MUSICOLOGY AND SISTER DISCIPLINES

PAST, PRESENT, FUTURE

Proceedings of the 16th International Congress
of the International Musicological Society
London, 1997

Edited by

DAVID GREER

with Ian Rumbold and Jonathan King

OXFORD
UNIVERSITY PRESS

The Timing Implications of Musical Structures

BRUNO H. REPP

For the last decade or so, the author has been engaged in empirical research on the small, unnotated variations in musical sound patterns that convey what is commonly referred to as expression in performance. The work has focused on piano performance, in part for methodological reasons and in part because the author is a pianist himself. Among the several parameters of expression, timing (rubato) has received the greatest attention. This paper summarizes the results of recent studies which have yielded four kinds of empirical evidence converging on a single hypothesis: that musical structures have specific timing implications which constrain the considerable variety of expressive timing patterns observed in individual performances.

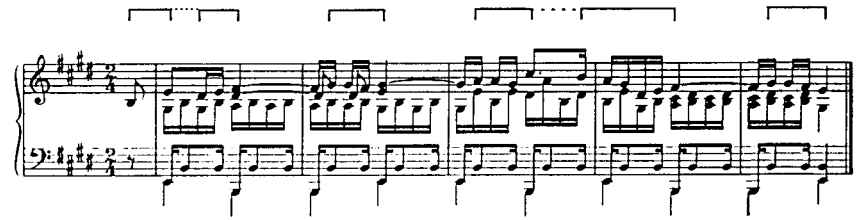
While musicologists are usually concerned with large works and entire *oeuvres* of composers, the present research, owing to the laborious empirical methods involved, is based on a mere five bars of music. Clearly, this raises questions about the generality of the results. While some of the findings have already been replicated with different musical excerpts, others are fairly new and in need of further investigation.

The music in question is the beginning of Chopin's Étude in E major, op. 10 no. 3. Example 1 shows a score of this excerpt, with slurs and expression marks omitted and a final chord added. With the exception of the initial quaver upbeat, the music proceeds entirely in semiquaver intervals. That is, at each semiquaver subdivision of each beat, there is a note onset in at least one of the several voices. As indicated below the score, these note onsets define thirty-six nominally equal inter-onset intervals (IOIs). This is convenient for the study of expressive timing, which in this case is simply the pattern of the actual durations of the thirty-six IOIs in a performance. (The initial upbeat was disregarded in all analyses.)

The music consists of a segmented melody in the soprano voice, a continuous accompaniment in the alto voice, and a repetitive rhythmic pattern in the tenor and bass voices. The melodic segments are indicated above the score. Each segment is terminated by (the onset of) a sustained longer note, during which the accompaniment continues. Obviously, the soprano melody is the

This research was supported by NIH grant MH-51230.

EXAMPLE 1. The initial five bars of Chopin's Étude in E major, op. 10 no. 3, without slurs and expression marks, terminated with a chord. Melodic segments and IOIs are indicated above and below respectively.



most salient voice, both psychologically and in terms of its dynamic level in performance, and the alto accompaniment tends to be more salient than the lower voices. The perceptually relevant semiquaver IOIs thus occur sometimes within the soprano voice, sometimes within the alto voice, and sometimes between these two voices. The harmonic structure of the music is largely congruent with the melodic segmentation, in that alternations between tonic and dominant coincide with the onsets of the segment-final long notes. The only exception is at the beginning of bar 4, where a harmonic change occurs near the beginning of a melodic segment.

Nearly all performances of this music by expert pianists exhibit large tempo modulations in the service of expression, which are reflected in the IOI durations. The pattern of the actual durations of the thirty-six semiquaver IOIs constitutes the 'timing profile' of a performance. The hypothesis addressed by the research was that a particular musical structure implies a particular timing profile which provides a norm for expressive performance. While it is generally agreed that expressive timing is subject to certain conventions and constraints, the proposal that these constraints originate in the musical structure itself may seem obvious to some but controversial to others. Let us now consider some of the empirical results.

The first kind of evidence comes from the measurement and statistical analysis of large samples of expressive performances. The author has measured the timing of 115 commercially recorded performances of the Chopin Étude,[1] and he has also recorded eighteen advanced student and amateur pianists who played the excerpt on a digital piano, essentially sight-reading the piece. The remarkable fact is that the average timing profiles of the two groups are extremely similar. This is shown graphically in Figure 1.[2] The only major difference is in the last IOI, which was longer in the students'

[1] B. H. Repp, 'A Microcosm of Musical Expression, I: Quantitative Analysis of Pianists' Timing in the Initial Measures of Chopin's Étude in E major', *Journal of the Acoustical Society of America*, 104 (1998), 1085–1100.

[2] Similar but preliminary data were reported in B. H. Repp, 'The Aesthetic Quality of a Quantitatively Average Music Performance: Two Preliminary Experiments', *Music Perception*, 14 (1997), 419–44.

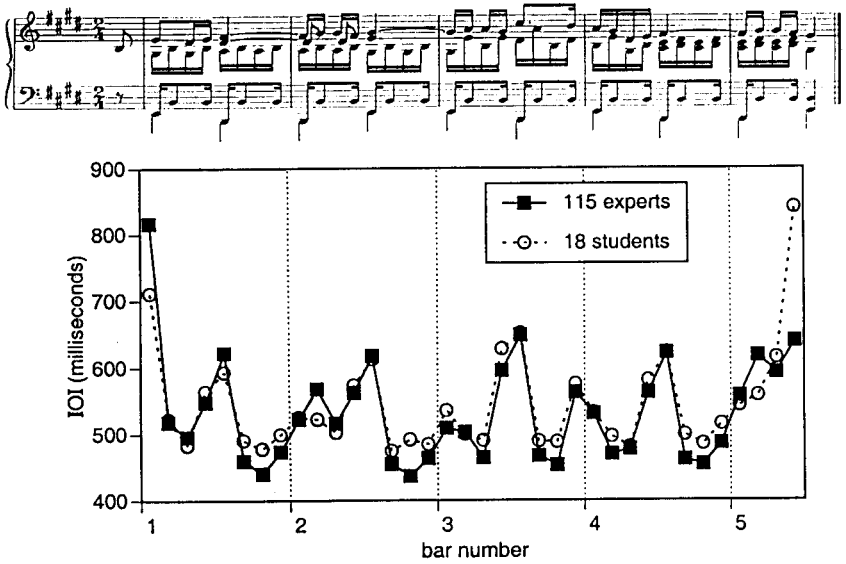


FIGURE 1. The average timing profiles of 115 expert performances (commercially recorded) and of eighteen performances by advanced student and amateur pianists (MIDI recordings).

performances because they played only the excerpt and therefore made a large final ritardando. Also, the initial downbeat was somewhat shorter in the students' performances. The correlation between the two profiles, without the final IOI, is .94. (The maximum possible value of a correlation is 1.) A similarly close correspondence of expert and student pianists' average timing profiles was found in previous analyses of two longer pieces played by fewer pianists, Robert Schumann's *Träumerei* and Claude Debussy's Prélude *La fille aux cheveux de lin*.[3]

An average profile is not unrepresentative of individual performances. The student performances in particular tended to be quite similar to their average. The commercially recorded performances showed more individual variation, but there were many famous pianists whose timing profiles were similar to the average profile. This profile may represent the expressive timing demanded by the music, as it were.

Let us have a closer look at such an average profile, which may also be considered a typical timing profile for the music. Figure 2 shows again the average profile of the 115 expert performances, but here the IOIs initiated by a

[3] B. H. Repp, 'Expressive Timing in Schumann's "Träumerei": An Analysis of Performances by Graduate Student Pianists', *Journal of the Acoustical Society of America*, 98 (1995), 2413–27; *idem*, 'Expressive Timing in a Debussy Prelude: A Comparison of Student and Expert Pianists', *Musicae scientiae*, 1 (1997), 257–68.

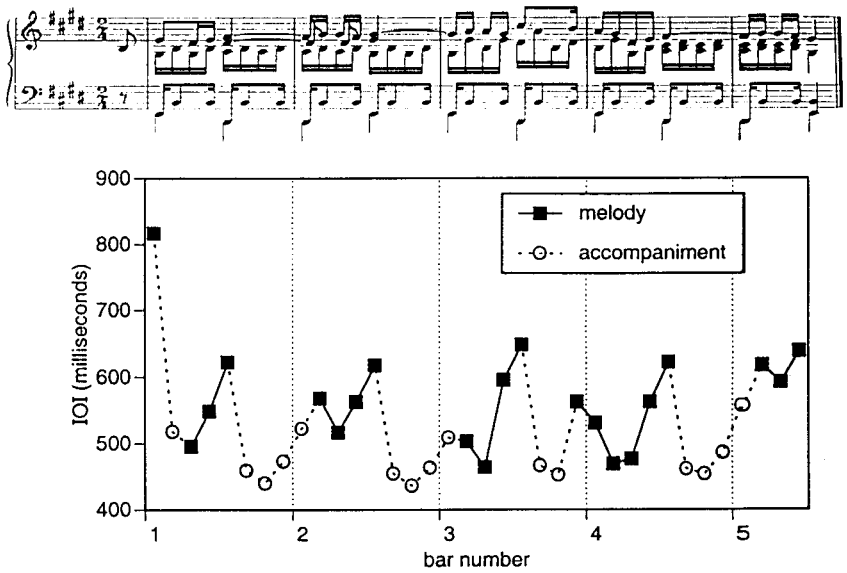


FIGURE 2. The expert average timing profile (cf. Figure 1), with IOIs initiated by melody notes (filled squares) distinguished from IOIs initiated by accompaniment notes (open circles).

soprano melody note (filled squares) are distinguished from those that are initiated by a note in the alto accompaniment (open circles). The initial downbeat IOI is usually greatly elongated. Each melodic segment typically shows a final ritardando, and the longest segment also shows an initial acceleration. There is also a tendency to decelerate in the accompaniment immediately before the beginning of a melodic segment. Clearly, all these expressive ritardandos are related to the musical structure, particularly the melodic segmentation, though harmonic progression and metre may also contribute. Further research with varied materials is needed to disentangle the effects of different aspects of musical structure on typical expressive timing.

The point here is simply that a particular musical structure is associated with a particular typical timing profile. This typical timing pattern is aesthetically pleasing.[4] It may serve as a norm or prototype that constrains the actual timing of pianists by serving as the focal point of attraction in a dynamic system of two conflicting tendencies: to be as faithful as possible to the composer's intentions (i.e. to the musical score) and yet to make an interesting and individual statement in performance. Typical timing, by definition, is neither interesting nor individual, but it is highly appropriate to the music. Deviation from this bland norm requires cognitive effort and

[4] See Repp, 'The Aesthetic Quality of a Quantitatively Average Music Performance'.

imagination, whereas adherence to the norm merely requires musical competence. The former is seen more often among experienced artists, the latter among students and amateurs, although some renowned artists also play with very unadventurous timing.

Performance analyses alone are not sufficient to prove that the typical timing profile is demanded by the music. Perhaps it reflects only a central tendency in a changing performance tradition and has no real functional role in music perception and cognition. Three other kinds of evidence, however, suggest that the typical timing profile is wedded to the musical structure at a more basic perceptual and motoric level.

One result comes, paradoxically, from inexpressive performances. Several researchers have noted that, when pianists are asked to play without expression, their performances still show small variations in timing that bear some resemblance to the timing of their original expressive performance.[5] In a recent study, six pianists (all advanced students or amateurs) were recorded playing the Chopin Étude excerpt twenty times on a digital piano, the first ten times with normal expression and the second ten times 'metronomically', i.e. in strict time.[6] The timing profiles of the ten performances in each set were averaged to reduce unsystematic timing variation. Figure 3 compares these average expressive and metronomic timing profiles for each of the six pianists. In each case, there was systematic residual timing variation in the metronomic condition, and the correlations between expressive and metronomic timing profiles (the 'r' values shown in Figure 3) were very high in five pianists and moderately high in one (T.C.) who played with little expressive timing to begin with. Although the individual timing profiles were different from each other, they were all rather similar to the typical timing profile shown earlier, as is characteristic of student and amateur pianists, and this was also true in the absence of any intention to play expressively. These data are consistent with the interpretation that the residual timing variation is truly unintended and reflects the timing demands of the musical structure, which have a small but unavoidable effect on pianists' actions. However, it could also be argued that the pianists did not succeed in completely suppressing their expressive intentions. Therefore, these results are suggestive but not conclusive.

Another, stronger kind of evidence comes from a listening task in which participants tried to detect momentary hesitations in an otherwise strictly metronomic performance.[7] The Chopin Étude excerpt was again used, synthesized on a digital piano under computer control. In each presentation of the excerpt, thirty-two of the thirty-six IOIs were equal in duration, while four were made slightly longer. These four detection targets occurred in unpredictable locations but never in close succession. In the course of a block of nine presentations, each of the thirty-six IOIs in the excerpt was lengthened once by the same amount, and a total of ten blocks was presented. Participants pressed a key on the computer keyboard whenever they heard a hesitation in the music, and received immediate feedback about the correctness of their response. The magnitudes of the duration increments were adjusted individually for each participant so as to result in an average detection score of about 60% correct.

[5] C. E. Seashore, *Psychology of Music* (New York, 1938; repr. New York, 1967); C. Palmer, 'Mapping Musical Thought to Musical Performance', *Journal of Experimental Psychology: Human Perception and Performance*, 15 (1989), 331–46; A. Penel and C. Drake, 'Sources of Timing Variations in Music Performance: A Psychological Segmentation Model', *Psychological Research*, 61 (1998), 12–32.

[6] B. H. Repp, 'Precision of Expressive and Metronomical Timing in Pianists', *Journal of Motor Behavior*, 31 (1999), 145–64.

[7] B. H. Repp, 'Detecting Deviations from Metronomic Timing in Music: Effects of Perceptual Structure on the Mental Timekeeper', *Perception and Psychophysics*, 61 (1999), 529–48.

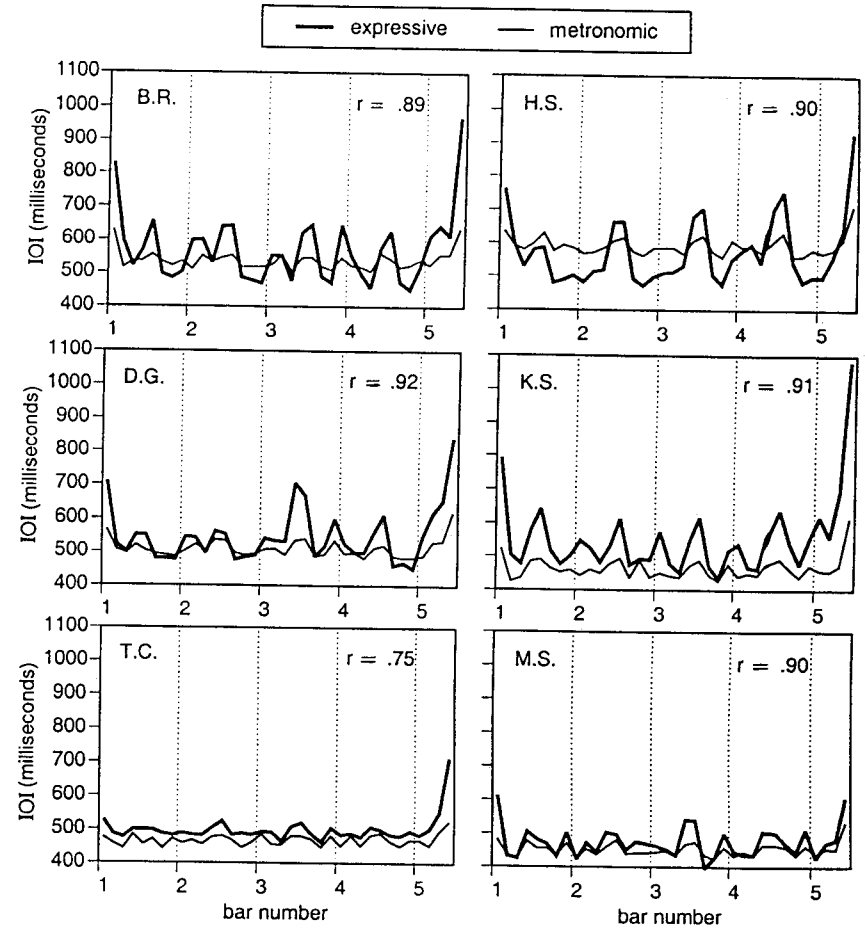


FIGURE 3. Comparison of expressive and metronomic timing profiles of six individual pianists, and correlations between the profiles.

By averaging the results across forty-three participants, a 'detection accuracy profile' was obtained which shows average percent correct detection as a function of location in the music. This profile is shown in the upper panel of Figure 4, whereas the lower panel shows an inverted typical timing profile, i.e. a tempo profile, based on expressive performances by eighteen student and amateur pianists. The remarkable finding is that these two profiles are closely related—a result which has been obtained in a number of earlier studies, also

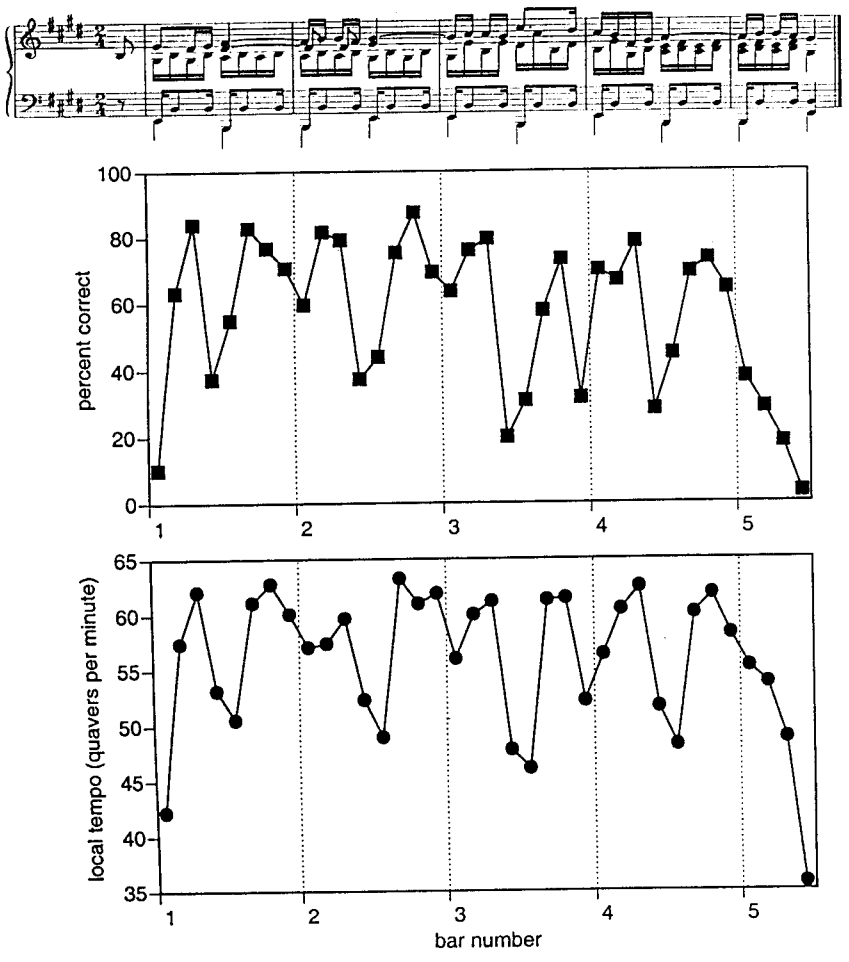


FIGURE 4. Upper panel: average detection accuracy profile of a group of listeners. Lower panel: average tempo (inverted timing) profile of 18 student and amateur pianists (cf. Figure 1).

with different musical materials.[8] The correlation between the two profiles is .90. What this means is that local hesitations in a metronomic performance are difficult to perceive where a slowing of local tempo is likely to occur in an expressive performance. This suggests that listeners expect to hear a typical timing pattern.

Where do these timing expectations come from? One might think that they reflect musically trained listeners' experience with expressive music performance—in other words, that they are cognitive expectations and that they would be malleable to some extent. However, the results of several recent experiments suggest that this is not so.[9] The perceptual effects seem to be obligatory: they do not depend on musical experience, and they also seem to be resistant to practice. Thus it is not merely the case that hesitations are aesthetically more acceptable where lengthened IOIs are expected; they really are very difficult to perceive in these locations.

These perceptual results then reflect indirectly what an average timing profile reflects directly, namely the fact that musical structures have specific timing implications. Hesitations are difficult to perceive in certain locations because the local musical structure demands a lengthening in that place. In expressive performance, musicians either obey these demands or deliberately resist them by controlling their movements during performance. Listeners, however, cannot resist them because they have no control over the auditory processes that govern simple detection. If there is no lengthening where it is demanded by the music, listeners show obligatory perceptual compensation: an IOI that should be long sounds relatively too short, and therefore a small lengthening is not easily detected because it merely restores subjective evenness. These perceptual compensation effects are subliminal in magnitude: a truly metronomic performance does not really sound uneven. Its perceptual unevenness is revealed only indirectly in the detection task.[10]

This leads to the fourth and final piece of empirical evidence. In the detection task, listeners clearly must employ a timekeeping mechanism in their brains to be able to detect deviations from metronomic regularity. If the period of this mental timekeeper is sensitive only to acoustic note onsets, then the variations in detectability of deviations from isochrony must arise at some subsequent stage in perceptual processing. Alternatively, it is possible that the

[8] B. H. Repp, 'Probing the Cognitive Representation of Musical Time: Structural Constraints on the Perception of Timing Perturbations', *Cognition*, 44 (1992), 241–81; *idem*, 'Detectability of Duration and Intensity Increments in Melody Tones: A Partial Connection between Music Perception and Performance', *Perception and Psychophysics*, 57 (1995), 1217–32; *idem*, 'Variations on a Theme by Chopin: Relations between Perception and Production of Deviations from Isochrony in Music', *Journal of Experimental Psychology: Human Perception and Performance*, 24 (1998), 791–811.

[9] Repp, 'Detectability of Duration and Intensity Increments in Melody Tones'; *idem*, 'Obligatory "Expectations" of Expressive Timing Induced by Perception of Musical Structure', *Psychological Research*, 61 (1998), 33–43; *idem*, 'Detecting Deviations from Metronomic Timing in Music'.

[10] See also B. H. Repp, 'The Detectability of Local Deviations from a Typical Expressive Timing Pattern', *Music Perception*, 15 (1998), 265–89.

period of the mental timekeeper is systematically modulated by the timing demands of the music. In other words, when the local musical structure calls for a slowing of the tempo, the mental timekeeper may predict a slight delay of the next event, and this may be the reason why an actual slight delay is difficult to detect. In a recent study, a motor synchronization task was used as a window on the operation of the mental timekeeper.[11] Participants were required to tap with their index finger on a response key in synchrony with the metronomically timed Chopin excerpt, aligning their taps as closely as possible with semiquaver onsets. By averaging the IOIs between taps across sixty presentations of the music, a 'tap-timing profile' was obtained for each participant, and averaging across participants yielded a grand average tap-timing profile. If the mental timekeeper is flexible and responsive to the timing demands of the musical structure, then there should be a positive correlation between the tap-timing profile and the typical expressive timing profile.

Figure 5 compares these two profiles. The tap-timing profile is in the upper panel. It shows systematic variations around the fixed IOI duration of 500 milliseconds. The first three tap IOIs are rather short, because the first tap tended to be late and it took three or four taps to catch up with the music. If these three IOIs are disregarded, the correlation between the two profiles is .58, which is significant. Thus there was a tendency to delay a tap where lengthening was expected. The modulations of the tap-timing profile are quite small, within a range of only ±10 milliseconds or ±2%, which is below the conscious detection threshold. The correlation is not as high as the one between the average detection accuracy profile and the typical timing profile (see Figure 4). One reason for this is the need to maintain synchrony in the tapping task: for example, it is not easily possible to lengthen two successive IOIs; rather, a lengthening of one IOI induces a strong tendency to shorten the next IOI, due to an automatic feedback and correction mechanism that is sensitive to sub-threshold deviations from synchrony. In expressive performance, by contrast, there is no strong constraint to adhere to a constant underlying tempo; the tempo is flexible and is controlled by an equally flexible timing mechanism in the musician's mind. The synchronization task reveals that the mental timekeeper is flexible even when it is tracking inflexibly timed music: it wants to follow the timing demands of the music and escape from the mechanical regimen, but can do so only within narrow limits that are set by the task requirements. These intriguing results are new and require further study.[12]

In summary, the data presented here provide considerable support for the hypothesis that a particular musical structure is associated with a particular

[11] Repp, 'Detecting Deviations from Metronomic Timing in Music'.

[12] See B. H. Repp, 'Relationships between Performance Timing, Perception of Timing Perturbations, and Perceptual-Motor Synchronisation in Two Chopin Preludes', *Australian Journal of Psychology*, 51 (1999), 188–203.

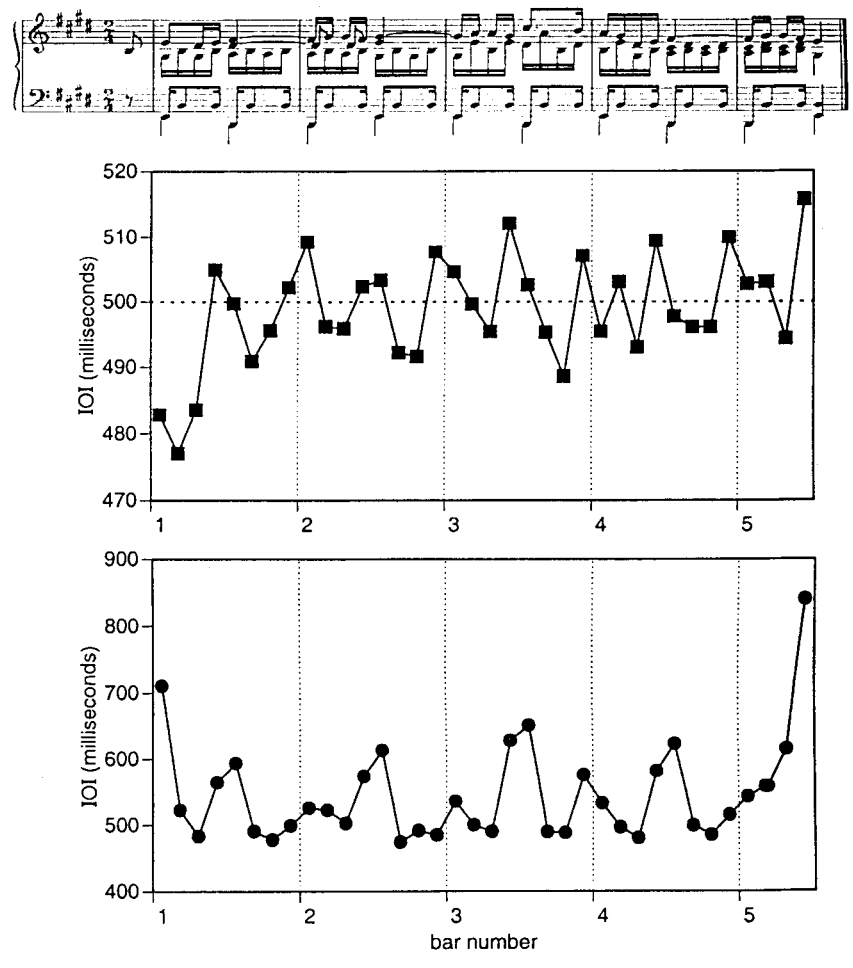


FIGURE 5. Upper panel: average tap-timing profile to a musical stimulus with IOIs constant at 500 milliseconds. Lower panel: the average timing profile of eighteen advanced student and amateur pianists (cf. Figure 1).

pattern of expressive timing. The timing demands of the music are revealed not only in typical performance but also in basic perceptual-motor effects that are independent of musical experience and conscious control. They may be the result of processes of auditory scene analysis applied to music,[13] that is, of the grouping of events, the segregation of voices, and the assessment of pitch contours and distances. Future research will have to clarify the relative importance of different aspects of musical structure for timing at this basic

[13] A. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound* (Cambridge, MA, 1990).

level, but rhythmic grouping is likely to be a major factor, as it is in performance timing. The perceptual-motor basis of these timing implications provides an objective foundation for the kind of expressive timing that is considered natural and aesthetically appropriate for a given musical structure. This typical timing may be seen as an amplification of the subliminal timing variations observed at the perceptual-motor level. A musician playing with typical timing may be said to play with intuitive, natural expression. Individuality and originality in expressive timing may arise in part from the additional application of cognitive faculties to the musical structure, though such an analytical approach may also merely reinforce a typical interpretation. Since cognitive processes may become automatized and subconscious with practice, it is difficult to separate cleanly cognition from perception or intuition in music performance. However, to the extent that the results of detection and synchronization tasks reflect a precognitive level of music perception, they suggest that typical expressive timing can be based on intuition rather than active analysis of musical structure: the music can 'speak for itself'.

Two more general, perhaps provocative implications follow from this research. With regard to Western (notated) music, there is a widespread tendency to think of expressive timing as deviations from a default metronomic timing implied by the score, deviations that are introduced deliberately by a performer. The present results suggest instead that typical expressive timing is an integral part of music, and that it is therefore neither a deviation nor deliberate, whereas a metronomic performance (even an imaginary one) is a distortion. A performer who wishes to be faithful to the score should (and usually will) play with typical timing, not metronomically.

Another implication of the perceptual-motor findings is that the typical timing profile for a given composition is likely to be independent of historical changes in performance practice. Thus the timing implications of Chopin's Étude were presumably exactly the same in Chopin's time as now. That is, even though performances of 150 years ago may have differed in many ways from today's performances, their expressive timing probably revolved around the same central norm. What may have changed in the intervening years is musicians' preferred basic tempo, the average magnitude of their tempo modulations, and their willingness to deviate from typical timing. In addition, of course, there may have been historical changes in other expressive aspects of performance (e.g. dynamics, articulation, pedalling) that are more instrument-dependent and less closely tied to the musical structure than timing.