and the musical features that drive such responses.

Repp (1994a) attempted to assess whether changes in global tempi could affect the timing of aspects of a piece of music. Schumann's *Träumerei* for the piano was chosen, as it was thought that this piece was not technically difficult, but demanded a great deal of interpretation by a performer. Two pianists played the piece three times, at slow, medium, or fast tempi that were chosen because they reflected ranges of appropriate tempi used by famous pianists in recorded performances of this piece (Repp 1992). Each tempo was introduced to the pianists by a metronome, so that they could get a feel for the specified tempi before they performed the piece. The metronome was switched off prior to their performance and the piece was recorded onto MIDI. In the perceptual condition, each performance was played to highly trained pianists who were asked to identify the original performances from those that had been artificially speeded up or slowed down to the same performed speeds the pianists had played.

The results showed that the tone onset timings were similar overall, but deviated from proportional tempi. The perceptual tasks showed no real pattern of identification, which suggests that very little difference in the musical expression was found when the tempo was increased. It may also suggest that musical ability does not seem to aid listeners' discriminations. The timing of the grace notes of the performances at all tempi showed a consistent pattern of regularity. It was thought that this was due to the relatively slow general tempi of the performances and may not have been the case had the piece included faster ornaments. Synchronisation of the chords in each performance varied little, even when they were particularly important to the overall expressive shape of that moment in the music. It was expected that as the speed became faster the legato overlap between note-to-note onsets would become less pronounced. The amount of overlap between legato notes was shown to differ as temporal variability changed at a local level. However, this was only the case for one of the pianists as this pianist's use of legato was of a more typical level, and such differences between each performer were thought to be the result of differences in technical ability. Another aspect that was not affected by global temporal changes was pedal timing which showed no real perceptible effects. Lastly, the performance intensities were also assessed by looking at changes in velocity and, although these were shown to increase, the pattern of changes was held relatively constant. These results seem to focus on the technical aspects of the musicians' performance and it is not clear whether there were musical features such as rhythm or pitch that might play a role. Desain and Honing (1994) found evidence to suggest that expressive timing in musical performance does not rely on global tempo. Their results imply that timing in human performance is related to the actual structure of the music rather than the global tempo chosen by the performer.

Repp (1994a) suggests the need to explore a variety of different types of music to evaluate these effects more fully. He also highlights that rhythmically unpredictable music that includes short notes, rests, and a variety of articulation might be affected more by changes in global tempi. However, the limited number of pieces included in both the Desain and Honing (1994) and the Repp (1994a) studies cannot substantiate this claim. Repp (1994b) implies that pianists will have an interpretation for familiar pieces of music, and when they are asked to play these pieces at a different tempo without much notice, their interpretation will not change. Had the musicians had more time to practice this at a different speed they may have changed their interpretation. In contrast, one might argue that the overall similarities between these performances and those recorded performances where global tempo seems to differ and the time spent on ensuring the interpretation was secure, lead one to believe that this might not be the case, certainly for this piece of music. A more typical argument would be that overall expressive nuances are the same in most performances of experienced musicians, but vary perceptually in listeners' levels of satisfaction, interpretation or the appropriateness of the piece (one aspect not assessed in the study).

Understanding the basic tempo is necessary in order to clarify whether there are interpretative tempi that are deemed to be most appropriate rather than only measuring the ability to match performances with other such performances in the musical domain. In line with Lapidaki, it could be argued that performers are biased towards particular tempi because of their familiarity with music. To avoid these biases, the need to explore unfamiliar pieces of music should be considered. Indeed, listeners were not asked whether the tempi being performed were deemed to be optimal or appropriate. As this might add to a body of knowledge on this topic, a thorough investigation whether there are optimal global tempi for pieces of music should be made.

Boltz (1998) investigated whether temporal and melodic accent structure also influences tempo discrimination. In her first study, listeners were asked to complete a paired tempo comparison task where the comparator melody differed from the standard in terms of the number of changes in pitch contour direction and the number of pitch skips. It was assumed that the number of contour changes and pitch skips involved in listening to music would affect the perceived tempo of a piece of music. Particularly, it was assumed that the greater the number of contour changes and the greater the number and magnitude of pitch skips, the more likely the piece would be perceived as being slower than a comparison melody. It was suggested that these changes could cause listeners to perceive prolonged accents and retards at phrase ending points and are enhanced or reduced further by actual changes in tempi. Boltz found that manipulations of pitch content did affect the perceived tempo. This offers some insight into the musical features that affect listeners' perception of tempi in music.

In her second study, Boltz (1998) examined the effect of rhythm in combination with changes in pitch on the perceived tempo of a melody. It was thought that compatible and incompatible rhythms might differentially affect the perception of tempo. Where a melody and rhythm complement one another they contain temporal and melodic combinations that allow the listener to perceive the music in a single coherent way in order to make sense of the music. Where this is not the case, the music is perceived as being ambiguous. Boltz assessed this by asking the listeners to judge the tempi of melodies that contained compatible or incompatible rhythms. She suggests that a rhythmic structure that conflicts with the melodic aspects within a piece can lead the listener to disengage from the anticipated flow of the tempo. This lack of cohesion leads the listener to perceive the music ambiguously.

It is the purpose of this paper to explore and identify those factors that are involved in the perception of the appropriateness of a given global tempo. The method of constant stimuli is used to measure this and avoid the bias encountered in the Lapidaki study. Participants have been asked to judge whether a piece of music was played too slow or too fast. The point at which the probabilities of the two responses are equal is here called the optimal tempo: it is the tempo at which the tempo is neither too fast nor to slow and is therefore presumably judged appropriate.

Three key issues are focused on. First, are listeners able to make reliable judgments about the appropriateness of the tempo of pieces of music? If listeners are capable of doing this, then is there agreement between listeners about the optimal tempo? Finally, does the optimal tempo vary from piece to piece? If it does, then the implication is that factors within the music (eg pitch contour and/or rhythm) are determining the optimal tempo.

2 Methods

2.1 *Participants*

Participants were undergraduate students from the University of Stirling. The number of participants varied from seventeen to twenty-one in each study. Most were psychology students; none was a music student. No attempt was made to either select or filter out musically trained subjects: the incidence of musical training in the sample was around 2%, and probably reflects the population incidence generally.

2.2 *Materials*

The stimuli were 23 extracts of Scottish fiddle music (see Appendix). They were all monophonic, were in various keys and had indicative score tempi that ranged from 31.5 to 132 beats min^{-1}. Although the genre of the music was familiar to the listeners, the stimuli were selected on the basis that the pieces selected would be unfamiliar to the listeners. They were constructed as MIDI data and played to participants via a digital synthesiser (MIDI instrument = Violin). All note durations were set to the full length of the note and velocity was set to a value of 80 on the MIDI data. The notes all started at the score time without variations such as rubato. Each piece of music was played at seven different tempi. In studies 1 and 2, the tempi were set to the score tempo $\pm$ 0, 10, 20, or 30 beats min^{-1}. In study 3, the tempi were set to variations about the optimal tempi found in studies 1 and 2, rather than about the score tempi. All expressive features were removed to assess the musical structures rather than human performance. Stimuli were played to participants through high-quality headphones at a volume level that was comfortable.

2.3 *Procedure*

Before the beginning of the experiment proper, each participant was given a short practice trial in order to familiarise him/her with the task. The stimuli were then played to the participants and after hearing each stimulus they were asked to make a response by indicating whether they thought it was played too fast or too slow by pressing designated keys on a computer keyboard. They were instructed to make just one of the two responses, even if they were unsure about their response.

2.4 *Design*

The experiment made use of a within-subjects design. There were two independent variables of musical extract and tempo. The dependent variable was a forced choice of "too fast" or "too slow" response. The design was constructed to avoid, as far as possible, the possibility of the stimulus set providing a normative tempo against which participants could make the tempo judgment. Participants heard the set of stimuli in a completely randomised order with neither extract nor tempo blocked. This avoids any effects of a tendency to judge the tempo of any given stimulus with respect to its immediate predecessors. Each participant heard each stimulus just once, to avoid any possible memory effect. Each participant was assigned to one of two groups. Both groups heard 4 of the extracts and the remaining extracts were split between the two groups so that group 1 heard an additional 10 melodies and group 2 heard 9 additional melodies. Each melody was played at seven different tempi making a total of 98 trials for group 1 and 91 trials for group 2. This was done to make the total length of time required from each participant manageable.

Three separate studies were conducted. In study 1, the participants were required to wait until the stimulus had been played in its entirety before they could make a response. In study 2, the participants were free to respond as soon as they were sure of their response, irrespective of whether the extract had finished or not. In study 3, the stimuli were made of a non-pitched drum beat rather than a pitched instrument, thereby removing the melodic component of the stimuli but leaving the rhythmic element unchanged.

3 Results

3.1 *Study 1*

The data were collected into psychometric functions for each extract, giving the proportion of participants responding "too fast" as a function of the tempo of the music. Three sample psychometric functions and the data are shown in figure 1. For each extract, the proportion rises as tempo increases; in most cases the psychometric functions produced a full range of responses from 0 "too slow" to 1 "too fast". Each psychometric function was analysed by finding the least-squares fitting cumulative normal function by using a simplex minimisation routine (this is equivalent to the pre-computer technique of probit analysis). The curve is then represented by two parameters, the PSE (the point at which the function crosses 50%) and the SD (inversely proportional to the slope of the curve). The PSE is the point at which the responses "too fast" and "too slow" are equally likely, and is thus the point at which the tempo is perceived as optimal by the participants.

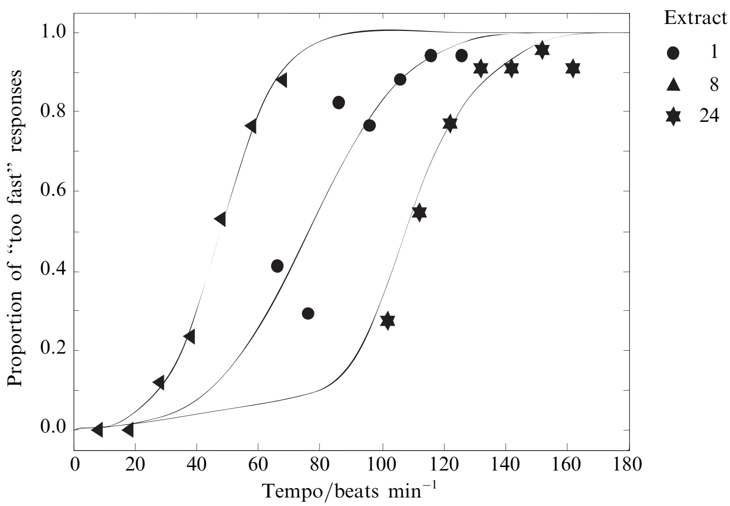

Figure 1. The proportion of "too fast" responses as a function of tempo for extracts 1, 8, and 24. The symbols are data points; the smooth curves are least-squares fits to the data. The psychometric functions rise as tempo increases and participants produced a full range of responses from 0 "too slow" to 1 "too fast". The optimal tempi can be obtained at the point where the psychometric curves cross 0.5 and produced a range from 45 to 120 beats min^{-1}.

Figure 2 shows the values obtained for the optimal tempi for each extract. The main observation is that these are uniformly scattered between 40 and 120 beats min^{-1}. The 95% confidence limits, derived by a bootstrap technique, also shown in figure 2, clearly indicate that variation in optimal tempi from extract to extract is systematic and not due to random sampling of some single underlying value. Each extract has its own optimal tempo, which is significantly different from the optimal tempi of other extracts. To assess the statistical significance of this observation, a global optimal tempo was calculated using all the data together and compared with the individual optimal tempi for the different extracts. Bootstrap analysis shows that the null hypothesis that all optimal tempi are the same can be rejected ($p < 0.0001$).

Before moving on, let us consider the issue whether each individual participant is making self-consistent judgments about the tempo of each extract. Since each participant made only one response to each extract and tempo combination, the psychometric functions for individual participants are comprised of only 0 and 1 values. If the judgments are self-consistent, each participant would have a pattern where all the responses for tempi less than some critical value would be "too slow" and all responses

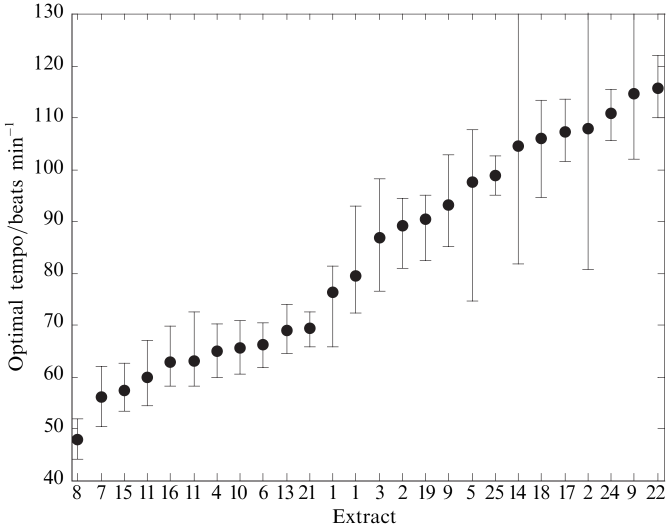

Figure 2. The optimal tempi for each extract with 95% confidence limits on each extract (shown as vertical bars). The extracts are ordered by their optimal tempo. Were there any clusterings of values, then these would appear as plateau in the graph, and it can be seen that the optimal tempi are uniformly distributed over a wide range. The conclusion is that each extract has its own characteristic optimal tempo.

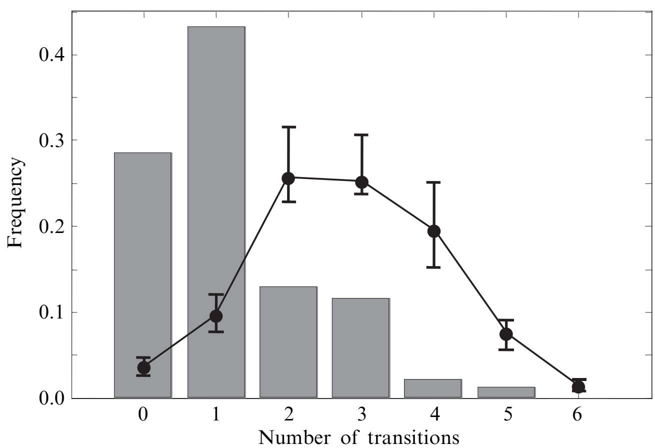

Figure 3. Each participant heard each extract once at each tempo. The psychometric function for an individual participant will switch between 0 and 1 as tempo increases. The number of transitions in either direction is a measure of how consistent the participant is in his/her response to tempo: one or less transitions would indicate perfect consistency. The graph shows the distribution of the number of transitions in response from "too slow" to "too fast" for each extract and each participant for study 1 and the expected distribution (with 95% confidence limits) on the null hypothesis that transitions are randomly distributed bars. On 60% of occasions participants made no more than 1 transition in response [compared with 15% expected by chance ($p < 0.001$)]. The conclusion is that participants are responding in a self-consistent manner.

for faster tempi would be "too fast": ideally there would be no more than one transition in response from 0 to 1 or vice versa as tempo increases. By counting the number of response transitions for each participant and for each extract, it is possible to assess how far each individual is behaving self-consistently. Figure 3 shows the distribution of the number of transitions per participant and per extract. It also shows the expected distribution on the null hypothesis that transitions are randomly distributed with 95%

confidence limits, as calculated by a bootstrap technique with a repeated random sampling of the responses made by participants. It can be seen that on 65% of occasions there was no more than one transition in response, compared with 15% expected by chance ($p < 0.001$).

The second issue is whether different participants show similar optimal tempi for the same extract. For each extract we can estimate the optimal tempo for each participant, by fitting a cumulative normal curve to the individual data. A simple measure of the extent to which these are similar is the spread of the distribution across participants. Figure 4 shows the distribution of individual optimal tempi about the mean optimal tempo for each extract for study 1. It also shows the distribution expected on the null hypothesis, with 95% confidence limits, obtained by bootstrap. It can be seen that the spread of individual optimal tempi is narrower than expected by chance ($p < 0.01$).

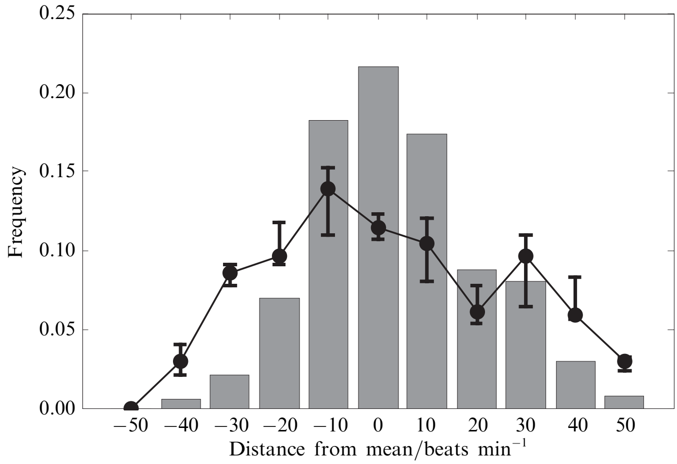

Figure 4. The dispersion of optimal tempo measures for individual participants and individual extracts about the optimal tempo for each extract for study 1 (bars). The figure shows the expected distribution on the null hypothesis that individual participants are not sensitive to difference between extracts. The spread of individual tempi is narrower than expected by chance ($p < 0.01$).

3.2 *Study 2*

This study was a repeat of the first study, with the difference that participants could respond at any time after the start of the stimulus. For each extract, a different range of test tempi was chosen, varying around the measured optimal tempi of study 1, rather than around the score tempi. This study was carried out principally to establish whether the measured optimal tempi for the various extracts were stable properties that could be replicated with a fresh sample of participants and a slightly different technique. Figure 5 shows the optimal tempo from study 2 plotted against the optimal tempo from study 1. The data are closely correlated, indicating that the measured optimal tempo for an extract can be replicated.

3.3 *Study 3*

In the final study the procedure was repeated, but with a non-pitched drum beat rather than a pitched instrument, thereby removing the melodic component of the stimuli. The optimal tempi for these stimuli are plotted in figure 6 against the optimal tempi measured in study 1. As can be seen, there is still some relation, but the scatter is larger than in study 2. We conclude that rhythm plays some part in the judgments of the optimality of tempo, but that the pitch contour is also important.

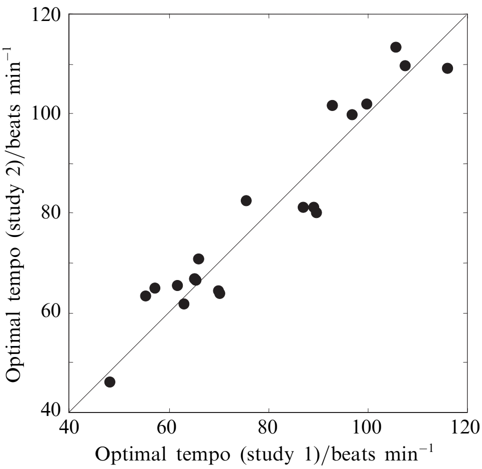

Figure 5. The optimal tempi for study 2 compared to the results of study 1. Optimal tempi for the same extract from the two studies are seen to be similar.

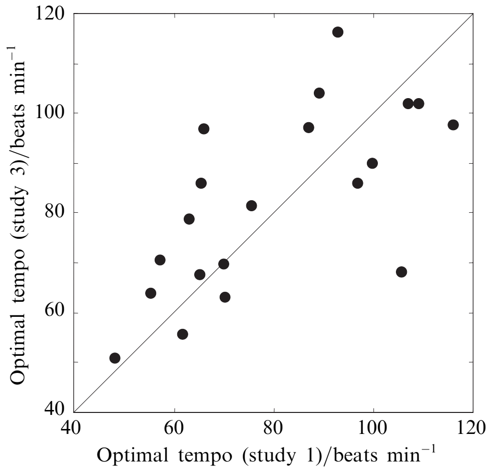

Figure 6. The optimal tempi for study 3 compared to study 1. The optimal tempi for the same extract from the two studies are similar. The scatter is wider than in figure 5, showing that rhythm alone can deliver the optimal tempo in some, but not all, cases.

4 Discussion

It has been shown that untrained (but not unexposed) listeners can make consistent judgments about the appropriateness of the tempo of pieces of music, and can judge whether a piece of music is played too fast or too slowly. The main result is that the judged optimal tempo varies significantly from piece to piece, and this variability is robust across three different studies. This result, although perhaps not entirely surprising in musical terms, has some significant implications.

An important part of the procedure was the randomisation of all stimuli so that neither piece nor tempo was blocked. This precaution means that the possibility that listeners were basing their judgments on a comparison between their memory for a previously heard extract with the extract they were responding to at that time is unlikely to explain the pattern of results. The variation of optimal tempi across different extracts showed that listeners were making responses that were independent of the context in which they heard each extract.

Having repeated the procedure with stimuli from which pitch variations were removed, leaving rhythm-only stimuli, we found that the basic result is very similar to that obtained with the melody present. The lack of melodic structure does not seem to inhibit listener's abilities to make decisions about the appropriateness of the tempo, in two-thirds of the cases. As previous studies show that ambiguity arises when rhythm and melodic structure are not compatible, it is interesting that, in absence of a melody, consistent "too fast" and "too slow" judgments could still be made.

One possible explanation for these findings might be drawn from a cultural perspective. It could be argued that listeners hold expectations about the specific tempo that Scottish music should be performed at because it is frequently associated with dancing. However, the results showed that the optimal tempo for each extract varied across a wide spectrum. This suggests that the results are not influenced by any cultural preconceptions.

The most likely musicological explanations do not prove useful in explaining these results (see table 1). It could have been expected that the major and minor distinction might have caused the results to be such that listeners prefer the music to be fast in a

Table 1. Correlation between optimal tempi in study 1 and musical features. Almost all the musical features are not correlated with the optimal tempo, except the number of descending intervals. Most musicological features do not account for the optimal tempo for extracts of music.

Musical feature	Correlation
Major/minor	−0.16
Number of pitches	0.41
Note length SD	0.40
Frequency of most common pitches	−0.28
Mean interval direction	−0.37
Mean interval size	0.31
Number of ascending intervals	0.37
Number of descending intervals	0.49*
Number of chord changes	0.17

Note: $*p < 0.05$.

major key or slow in a minor key. This is not the case in our studies. Although chords were not played to the listeners explicitly, it is possible that the number of harmonic changes implied by the musical structure may have provided them with a cue to judgments about tempo. A greater number of harmonic changes might cause listeners to judge the music as faster than those with fewer changes. However, the number of chord changes did not correlate with the optimal tempo.

Boltz (1998) found that the number of changes in pitch contour direction (ascending or descending) and the number and magnitude of pitch changes contained in the music influence the perceived tempo of the music. In contrast to this, our results showed that the number of pitches, the mean interval size, and the mean interval direction were not correlated with perceived optimal tempi. However, the number of descending intervals did partially correlate with optimal tempo, although the number of ascending intervals did not.

Lastly, no correlation was found between the variability of note durations (standard deviation of the length of the notes) or the frequency of the most common notes with optimal tempo. Apart from pitch contour, the most likely musicological features do not seem to be involved in the perception of tempo. In absence of any clear explanation, we need to turn to a perceptual account in order to explain these results more fully.

Having ruled out most musicological accounts for the main result, we turn to consider possible mechanisms for the formation of a reliable judgment of tempo optimality. In order to assess the optimal tempo of music, the listeners might first decide how fast the music is being played: the tempo of the stimulus. A further stage would be necessary, as a tempo measurement would only state how fast the music is being played, not whether it is "too fast" or "too slow". In order to make a response, listeners would need to make a comparison between the supposed tempo measurement and some criterion optimal tempo that they consider to be appropriate. The present findings show that the criterion by which they make a response is driven by the stimulus and not internally generated.

The notion that internal devices, either in terms of an internal measuring device or clock (Clarke 1987) or an internal oscillator (Drake et al 2000; Large and Jones 1999; Schulze 1978), are involved in the experience of temporal aspects of music has proven useful. However, the present study begins to shape an entirely different picture of the usefulness of internal mechanisms for perceiving tempo in music. If internal mechanisms are involved, either in terms of a measuring device or an oscillator, it would be expected that such a measurement would be highly accurate and precise. Indeed, oscillatory attending is thought to be highly accurate where the musical stimulus has a regular rhythm (Large and Jones 1999). As our stimuli contained no variations in inter-onset intervals or changes in velocity that are produced in expressive performances, then it could be argued that oscillatory attending is heightened further. In these circumstances, the way in which the listener attends to the music is extremely focused and can accurately predict the occurrence of the underlying temporal regularity.

Such exact predictions or measurements of the stimulus tempo may or may not be a part of the process that delivers a "too fast" response, but it is not the whole process. If a tempo measurement is made, then some other measurement, indicating the preferred tempo (rather than the actual tempo), also needs to be made, in parallel, from the same stimulus. This is possible, but we would like to explore a different approach: that the tempo of the stimulus is measured, not in absolute terms as events per unit time from which "too slow" or "too fast" is derived, but instead is measured directly as being "too slow" or "too fast". We wish to consider the possibility that tempo relative to optimum is perceived directly rather like a higher-order invariant (Gibson 1966).

This direct perception of tempo relative to optimal could arise because the events were occurring faster or slower than predicted by some internal device, but there are no grounds for supposing that such devices would make predictions that differed from the actual timing of the stimulus events. An alternative is that the nature of events (such as individual pitches) coupled with their temporal sequence, are used to make predictions about the timing of the next event. Where the event nature and timing are incompatible, the tempo is perceived to be suboptimal.

Our results indicate that the perception of the appropriateness of the tempo for an extract of music is determined by the contents of the extract itself. We have not been able to identify any simple musical correlate of this and, instead, turn to consider perceptual features of the music. It might be thought that variations in the speed with which the different musical extracts can be processed sets different optimal tempi, but there is no evidence of a failure to process the stimuli. Furthermore, the stimuli, being simple tonal and monophonic, are very much simpler than much music that people encounter.

Suppose that music has events that vary in their character, possibly along several simultaneous dimensions. The perceptual effect of the music will be due to the time-varying nature of this character, but also to the speed with which the character varies. To say this is to say little, except that it then opens up the possibility that there is an important perceptual interaction between event character and event duration. The overall effect of a piece of music depends on getting the tempo right for the character of the particular events in the music. This can be illustrated by supposing that events in musical stimuli vary in strength, and that strong events lose effectiveness by being rushed.

It is then reasonable to suggest that a piece of music that is played too quickly and is filled with strong events would be perceived by the listener as "too fast". Likewise, a piece of music that is played slowly and is filled with weak events would be perceived as "too slow". The strength of events within a piece of music lead the listener to determine what tempi are appropriate. The relative strengths of these events may require them to unfold more or less quickly in order to construct a perceptual representation of them that makes sense to the listener. These assumptions are speculative and necessitate further research in order to establish what event strength is and what is its relationship with perceived tempo in music.

References

Boltz M G, 1998 "Tempo discrimination of musical patterns: Effects due to pitch and rhythmic structure" *Perception & Psychophysics* **60** 1357 – 1373

Clarke E F, 1987 "Levels of structure in the organization of musical time" *Contemporary Music Review* **2** 211 – 238

Desain P, Honing H, 1994 "Does expressive timing in music performance scale proportionally with tempo" *Psychological Research* **56** 285 – 292

Drake C, Jones M R, Baruch C, 2000 "The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending" *Cognition* **77** 251 – 288

Gibson J J, 1966 *The Senses Considered as Perceptual Systems* (Boston, MA: Houghton and Mifflin)

Jones M R, 1976 "Time, our lost dimension: toward a new theory of perception, attention, and memory" *Psychological Review* **83** 323 – 355

Jones M R, 1987 "Dynamic pattern structure in music: recent theory and research" *Perception & Psychophysics* **41** 631 – 634

Jones M R, 1990 "Learning and the development of expectancies: an interactionist approach" *Psychomusicology* **9** 193 – 228

Jones M R, Boltz M, 1989 "Dynamic attending and response to time" *Psychological Review* **96** 459 – 491

Kamenetsky S B, Hill D S, Trehub S E, 1997 "Effect of tempo and dynamics on the perception of emotion in music" *Psychology of Music* **25** 149 – 160

Lapidaki E, 2000 "Stability of tempo perception in music listening" *Music Education Research* **2** 25 – 44

Large E W, Jones M R, 1999 "The dynamics of attending: How people track time-varying events" *Psychological Review* **106** 119 – 159

Large E W, Palmer C, 2002 "Perceiving temporal regularity in music" *Cognitive Science* **26** 1 – 37

Levitin D J, Cook P R, 1996 "Memory for musical tempo: Additional evidence that auditory memory is absolute" *Perception & Psychophysics* **58** 927 – 935

London J, 2002 "Cognitive constraints on metric systems: some observation and hypotheses" *Music Perception* **19** 529 – 551

Repp B H, 1992 "Probing the cognitive representations of musical time: Structural constraints on the perception of timing perturbations" *Cognition* **44** 241 – 281

Repp B H, 1994a "Relational invariance of expressive microstructure across global tempo changes in music performance: An exploratory study" *Psychological Research* **56** 269 – 284

Repp B H, 1994b "On determining the basic tempo of an expressive music performance" *Psychology of Music* **22** 157 – 167

Repp B H, 1999 "Relationships between performance timing, perception of timing perturbations, and perceptual – motor synchronization in two Chopin preludes" *Australian Journal of Psychology* **51** 188 – 203

Repp B H, Windsor L, Desain P, 2002 "Effects of tempo on the timing of simple musical rhythms" *Music Perception* **19** 565 – 593

Schulze H-H, 1978 "The detectability of local and global displacements in regular rhythmic patterns" *Psychological Research* **40** 173 – 181

Timmers R, Ashley R, Desain P, Heijink H, 2000 "The influence of musical context on tempo rubato" *Journal of New Music Research* **29** 131 – 158

Trehub S, Thorpe L A, 1989 "Infants' perception of rhythm: Categorization of auditory sequences by temporal structure" *Canadian Journal of Psychology* **43** 217 – 229

Vos P G, Assen M V, Fraňek M, 1997 "Perceived tempo change is dependent on base tempo and direction of change: Evidence for a generalized version of Schulze's (1978) internal beat model" *Psychological Research* **59** 240 – 247

Appendix

Extracts 1 to 16 are from Hunter (1979); extracts 17, 18, 19, 21, 24, and 25 from Martin (2002a); and extracts 20, 22, and 23 from Martin (2002b).

1. *Angus Cameron's Complements to Alex Webster*
2. *Bonnie Glenfarg*
3. *Lament for the Death of Rev. Archie Beaton*
4. *Lament of Flora MacDonald*
5. *MacPherson's Rant*
6. *Mairi Bhan Og Mary, Young and Fair*
7. *Mr Garden Troup's Farewell to France*
8. *Mrs Hamilton of Pentcaitland*
9. *Mrs Helen N Robertson*
10. *Roslin Castle*
11. *Sitting on the Stern of a Boat*
12. *The Duchess of Bedford*
13. *The Duchess of Manchester's Farewell to the Highlands of Scotland*
14. *The Fallen Chief*
15. *The Marchoness of Huntly's Favourite*
16. *The Marquee of Huntly's Snuff Mill*
17. *The Kilworth Hills*
18. *Miss Drummond of Perth*
19. *Loch Torridon*
21. *Queen Victoria's Diamond Jubilee*
22. *Stornoway Castle*
24. *Bob Steele*
25. *Lochanside*

References for this Appendix

Hunter J, 1979 *The Fiddle Music of Scotland* (Edinburgh: T and A Constable)
Martin C, 2002a *The Fiddle Music of the Scottish Highlands; Ceol Na Fidhle* 1 and 2 (Isle of Skye: Taigh Na Teud)
Martin C, 2002b *The Fiddle Music of the Scottish Highlands; Ceol Na Fidhle* 3 and 4 (Isle of Skye: Taigh Na Teud)

ISSN 0301-0066 (print) ISSN 1468-4233 (electronic)

PERCEPTION
VOLUME 35 2006

www.perceptionweb.com