15

Ensemble performance: Interpersonal alignment of musical expression

Peter E. Keller

<A>Introduction

Music performance is typically a collective affair, with multiple individuals coming together in ensembles of varying size and composition. Although aesthetic and social goals of ensemble performance may vary as a function of piece, genre, socio-cultural context, and performers' predilections, ensemble musicians commonly aim to communicate information about musical structure and expressive intentions to co-performers and audience members. As in solo performance, this aim is pursued by deliberately introducing variations in performance parameters such as tempo, intensity, articulation, and sound quality. The special challenge for ensemble performers, however, is that these expressive devices are no longer merely a matter of individual variation, but rather inter-individual co-variation.

The current chapter addresses the question of how experienced ensemble musicians coordinate their actions to bring about the optimal co-variation of expressive performance parameters. This question is considered from the perspective of western classical music (though other genres are touched upon) and at two levels, one descriptive and the other explanatory. The descriptive level encompasses quantitative measures of behavioral cues related to the expressive co-variation that characterizes ensemble performance. The explanatory level deals with psychological mechanisms that facilitate such co-variation by enabling ensemble performers to achieve precise interpersonal coordination in fundamental musical parameters (timing, intensity, and intonation), while simultaneously displaying the flexibility required to match artful stylistic expression. Finally, cognitive, motor, and social factors that potentially constrain or influence the operation of these mechanisms will be discussed. Figure 1 provides an overview of key concepts broached in this chapter. This overview provides a heuristic framework for investigating the psychology of ensemble performance, and the research reviewed in the following sections of the chapter represents our current understanding of this topic.

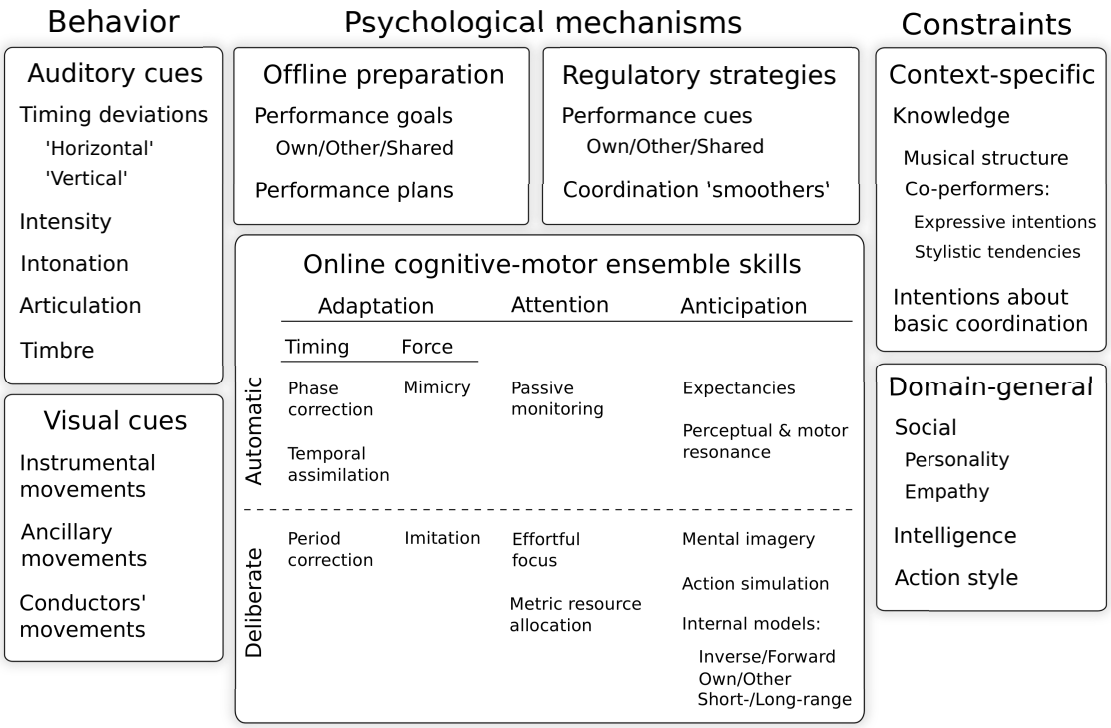

Figure 1: Overview of concepts related to behavioral cues, psychological mechanisms, and factors that may constrain or influence interpersonal coordination in expressive ensemble performance. Details about these concepts and their interaction with one another are given in the main body of the text.

<A>The nature of expressive ensemble performance

Aesthetically motivated variations in performance parameters are optimal to the extent that they foster the effective communication of musical structure and expressive intentions. Pertinent structural information includes (1) formal units such as musical motives, phrases, and sections, (2) temporal units of underlying metric frameworks (i.e., hierarchically nested measures, beats, and beat-subdivisions), and (3) surface features such as melodic, duration- and intensity based accents (i.e., points of local emphasis). Expressive intentions refer to the specific musical character with which a performer—or group of performers—wishes to imbue a piece. These intentions are realized during performance through deviations from what would be considered to be a prototypical interpretation of the piece in the prevailing cultural context.

Variations that are designed to convey information about musical structure and expressive intentions must be matched across co-performers when the aim of an ensemble is to sound unified. Given that music making typically involves the production of complex sound sequences through bodily motion, this co-variation in performance parameters spans different modalities (e.g., auditory and visual) and dimensions (e.g., timing and intensity). Ensemble performers so communicate information about musical structure and expression to co-performers and the

audience via multiple cues (Figure 1, left). Such communication may not only serve aesthetic goals, but also function to regulate the coordination of ensemble performers. The interpersonal alignment of fundamental and expressive performance parameters fosters ‘ensemble cohesion’.

<B>Auditory cues

The dimensions of sound that musicians modulate for structural and expressive purposes include basic properties such as timing and intensity, as well as subtle features associated with intonation, articulation, and timbre.

<C>Timing

Timing deviations in ensemble performance affect horizontal and vertical relations between sounds. Horizontal relations pertain to the timing of successive sounds within each voice or instrumental part, while vertical relations are a matter of the degree of synchronization of sounds in separate parts. Horizontal and vertical deviations are enmeshed in the sense that the greater the horizontal deviation, the more challenging it is to maintain optimal vertical relations:

tempo changes create a moving synchronization target. Nevertheless, systematic modulations of both types of relation give music vitality and aesthetic appeal.

A common form or horizontal deviation involves intentionally varying the length of metric inter-beat intervals to produce systematic modulations in local tempo (Clarke 1988; Gabrielsson 2003). For example, tempo typically decreases in the vicinity of boundaries in musical structure and the amount of decrease reflects hierarchical relations, with greater slowing associated with events at relatively high levels in the music's structural organization (e.g., phrase endings) (Todd 1985). Such horizontal timing deviations are remarkably consistent across performances of a piece by the same individual and different individuals (Seashore 1938). Timing deviations in performances of the same piece by different musicians display qualitative similarities, while leaving room for individuality (Repp 1992).

Vertical timing deviations affect the temporal relationship between sounds that are nominally synchronous in the ensemble texture. The fact that these deviations are observed in competent ensembles suggests that strict simultaneity is not only impossible (due to perceptual-motor limitations), but also undesirable in human music making. As noted by Rasch (1988 p.81), "The

asynchronization of simultaneous tones should be regarded as one of the vital deviations in the performance of music".

Ensemble cohesion may be perceived to be optimal in the presence of specific types of vertical timing deviations that vary systematically according to the musical context. Thus, in various genres—including classical art music and jazz—vertical timing deviations serve functions related to defining the music's expressive character and giving it the vital drive, or 'groove' (see Janata et al. 2012), that induces pleasure and the urge to move in listeners.

Vertical timing deviations can be quantified by computing the asynchrony—for example, the time difference in milliseconds (ms)—between sound onsets within pairs of nominally synchronous sounds in a piece. To do so, it is first necessary to define the time points at which sound onsets occur in each part. This can be challenging when working with acoustic recordings, where identifying the perceptual onset of sounds with different rise times and intensities is a nontrivial matter. This issue can be avoided to some degree by using movement sensors, motion capture technology, or digital instruments (such as electronic pianos) where onset times can be estimated based on the time code associated with the action (e.g., a keystroke) that triggers sound (cf. chapter 13 in this volume).

Computing asynchronies involves subtracting the onset time for a sound in one part from the onset time of a nominally synchronous sound in another (referent) part. If the referent part sounds *ahead* of the other part, then the asynchrony will be *negative* in sign. If the referent part sounds *after* the other part, then the asynchrony will be *positive*. Signed asynchronies therefore indicate which part is leading in the ensemble texture. If there is no consistent lead part, then the asynchrony will be close to 0 ms on average, though the variability of asynchronies (as indexed by their standard deviation, for example) will obviously be greater than zero. A more general indication of synchrony is given by computing the absolute (unsigned) value of asynchronies, a measure that ignores leader-follower relations and is useful for providing an index of the average temporal separation between nominally synchronous sounds even when there is no global leader. Generally speaking, signed and unsigned mean asynchrony measures provide inverse indices of synchronization accuracy (i.e., the closeness of events in one part to events in the referent part), while the variability of asynchronies is an inverse index of the precision (or stability) of synchronization.

Only a handful of published studies report detailed measurements of asynchronies between parts during naturalistic ensemble performance. These include studies of professional recorder, woodwind, and string trios (Rasch 1988), piano duos (Keller & Appel 2010; Shaffer, 1984), and commercial recordings of mixed jazz ensembles (e.g., Ashley 2002; Butterfield 2010; Friberg & Sundström 2002; Prögler 1995; also chapter 4 in this volume [Fabian 201*]). Further studies have measured asynchronies in the context of controlled experimental tasks requiring the production of simple musical sequences by pairs of pianists (Goebl & Palmer 2009; Loehr & Palmer 2011), violinists (Moore & Chen 2010), a string quartet (e.g., Marchini et al. 2012), and the synchronization of one part of a piano duet with a recording of the complementary part (Keller et al. 2007).

The combined results of these studies indicate that unsigned asynchronies typically range from 0 ms to around 100 ms, though occasionally large asynchronies of up to about 200 ms can be observed (Figure 2). Intentional asynchronies larger than 200 ms are rare, but they do occur under some circumstances, such as when a jazz soloist anticipates or delays an entry at the beginning of a phrase. The average value of unsigned asynchronies in expressively timed performances typically falls within the range of 30-50 ms. Mean signed asynchrony, however, is

often close to 0 ms, and falls within the range of -5 to +5 ms unless fixed leader-follower relations (e.g., melody lead) characterize the performance. The standard deviation of signed asynchronies—like the mean unsigned asynchrony—often lies within the 30-50 ms range. This variability is not completely random, however. Similar to the case with horizontal timing deviations, vertical timing deviations are applied systematically in expressive performance, as evidenced by findings that interpersonal asynchrony time series are correlated across repeat performances (Shaffer 1984).

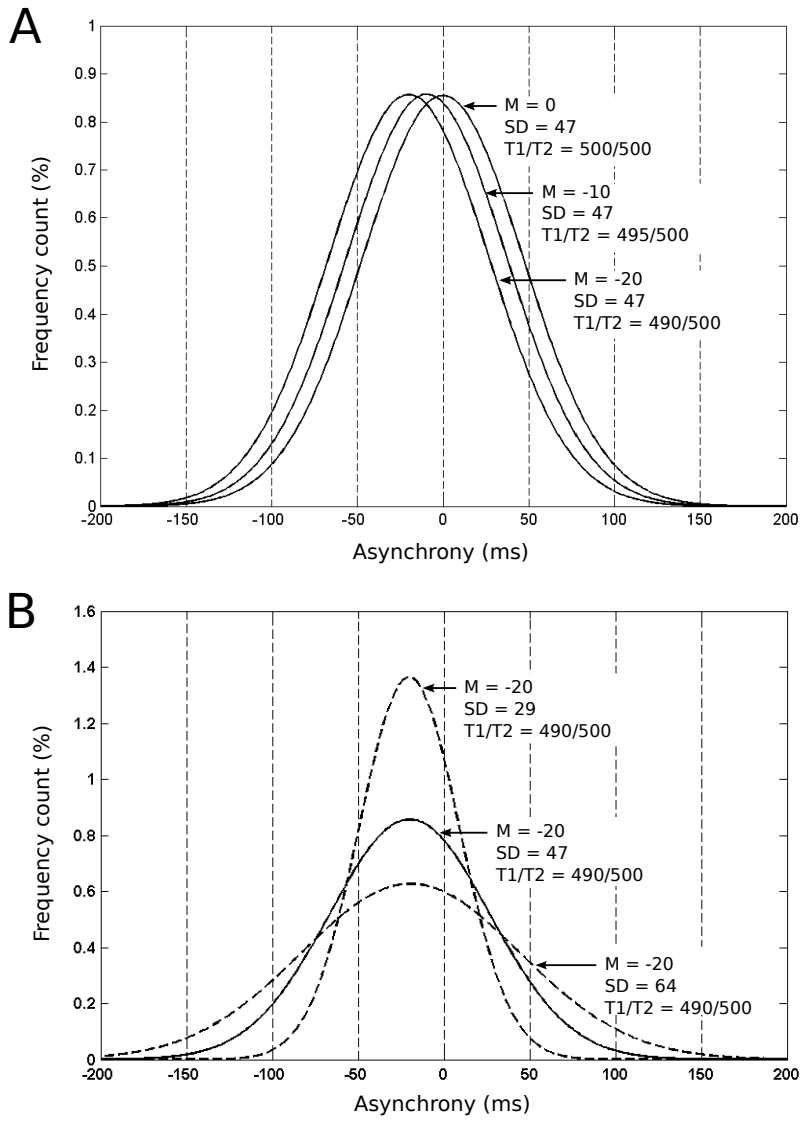

Figure 2: Schematic frequency distributions of vertical timing deviations (signed asynchronies, in ms) representing values reported in studies of expressive ensemble performance. Each

distribution was generated via a computer simulation (based on the procedure used by Repp & Keller 2008) that produced 50,000 asynchronies. (Panel A) The standard deviation (SD) of each distribution is the same, but the mean (M) varies to reflect different degrees of melody lead (0 to 20 ms). The tempo (base inter-beat interval, in ms) of the simulated melody part (T1) was to be to the same or faster than the accompaniment (T2) to simulate varying degrees of melody lead (cf. Vorberg & Wing 1996). (Panel B) The means are constant but different standard deviations were produced by scaling the noise parameter of the simulated internal timekeeper. The solid-line distribution represents a typical degree of variation (which may be considered a synchronization 'sweet spot'), while the dashed-line distributions represent relatively narrow and wide ranges of variability.

Vertical timing deviations have multiple determinants apart from perceptual and motor limitations. Additional factors include the degree to which ensemble members can see and hear one another during performance (Goebl & Palmer 2009), mechanical aspects of production such as bowing (Moore & Chen 2010), and fine acoustic considerations such as differences in the rise times of instrument sounds and the time delay between the production of a player's own tones and the perception of tones produced by others (Rasch 1988). Vertical timing deviations may

also be affected by factors that are related to musical expression, including tempo, musical structure, style, and leader-follower relations between parts in the ensemble. Some research has investigated these factors through measurements in the context of expressive ensemble performance, but the bulk of relevant work has been done using sensorimotor synchronization tasks that require participants to produce simple movements (typically finger taps) in time with computer-controlled pacing sequences.

<C>Tempo

There is a general tendency for vertical timing deviations to be larger and more variable at slow than at fast tempi (Rasch, 1988). Although this tendency has not been explored extensively in research on musical ensembles, the relationship between asynchrony and tempo has been examined thoroughly in sensorimotor synchronization studies using paced finger tapping paradigms (Repp 2006a; Semjen et al. 2000). These studies have revealed that synchronization with isochronous auditory pacing sequences can be maintained at a range of tempi, with the upper rate limit being with inter-onset intervals of 100-125 ms (16th notes at 150-120 beats per minute [bpm]) and the lower limit—assuming that the intervals are not subdivided—lying

around 1800 ms (33 bpm). Synchronization accuracy and precision are generally better at fast tempi within this range.

Sensorimotor synchronization is also affected by the direction of tempo change (acceleration or deceleration) in the sense that unsigned asynchronies are larger for accelerations than decelerations (Loehr et al. 2011; Madison & Merker 2005). Furthermore, there is a tendency to 'overshoot'—that is, tap too fast for accelerations and too slow for decelerations—at transitions between tempi when the direction and/or magnitude of the tempo change is unpredictable (Repp & Keller 2004; Schulze et al. 2005). During gradual tempo changes, taps tend to lag for accelerations and to lead for decelerations (Madison & Merker 2005; Schulze et al., 2002). Although these tendencies may have limited relevance when experienced ensemble performers who have the opportunity to rehearse specific patterns of tempo change, they may hold pedagogical implications for less experienced ensemble performers to the extent that expressive tempo variations will affect their ability to coordinate.

<C>Musical structure

Structural factors that have been found to affect the magnitude of vertical timing deviations between ensemble performers include phrase boundaries, metric location, and rhythmic grouping. Phrase entries and endings present obvious challenges for ensemble coordination. Coordination at phrase endings is difficult because tempo may slow to varying degrees, depending on the performers' expressive intentions, as phrase boundaries are approached (Repp 1992). The coordination of phrase entries is also difficult; especially if preceded by a rest or longer period of silence (Rasch 1988). Phrase boundaries often coincide with structural boundaries that Schögler (1999-2000) refers to as 'points of change'. These points are characterized by intense communicative interaction to effect a qualitative transformation in the 'feel' of the music. Schögler (1999-2000) presents evidence from improvising duos that playing becomes more synchronous immediately prior to points of change. Ensemble performers apparently engage in a form of temporal mobilization in preparation for upcoming coordination challenges.

Vertical timing deviations may also vary as a function of location within the music's metric structure; specifically, the beat or beat subdivision on which nominally synchronous sounds occur within a bar. Analyses of the temporal relationship between instrumental parts in jazz

recordings have revealed that soloists lag behind the ride cymbal on downbeats but synchronize on off-beats (Friberg & Sundström 2002).These effects are modulated by tempo, with soloists' downbeat delays being larger at tempi in the range of 100-200 bpm than in the faster 250-350 bpm range. Such relations between metric location and asynchrony are perhaps not solely stylistic in origin, as work with basic sensorimotor synchronization paradigms has found that finger taps are delayed at downbeats relative to other beats when tapping in time with metrically structured pacing sequences (Keller & Repp 2005; Snyder et al. 2006).

The temporal grouping structure of the surface rhythm may produce even more pronounced effects upon asynchronies than those associated with metric location. A study that required the synchronization of finger taps with uneven rhythms (containing intervals related by 2:3 ratios, as is common in Balkan music) found that the timing of taps associated with group-initial tones (i.e., tones occurring after the relatively long silent interval separating rhythmic groups) was late relative to taps associated with group-medial and group-final tones (Repp et al. 2005). These effects, which can be attributed to grouping-based accents (Povel & Essens 1985), prevailed even when grouping structure conflicted with the instructed metrical interpretation of the patterns.

Other studies have demonstrated that sensorimotor synchronization with cyclically presented rhythms becomes more difficult with increasing complexity in terms of the number of elements and different durations that each cycle contains, as well as the mathematical complexity of the ratios between these durations (Repp 2006a, 2006b). In addition to the complexity of the pacing rhythm, the required mode of coordination can affect synchronization. Qualitatively different patterns of asynchrony arise depending on the metric level (bar, beat, or beat-subdivision) that serves as the synchronization target (e.g., Large et al. 2002), with accuracy being greater at higher metric levels (the beat and bar) than at lower levels (beat subdivisions) (Rankin et al. 2009). Anti-phase coordination is challenging at fast tempi (Keller & Repp 2004), such as when playing the 'pahs' in an 'Oom-pah' band or when strumming the guitar upstrokes in ska and reggae music. Producing rhythms that form polyrhythmic ratios, as in sub-Saharan African drumming and minimalist music, is especially demanding, and it is perhaps for this reason that studies of the production of polyrhythms in ensembles are scarce. There is, however, a tradition of studying such unstable coordination modes in the context of dynamical systems approaches to visually guided interpersonal coordination (e.g., Oullier & Kelso 2009; Schmidt & Richardson 2008), and some of the principles discovered through that line of enquiry may apply to musical

ensemble coordination.

<C>Musical style

The size of purposeful vertical timing deviations varies with musical style and period. Certain subgenres of jazz and popular music are characterized by consistent asynchronies between the bassline and drums (hi-hat or ride cymbal). For example, consistent signed asynchronies of around 15-30 ms have been noted in various musical genres of 'groove' based music, including Afro-Cuban percussion music, funk, bebop, cool jazz, and hip-hop(Butterfield 2010; Hove et al. 2007).

It had been claimed that asynchronous timing between members of the rhythm section (e.g., bass and drums), and between the rhythm section and soloists, signals 'participatory discrepancies' (Keil 1995) that produce the rhythmic tension that is essential to the experience of groove (Doffman 2009; Prögler 1995). Participatory discrepancies may have functional as well as aesthetic benefits in ensemble performance. There is evidence that the 'wide-beat' provided by asynchronous onsets between nominally synchronous tones facilitates the stability of

movement timing during sensorimotor synchronization (Hove et al. 2007; Large & Palmer 2002).

<C>Leader-follower relations

Vertical timing deviations can serve to indicate or reinforce specific leader-follower relations between instrumental parts in ensembles. Melody lead is a relevant phenomenon that has received considerable attention. Early investigations of melody lead focused on the tendency for keystrokes associated with the melody part to occur ahead of accompanimental keystrokes in solo piano performances (Palmer 1989; Repp 1996), which may be a form of temporal precedence or a consequence of more forceful keystrokes in the melody part (Goebl 2001). Melody lead has also been observed across pianists in duos (Keller & Appel 2010).

Melody lead is not obligatory. Rasch (1988) observed melody lead when there was a clear hierarchical structuring of melody versus accompaniment (in woodwind and string trios), while polyphonic music was characterized by the lowest voice leading (in a recorder trio). Under some conditions, the tendency for a melody to lead is completely reversed. For example, the melody can lag behind the accompaniment when a ‘laid back’ feel is adopted in jazz (Friberg &

Sundström 2002) and in certain historical performance practices in chamber music (see Chapters 4 & 5 in this volume). Leader-follower relations are thus mutable.

Leadership roles may be more readily exchangeable in heterarchical groups such as piano duos (Blank & Davidson 2007; Williamon & Davidson 2002) than in larger mixed groups, such as string quartets, where first violin is commonly the designated leader. Nevertheless, even under such established hierarchies, the assignment of leader-follower roles may be dictated by the nature of the part being played by each instrument in a particular piece. Varni and colleagues (2010) found that the viola, whose part consisted of short notes with regular rhythm, assumed the lead when a string quartet was asked to exaggerate the expressive parameters of a performance of a Schubert quartet. In this case, the structure of the viola part was well suited to steering the rubato. This illustrates flexibility in leader-follower relations depending on the nature of the music's compositional structure. Additional factors that influence the identity of the musical leader include stereotypical instrument roles and the personality, status, and skill level of individual performers within the ensemble (Goodman 2002; Loehr & Palmer 2011; Maduell & Wing 2007).

<C>Intensity

Sound intensity can provide auditory cues to various aspects of musical expression in ensemble performance. Musical 'dynamics' in ensembles, as in solo performance, are conveyed through variations and contrasts in perceived loudness, which is proportional to sound intensity and (on acoustic instruments) the force used in generating the sound. Some fluctuations in dynamics are—like expressive timing deviations, but to a less consistent degree—systematically linked to structural features such as musical phrases, rhythmic groups, and metric locations through intensity-based accents and gradual intensity changes (crescendos and decresendos) (Keller 2012b; Palmer 1997). Other fluctuations in dynamics are random (see Chaffin et al. 2007; Keller et al. 2011) and influence the perceived spontaneity of a performance (Engel & Keller 2011).

Co-performers' dynamic variations influence ensemble cohesion by affecting the balance between instrumental parts in terms of their relative loudness. The nature of the ideal balance between parts depends on their relationship in the ensemble texture. In homophonic textures where compositional structure and the performers' expressive intentions demand clear hierarchical relations between melody and accompaniment, the relative intensity of parts must

be regulated in such a way that the accompaniment does not mask, or 'drown out', the melody. The principal tuba player of the Chicago Symphony Orchestra, Gene Pokorny [interviewed on 4 April 2012], states that 'whether it's a small or large ensemble, I try to be as chameleon-like as possible until I have a solo, and even then I'm still lending an ear to the other things that are happening.' In homophonic textures characterized by equality between instrumental parts, sound intensity must be matched across parts to ensure that they blend smoothly (notwithstanding differences in tessitura and timbre). Even in homogenous textures, however, the communication of musical expression may require subtle variations in intensity between parts. In polyphonic textures—as in a fugue, where instrumental parts are ostensibly independent but equal—individual voices may be required to take precedence and loom momentarily into the limelight.

The question of how ensemble performers vary the intensity of their sounds to achieve optimal blend is complicated by several factors. One issue is the so-called 'self-to-other ratio', which reflects the degree to which an individual can hear their own sounds amongst co-performers' sounds (Ternström 2003). The need to optimize this ratio is balanced against the need for parts to blend. Further complications stem from the fact that individual performers have preferred

intensity profiles that are reflected in the way in which they spontaneously shape the dynamics of a piece (Repp 2000). Individuals coming together in an ensemble must reconcile potential differences related to idiosyncrasies in their preferred intensity profiles. This necessitates interdependency in fluctuations in dynamics between parts in the ensemble (Lee & Schögler 2009; Papiotis et al. 2012). Goodman (2002) compared dynamics in a solo performance and an ensemble performance of the piano part from a cello sonata. Although fluctuations in dynamics were correlated across the two performances, the pianist played some passages louder and others softer when paired with the cellist than when performing alone. This suggests that expressive features of intensity profiles produced when playing solo may be dampened or exaggerated in response to contingencies arising through interaction with co-performers.

<C>Intonation, articulation, and timbre

The relationship between instrumental parts with respect to intonation, articulation, and timbre can make or break ensemble cohesion. Playing in tune with co-performers is a fundamental requirement. To do so, ensemble performers must adopt the same system of musical temperament, or, in other words, they must (tacitly) agree on how to adjust the ratios of

acoustic frequencies defining the pitch intervals that they produce (Mason, 1960). This choice is tightly constrained in ensembles containing at least one instrument that is tuned according to the equal temperament system (e.g., the piano) but more freedom can be exercised in other groups (e.g., choirs). Achieving good ensemble intonation in such groups is challenging because, in many musical traditions, the tuning of a particular pitch can vary according to the harmonic context in which it occurs (Ternström 2003).To complicate matters further, it is also often conventional to employ purposeful deviations in tuning for expressive purposes (Morrison & Fyk 2002). Ensemble performers need to be sensitive to how their co-performers are using such microtonal expressive devices in order to make appropriate compensatory adjustments (Papiotis et al. 2012). Performers playing a subsidiary role (e.g., second violins) may adjust their intonation to match those playing a primary role (first violins) in skilled ensembles, while bi-directional adjustments may be required in less skilled groups (Papiotis et al. 2011).

Articulation (i.e., the degree of overlap/separation of successive sounds) makes less obvious contributions to ensemble cohesion than expressive timing, intensity and intonation. Musicians may vary articulation to alter the expressive character of a performance. Short, separated ‘staccato’ sounds may imbue a passage with a lighthearted or agitated quality, while sustained,

overlapping ‘legato’ sounds may suggest a longing or solemn quality (Gabrielsson & Lindström 2001). Expressive goals may require articulation to be matched (e.g., all staccato) or mismatched (a composite of staccato and legato passages) across instrumental parts. Matching articulation entails aligning the perceptual onsets and offsets of sounds produced by co-performers. Such alignment is nontrivial because sounds produced on different instruments have different amplitude envelopes. Thus, differences in attack (or rise time), steady state, and decay portions of sounds need to be taken into account to achieve consistency in perceived articulation across the ensemble.

Like articulation, timbre (i.e., a tone’s color, determined by its spectral and temporal envelope) can have relatively subtle effects upon ensemble cohesion. Choral singing is an arena where timbre can nevertheless be decisive. Choral tone and blend are influenced by a host of acoustic factors related to the frequency spectra of singers’ voices (Ternström & Karna 2002). Singers alter the relative intensity of formants (i.e., peak amplitude regions in the voice’s frequency spectrum) depending on whether they intend to blend with others in a choir or to be heard as a soloist (Ternström 2003). Classically trained male singers, for example, increase the energy in a

high-frequency formant (around 2500-3500 Hz) when required to add brilliance and carrying power to the voice (Sundberg 2006).

<B>Visual cues

Body movements associated with music performance provide visual cues that, in combination with the auditory cues described above, facilitate the multimodal communication of musical structure and expressive intentions, as well as regulating the coordination between ensemble performers. A distinction can be drawn between so-called *instrumental movements*, which are directly related to the production of musical sounds, and *ancillary movements*, which are not technically necessary for sound production (Nusseck & Wanderley 2009). One function of ancillary movements is to generate kinesthetic feedback that aids the performer in regulating technical and expressive parameters of sound production. Thus, oscillatory body sway or foot tapping may regulate rhythm and tempo, and muscular tension may assist in controlling musical dynamics. Ancillary movements also provide visual communicative signals that disambiguate, reinforce, or augment auditory information related to musical structure and expression (see Davidson 1994, 2009). Movements of different body parts are well suited to represent

structural and expressive information at multiple hierarchical levels (Leman & Naveda 2010; Toiviainen et al. 2010; Williamon & Davidson 2002). Head nods, body sway, and limb gestures (e.g., shoulder, elbow, and hand movements) may thus influence the perception of phrase structure, rhythm, meter, and tempo, as well as intended musical character (Broughton & Stevens 2009; Dahl &Friberg 2007). With regards to the latter, the extent of ancillary movements typically increases with the intensity of musical sound expression (e.g., from deadpan, through normal, to exaggerated; Davidson, 1994).

In ensembles, ancillary movements provide visual cues that support the coordination of basic timing and expressive parameters between co-performers (Davidson & Malloch 2009). The coupling of ancillary movements between individuals, which can be recorded using motion capture technology (see Chapter 13), provides an index of the quality of interpersonal coordination in ensembles (Keller, 2008). Research with various types of small ensemble has found that head nods and body sway, along with gaze patterns, play a role in establishing and maintaining interpersonal synchrony and leader-follower relations (e.g., King & Ginsborg 2011; Moran forthcoming; Wöllner & Cañal-Bruland 2010). Eye contact is particularly important for the coordination of entrances and exits and for increasing awareness of co-performers'

spontaneous musical ideas (Blank & Davidson 2007; Clayton 2007a). Accordingly, eye contact occurs in anticipation of structural boundaries and transition points in the music (Clayton 2007b; Williamon & Davidson 2002). When co-performers are denied visual contact, increases in body sway serve as a compensatory mechanism to regulate each performer's timing (Keller & Appel 2010).

The prevalence of visual communicative cues in ensembles is affected by expertise and by familiarity with the music and with co-performers. Body sway and eye contact become increasingly frequent—and more synchronized across co-performers—over the course of rehearsals, and the use of such cues is more pronounced in performances than in rehearsals (Blank & Davidson 2007; Williamon & Davidson 2002). Furthermore, the range and frequency of communicative gestures is greater in duos comprising partners who are familiar and matched in expertise than for partners who are unfamiliar or mismatched in expertise (King & Ginsborg 2011). These findings suggest that familiarity gained during rehearsals leads to the emergence of a common repertoire of gestures that can function as a shared representation of the music. When co-performer familiarity is especially high, however, the perceived need for visual cues can be low. Pianists Katia and Marielle Labèque are sisters who shared a musical childhood and

have since then forged an international reputation as a duo. Marielle [interviewed on 28 May 2010] admits that 'We do very little cuing, but we are not a typical example', implying that ensembles whose members are less familiar with each other may rely to a greater degree on visual cues.

An obvious domain where the role of visual cues in interpersonal coordination comes to the fore is the partnership of a conductor and an ensemble. During performance, the conductor leads the ensemble by providing hand (and baton) movements, body gestures, and facial expressions that function dually to provide cues concerning the goal interpretation of the work—which is typically communicated during rehearsal—and to assist the ensemble members in achieving coordination in terms of basic timing and musical expression. The leadership role of the conductor is reflected in the fact that ensemble musicians' movements and sounds tend to lag behind the conductor's beat (D'Ausilio et al., 2012; Luck & Toiviainen 2006). The temporal relationship between conductor and ensemble members can, however, vary. It has been found that alignment of both movements and sounds with conductors' gestures is closer for gestures that are high in beat clarity, according to subjective perceptual judgments and objective kinematic measures (e.g., maxima in the vertical acceleration of the baton), than for gestures

that are low in clarity (Luck & Toiviainen 2006; Wöllner et al. 2012).

<A>Psychological mechanisms in ensemble performance

The foregoing section of this chapter described auditory and visual cues that co-performers use and modulate to achieve ensemble cohesion via interpersonal coordination at the levels of basic action timing and artful expression. The current section gives an overview of the psychological mechanisms that facilitate such temporal and expressive coordination (Figure 1, center). This overview, which builds on the musical joint action framework formulated by Keller (2008), takes into account the cognitive, motor, and social processes that enable ensemble musicians to coordinate their actions with high precision while maintaining the flexibility required to align expressive performance parameters. Two broad classes of mechanism can be distinguished: strategies, including those that occur 'offline' (prior to performance), and real-time ensemble skills that operate 'online' (during performance).

<B>Offline preparation and regulatory strategies

Musicians spend considerable time preparing for ensemble performance in order to gain familiarity with the music, to establish performance goals, to form performance plans that fulfil these goals, and to develop shared performance cues that aid the realization of these plans. These tasks are accomplished through a combination of individual private practice and collaborative group rehearsal.Private practice can facilitate familiarity with co-performers' parts through the study of musical scores and sound recordings. Collaborative rehearsal is then typically geared towards establishing a shared performance goal, i.e., a unified conception of the ideal integrated ensemble sound (Keller 2008; Williamon & Davidson 2002). The degree to which a detailed performance goal can be truly shared across ensemble members, however, varies as a function of the musical context. Members of a symphony orchestra, for instance, do not necessarily know the intricacies of each part in the ensemble texture; instead, the conductor functions as a repository of the global performance goal. Furthermore, in freely improvised music, co-performers eschew fully preconceived goals in favour of transient shared goals that evolve spontaneously through mimicry and other processes during performance (Schoegler 1999-2000; Smith & Dean 1997).

When shared performance goals are strategically pursued during rehearsal, ensemble musicians enter into a process of becoming familiar with one another's parts and the manner in which these parts will be played. This process primarily entails nonverbal communication through body movements and musical sounds, though a limited amount of verbal communication does take place (Price & Byo 2002; Williamon & Davidson 2002). As musicians coming together to rehearse a piece bring their own preconceptions of the music, they must find a way to reach consensus on how expressive performance parameters should be modulated (Ginsborg et al. 2006). A mixture of social, conventional, and pragmatic considerations govern this process of negotiation. Social factors—including personality, pre-existing interpersonal relationships, verbal and nonverbal communication styles, and gender and instrument stereotypes—are relevant to the extent that they influence the effectiveness of information exchange during rehearsal (Blank & Davidson 2007; Davidson & Good 2002; Davidson & King 2004; Ginsborg et al., 2006; Goodman 2002; Williamon & Davidson 2002).

In professional ensembles, the impact of social factors on communication is moderated by well-established conventions concerning matters of organization, administration, repertoire choice,

and musical interpretation (Blank & Davidson 2007). Chicago Symphony horn player David Griffin [interviewed on 5 April 2012] describes an optimal scenario in an orchestral section:

> What goes on a lot in our horn section is non-verbal communication. We have two excellent principal horn players, and I think one of the reasons why they sound so good is because we have section players who put them on a pedestal. We make our playing complement what they do just by listening to what they're doing, and in the rehearsal it's mostly about the section getting together and trying to sound as one.

Once shared performance goals are consolidated, they reside in each individual's memory as idealized mental representations of the sounds constituting a musical piece. Co-performers thus come to co-represent elements of each other's parts (Keller, 2008; Loehr & Palmer 2011; Sebanz et al. 2006).

More generally, performance goals embody, to varying degrees, a performer's intentions and expectations about expressive parameters of (1) his or her own sound, (2) co-performers'

sounds, and (3) the overall ensemble sound. With such goals in mind, musicians develop performance plans that guide the motor processes involved in translating the goal representations into body movements that are appropriate for generating the ideal sound. As the parts played by individual musicians in ensembles are often complementary (rather than identical), performance plans will differ across co-performers. Ensemble musicians therefore develop systems of shared performance cues to regulate and coordinate their actions (Ginsborg et al., 2006).

Performance cues are features of the music (e.g., intensity changes, phrase boundaries, and the location of breaths) that the musician prospectively chooses to attend to during performance in order to ensure that things take place as planned (Chaffin & Logan 2006; Ginsborg et al., 2006). The selected features provide landmarks in a mental map that reflects the hierarchical organization of sections and subsections in a piece's formal structure. Hierarchies of performance cues thus serve as retrieval schemes that allow performers to deal with the real-time demands of performance by utilizing domain-specific expert memory processes (Lehmann & Ericsson 1998). In ensembles, shared performance cues remind co-performers of shared performance goals and link individual performance plans into a common scheme that can be

used to regulate the interplay between musicians. The successful execution of shared goals and plans is mediated during performance by a set of online ensemble skills.

<B>Online ensemble skills

Keller (2008) proposed that three core cognitive-motor skills interact to determine the quality of real-time interpersonal coordination in musical ensemble performance. These cognitive-motor skills allow a performer to anticipate, attend, and adapt to auditory and visual cues generated by co-performers. Each skill comprises mechanisms that operate automatically and mechanisms that require conscious control. These mechanisms overlap with those that support interpersonal coordination more generally (Knoblich et al. 2011).

<C>Adaptive mechanisms

Adaptive mechanisms enable ensemble performers to react to intentional and unintentional variations in each other's action timing, as well as to variations in intensity, intonation, articulation, and timbre. Adaptive timing is the most fundamental and assiduously studied

process. Mutual temporal adaptation in ensembles entails adjusting the timing of one's movements in order to maintain synchrony in the face of small random irregularities and expressively motivated deviations in local tempo, as well as larger tempo changes and errors disrupting rhythm. Adaptive timing is mediated by temporal error correction mechanisms that enable internal timekeepers in co-performers' brains (i.e., oscillations of neural populations at timescales compatible with musical meter) to remain coupled under such variable conditions (Large 2008; Repp & Keller 2008; Vorberg & Wing 1996). These mechanisms support neural entrainment and behavioral synchronization with a range of auditory and visual signals, though they may be most effective in the auditory modality (Hove et al. 2010) and at moderate to slow tempi (Repp et al. 2012).

Two separate error correction mechanisms subserve adaptive timing in the performance of pulse-based ensemble music (Repp 2006a). One mechanism, termed *phase correction,* is an automatic and obligatory process that adjusts the way in which the sequence of pulses generated by an internal timekeeper in one performer is aligned against a sequence of pulses generated by a timekeeper in a co-performer. Phase correction supports precision in basic temporal coordination by keeping vertical timing deviations within an optimal zone. The other

mechanism, *period correction*, involves consciously controlled adjustments to the duration of timekeeper intervals, and is invoked when a performer intentionally adapts to tempo changes produced by a co-performer. Period correction may facilitate the flexibility that is required for co-performers to match expressively motivated horizontal timing deviations through deliberate imitation.

In addition to temporal error correction, adaptive timing involves a form of temporal assimilation whereby co-performers automatically mimic small fluctuations in the timing of each other's actions. Such mutual temporal assimilation is a phenomenon that has been observed in experimental tasks requiring piano duet performance (Goebl & Palmer 2009; Loehr & Palmer 2011) and dyadic sensorimotor synchronization (Konvalinka et al. 2010; Merker et al. 2009; Nowicki et al. in press). In these contexts, it has been found that one individual tends to copy timing variations produced by an interaction partner. This process takes place on a millisecond timescale, and presumably supports basic coordination rather than expressive goals. Mutual temporal assimilation may be a form of non-conscious behavioral mimicry that facilitates ensemble cohesion by making multiple individuals sound collectively as one (Nowicki et al. in

press). Non-conscious mimicry and deliberate imitation may also characterize adaptation to co-performers' variations in intensity and intonation (Maduell & Wing 2007; Ternström 2003).

<B>Attention

A more cognitively advanced ensemble skill is a form of divided attention that involves concurrently paying attention to one's own actions (high priority) and those of others (lower priority) while monitoring the overall ensemble sound. This attentional strategy, which has been termed 'prioritized integrative attending' (Keller 2001), is assumed to facilitate ensemble cohesion by allowing musicians to adjust their performances based on the online comparison of mental representations of the ideal sound (i.e., the performance goal) and incoming perceptual information about the actual sound. Performers are thus able to deal with changes in the momentary demands of their own parts and the relationship between their own and others' parts in terms of timing, intensity, intonation, articulation, and timbre.

It has been proposed that the dynamic allocation of attentional resources required for prioritized integrative attending is guided by metric frameworks (Keller 2001, 2008). Specifically,

in the context of pulsed music, the coupling of internal timekeepers to periodicities associated with the music's metric structure may provide a resource allocation scheme for modulating the amount of attention that is available at a particular point in time (by affecting autonomic arousal) and the amount of attention that is actually invested at this time (by affecting the intensity of attentional focus). These modulations ensure that attention is deployed in a manner that is conducive to the flexibility required to monitor multiple levels of the musical texture concurrently (Keller 2001).

Prioritized integrative attending is cognitively demanding (Keller & Burnham 2005) and it is quite likely applied only sporadically during ensemble performance. For much of the time, co-performers may rely upon passive, pre-attentive processes that allow multiple streams of information to be monitored in parallel. Katia Labèque touches on this form of automaticty: I'm listening to my sister much more when I practice at home than when I am on stage. I of course listen to what she does—I know perfectly what she does—but we do not consciously plan the interaction. We are able to react spontaneously to whatever happens.' The need to actively shift attention between streams, or to integrate their contents effortfully, may arise mainly in relation to shared performance cues (e.g., when coordinating entries, spontaneous tempo

changes, and expressive parameters). The focus of attention may broaden in the vicinity of these locations to accommodate visual cues provided by ancillary body movements.

A challenge for ensemble performers is to balance the tasks of attending to expression in their own playing and in others' playing. Chicago Symphony horn player Dan Gingrich [interviewed on 5 April 2012] characterizes the problem as follows:

> You have to be aware of context, otherwise the ensemble really isn't an ensemble. There has to be that cooperation and a sense of what else is going on, from sound quality to dynamic levels, and whether the tempo leans forward or leans back. Those are just basic things, but there has to be an awareness, otherwise the music really won't be of a superior level. On the other hand, you can't get so lost in what else is going on that you neglect the concentration and the focus that needs to go into what you are responsible for yourself.

<C> Anticipatory mechanisms

Anticipatory cognitive-motor mechanisms enable ensemble performers to plan the production of their own sounds and to generate online predictions about the upcoming sounds of co-performers (Keller 2008, 2012a). These predictions evolve along two routes (Phillips-Silver & Keller 2012). One route supports automatic expectancies that are triggered by the perception of auditory and visual cues associated with actual sounds and body movements. These expectancies, which evolve at short time-scales (e.g., anticipating the next tone or end point of a conductor's beat gesture), are driven by processes that cause perceptual and motor regions of the brain to resonate within coming auditory and visual information. The other route is characterized by the deliberate use of mental imagery to anticipate the future course of co-performers' actions. Such anticipatory imagery entails running internal simulations that are experienced, in phenomenological terms, as auditory and motor images of musical sound sequences and related movements that unfold over relatively long time-scales (e.g., body sway linked to metric bars or phrase structure). It is through anticipatory musical imagery that ensemble musicians activate internal representations of shared goals, plans, and cues during performance. The importance of this process is underscored by the fact that some elite ensemble musicians claim to imagine sounds produced by co-performers even during private

practice (Trusheim 1993).

Anticipatory musical imagery is an advanced cognitive-motor skill that is refined through musical experience (Keller 2012a). The mechanisms underpinning this learning involve the training of 'internal models' instantiated in the performer's central nervous system.Internal models represent sensorimotor associations between motor commands that issue from the brain and the sensory experience of bodily states and events in the immediate environment (Wolpert et al. 2003). 'Forward models' represent the causal relationship between motor commands and their effects on the body and environment. 'Inverse models' represent transformations from intended action outcomes (sounds, in the case of music) to the motor commands that produce them. Instances where instrumentalists 'sing along' with their own performances—as Canadian pianist Glenn Gould was prone to do—may arise due to uninhibited output of internal models.

Separate classes of forward and inverse models are harnessed to simulate one's own and others' actions in advance of their production (Keller 2008). The coupling of 'own' and 'other' internal models facilitates fluent interpersonal coordination in ensembles by allowing error

correction to be applied on the basis of anticipated relations between one's own and others' actions rather than in response to the perception of actual discrepancies. Internal models may operate at different hierarchical levels and timescales (cf. Pacherie 2008; Shaffer 1984), with own-other couplets devoted to long-range performance goals and plans (concerning musical phrases and ancillary body sway, for example) and short-range goals and plans (pertaining to instrumental movements and individual sounds or brief sequences).

According to the current conceptualization, the quality of ensemble cohesion is constrained by co-performers' abilities to use anticipatory imagery to predict basic structural properties and expressive features of each other's productions. Mental imagery steers action simulation; internal models provide the motor that drives it. The results of studies investigating the relationship between behavioural indices of auditory imagery, temporal prediction, and interpersonal coordination support this view. One study found that individual differences in the coordination of keystrokes and body sway between pianists in duos were positively correlated with performance on a separate task designed to measure the vividness of anticipatory auditory imagery (Keller & Appel 2010). Another series of experiments found that the accuracy and precision of dyadic sensorimotor synchronization were dependent upon the paired individuals'

temporal prediction abilities (Pecenka & Keller 2011), which, in turn, were correlated with their auditory imagery abilities (Pecenka & Keller 2009).

<B>Factors that affect ensemble strategies and skills

A number of factors may influence ensemble cohesion by affecting preparation strategies and online ensemble skills. These factors are potential sources of individual differences in basic coordination and the communication of musical expression during ensemble performance. Factors that are specific to the musical context and are, in principle, controllable by performers (e.g., familiarity with a piece) offer soft constraints, whereas factors that relate to domain-general and relatively fixed characteristics of individual performers (e.g., personality) may impose hard constraints (Figure 1, right). The degree to which constraints on coordination can be overcome, and the use of strategies to enhance the communication of musical expression, presumably depends upon the performers' level of musical skill and experience.

<C>Context-specific factors

The successful application of ensemble strategies and skills is obviously constrained by knowledge of the structure of the music and familiarity with the expressive intentions and stylistic tendencies of co-performers. These two varieties of knowledge—structural and personal—serve different functions and can have dissociable effects on ensemble coordination (Uhlig et al. 2012). Knowledge of musical structure is acquired through collaborative rehearsal and the private study of co-performers' parts in scores and recordings. Structural knowledge is essential to the integrity of shared performance goals. Furthermore, such knowledge allows a performer to adjust basic and expressive features in their performance plans in accordance with the role of their part in the ensemble texture, as well as to utilize performance cues related to others' parts. Structural knowledge is therefore well suited to the control of anticipatory processes at relatively long time-scales corresponding to important structural locations associated with high-level metric units and phrases.

Knowledge of co-performers' expressive intentions and familiarity with their stylistic tendencies (i.e., idiosyncrasies in the use of expressive devices) allows features of their performances to be anticipated and accommodated in one's own performance plans. Acquiring such knowledge entails a performer learning—through the calibration of internal models—to simulate another's

action style. The results of studies examining perception and action in relation to recordings of a musician's own versus others' performances suggest that higher fidelity simulations lead to increased perceptual sensitivity to rhythmic timing and better sensorimotor synchronization (Keller et al., 2007; Repp & Keller 2010). Knowledge of co-performers' stylistic idiosyncrasies thus impacts upon anticipation at short (millisecond) time-scales. Such intimate knowledge can also have social implications that affect performance expression. Katia Labèque [interviewed on 28 May 2010] hints at the immediacy with which she can sense her sister Marielle's psychological state: 'I know when I start playing if she's nervous, if she likes the piano, if she feels well. I know it—I feel it—immediately.'

Another relevant factor concerns the performers' intentions about the interpersonal alignment of basic and expressive sound parameters. As discussed earlier, ensemble performers typically favor artful deviations in vertical timing over strict synchrony. In extreme cases, performers may attempt to avoid entrainment (e.g., in some polymetric Western contemporary Art music and freely improvised music; see Smith & Dean 1997), presumably by attending exclusively to their own part to minimize adaptive timing. Related phenomena occur in Indian raga (Clayton, 2007a)

and in multi-religious rituals when separate groups wish to exert their identity (Lucas et al. 2011).

At the other extreme, performers in Western classical chamber music ensembles listen attentively to other parts to maximize mutual adaptation, and use regulatory strategies to adjust their manner of playing to facilitate ensemble cohesion. One method for doing so is to employ coordination 'smoothers', that is, behavioral modifications that simplify interpersonal coordination by making actions easier to predict(Vesper et al. 2010). Musical coordination smoothers come in auditory and visual flavors. An example of an auditory smoother is when musicians dampen their use of expressive devices, such as rubato, during ensemble performance relative to when they perform their part alone (Goodman, 2002). Another example occurs when an ensemble leader sharpens the contrast in rhythmic durations in order to communicate clear expressive intentions and unambiguous synchronization targets (Marchini et al. 2012).Visual smoothers are exemplified by the ensemble leader exaggerating their instrumental movements and simplifying their ancillary movements (Goebl & Palmer 2009; Glowinski et al., 2010).

<C>Domain-general factors

The impact of social-psychological factors, including personality, upon interpersonal interactions during joint preparation for ensemble performance is well documented, but the degree to which such factors directly constrain online ensemble skills is an open question. Several considerations are worth noting in this regard.

Personality has been found to be associated with the way in which people move their bodies to music. Individuals who are high in extroversion, for example, tend to produce fast and expansive movements while individuals high in neuroticism produce constricted and jerky movements (Luck et al. 2010). Other work has linked personality stereotypes to instrument and ensemble choice (MacLellan 2011). Aspects of personality such as locus of control (the degree to which life events are perceived to be a consequence of one's own actions) and social competence may influence whether a performer is better suited to play the role of soloist or accompanist by affecting the degree to which he or she adapts to other's action timing (Fairhurst et al. 2012; Schmidt & Richardson 2008).

Personality may therefore predispose individuals to play different roles in different types of ensembles, depending on their action style, motivation, self-focus, and ability to attend to multiple sources of information simultaneously. This latter ability may vary more specifically as a function of general intelligence, which determines (among other things) an individual's capacity to manipulate information in working memory (Cowan et al. 2005).General intelligence may also influence basic coordination through its relationship with the ability to produce a steady tempo (Ullén et al. 2008), while emotional intelligence may affect the ability to recognize cues to musical expression (Resnicow et al. 2004).

The related concept of empathy—understanding others' thoughts and feelings—has been linked to music perception (Leman 2007), and may also have implications for ensemble performance (Bablioni et al., 2011). A neurophysiological study of piano duos found that the degree to which one pianist represented the other's part in their motor system was correlated with scores on a questionnaire assessing empathy (Novembre et al. 2012). Another study revealed that perspective taking—a dimension of empathy—was correlated with objective behavioral measures of the degree to which an individual predicted event timing during sensorimotor synchronization with auditory sequences containing tempo changes (Pecenka &

Keller 2011). These results suggest that empathy covaries with ensemble skills due to a link with anticipatory mechanisms related to action simulation.

Additional hard constraints that may affect ensemble cohesion include pre-existing similarities between co-performers in terms of their stylistic tendencies. These tendencies, and co-performers' action styles more generally, may influence the ability to align expressive performance parameters across ensemble members (Keller et al., 2007). For example, it has been shown that pianists who have similar preferred tempi in solo performance are more closely coordinated and display more mutual temporal adaptation during duo performance (Loehr & Palmer, 2011).

An individual's action style is determined in part by biomechanical constraints related to physical body characteristics. Anthropometric factors such as body mass, shoulder width, and limb length influence spontaneous motor tempo and preferred beat rates when listening to music (Todd et al. 2007; van Noorden & Moelants 1999). This suggests that the physical characteristics of ensemble members (and presumably their instruments) may affect the performance tempi that they find comfortable, and inter-individual differences in physical

characteristics could therefore have subtle effects upon ensemble cohesion. Experimental manipulations of physical properties of objects in studies of unintentional coordination via the visual modality (e.g., in pairs of individuals swinging hand-held pendulums or seated in rocking chairs) show that mismatches in comfortable movement tempo induce systematic phase relations in interpersonal coupling (the faster individual leads) (Schmidt & Richardson 2008). Although ensemble musicians engaged in carefully planned coordination via multimodal auditory and visual cues is a different matter, it may be the case that individual differences in preferred tempo produce related effects (Loehr & Palmer 2011). Thus, differences in real or imagined physical constraints on movement tempo across co-performers may assist in sustaining the vertical timing deviations (see Figure 2) that contribute to the vitality of expressive ensemble music.

<A>Conclusions

The current chapter addressed the question of how ensemble musicians achieve precision in basic interpersonal coordination without sacrificing the flexibility required for inter-individual co-variation in expressive performance parameters. As in solo performance, ensemble

musicians signal their expressive intentions via auditory and visual cues, but in ensembles these behavioral cues are imbued with additional communicative functions related to ensemble cohesion. The fulfillment of these functions is enabled by psychological mechanisms including preparatory and regulatory strategies, and real-time ensemble skills that facilitate mutual adaptation, attention, and anticipation. The operation of these mechanisms is subject to cognitive, motor, and social constraints, which may nevertheless be surmounted to varying degrees by plasticity in ensemble skills and by the performer's motivation to develop compensatory strategies (e.g., studying musical scores and practicing with sound recordings, at a range of tempi, and with varying expression). Indeed, while automatic processes may be sufficient for basic interpersonal coordination, the interpersonal alignment of expressive performance parameters most likely relies upon effortful processes. This implies that the large amount of experience that is necessary to attain mastery in solo performance must be supplemented by a complementary regimen of specialized training dedicated to the development of strategies and skills for artistry in musical ensemble performance.

<A>References

Ashley, R. (2002). Do[n't] change a hair for me: The art of jazz rubato. *Music Perception, 19*, 311-322.

Babiloni, C., Vecchio,F., Infarinato, F., Buffo, P., Marzano, N., Spada, D., Rossi, S., Bruni, I., Rossini, P. M., & Perani, D. (2011). Simultaneous recording of electroencephalographic data in musicians playing in ensemble. *Cortex, 47*, 1082-1090.

Blank, M., & Davidson, J. W. (2007). An exploration of the effects of musical and social factors in piano duo collaborations. *Psychology of Music, 35*, 213-230.

Broughton, M., & Stevens, C. (2009). Music, movement and marimba: an investigation of the role of movement and gesture in communicating musical expression to an audience. *Psychology of Music, 37*, 137-153.

Butterfield, M. (2010). Participatory discrepancies and the perception of beats in jazz. *Music Perception, 27*, 157–176.

Chaffin, R., Lemieux, A., & Chen, C. (2007). “It’s different each time I play”: Spontaneity in highly prepared musical performance. *Music Perception*, *24*, 455-472.

Chaffin, R., & Logan, T. (2006). Practicing perfection: How concert soloists prepare for performance. *Advances in Cognitive Psychology, 2*, 113-130.

Clarke, E. F., (1988). Generative principles in music performance. In J. Sloboda (Ed.). *Generative processes in music* (pp. 1–26). Oxford, UK: Clarendon Press.

Clayton, M. (2007a).Observing entrainment in music performance: video-based observational analysis of Indian musicians' tanpura playing and beat marking. *Musicae Scientiae,11*, 27-59.

Clayton, M. (2007b). Time, gesture and attention in a khyal performance. *Asian Music, 38*, 71-96.

Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology, 51*, 42-100.

Dahl, S., & Friberg, A. (2007). Visual perception of expressiveness in musicians' body movements. *Music Perception, 24*, 433-454.

Davidson, J. W. (1994). What type of information is conveyed in the body movements of solo musician performers? *Journal of Human Movement Studies, 6*, 279-301.

Davidson, J. W. (2009). Movement and collaboration in musical performance. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (pp. 364-376), Oxford, Oxford University Press.

Davidson, J. W., & Good, J. M. M. (2002). Social and musical co-ordination between members of a string quartet: An exploratory study. *Psychology of Music, 30*, 186–201.

Davidson, J. W., & King, E. C. (2004). Strategies for ensemble practice. In A. Williamon (Ed.), *Enhancing musical performance* (pp. 105–122). Oxford, UK: Oxford University Press.

Davidson, J. W., & Malloch, S. (2009). Musical communication: The body movements of performance. In S. Malloch & C. Trevarthen, (Eds.), *Communicative musicality: Exploring the basis of human companionship* (pp. 565–583). Oxford: Oxford University Press.

Doffman, M. (2009). Making it groove! Entrainment, participation and discrepancy in the 'conversation' of a jazz trio. *Language and History, 52*, 130–147.

D'Ausilio A., Badino L., Tokay S., Aloimonos Y., Li Y., Craighero L., Canto R., Fadiga L. (2012). Leadership in orchestra emerges from the causal relationships of movement kinematics. *PLoS ONE 7(5):* e35757.

Engel, A., & Keller, P. E. (2011). The perception of musical spontaneity in improvised and imitated jazz performances. *Frontiers in Psychology, 2:83*. doi: 10.3389/fpsyg.2011.00083

Fairhurst, M. T., Janata, P., &Keller, P. E. (submitted). *Leading the follower: an FMRI investigation of dynamic cooperativity and leader-follower strategies in synchronization with an adaptive virtual partner*.

Friberg, A., & Sundström, A. (2002). Swing ratios and ensemble timing in jazz performance: Evidence for a common rhythmic pattern. Music Perception, 19, 333–349.

Gabrielsson, A. (2003). Music performance research at the millennium. *Psychology of Music, 31*, 221–272.

Gabrielsson, A., & Lindstrom, E. (2001). The influence of musical structure on emotional expression. In P. Juslin & J. Sloboda (Eds.),*Music and emotion: Theory and research* (pp. 223–248). Oxford: Oxford University Press.

Ginsborg, J., Chaffin, R., & Nicholson, G. (2006). Shared performance cues in singing and conducting: A content analysis of talk during practice. *Psychology of Music, 34*, 167-192.

Glowinski, D., Camurri, A., Chiorri, C., Coletta, P., Schenone, A., &Volpe G. (2010). Multi-scale entropy analysis of dominance in social creative activities. *Proceedings of the ACM Multimedia International Conference*. Firenze: ACM Press.

Goebl, W. (2001). Melody lead in piano performance: Expressive device or artifact? *Journal of the Acoustical Society of America, 110*, 563–572.

Goebl, W., & Palmer, C. (2009). Synchronization of timing and motion among performing musicians. *Music Perception, 26*, 427-438.

Goodman, E. (2002). Ensemble performance. In J. Rink (Ed.), *Musical performance: A guide to understanding* (pp. 153-167). Cambridge, UK: Cambridge University Press.

Hove, M. J., Keller, P. E., & Krumhansl, C. L. (2007). Sensorimotor synchronization with chords

containing tone-onset asynchronies: The role of P-centers.*Perception and Psychophysics, 69*, 699-708.

Hove, M. J., Spivey, M. J., & Krumhansl, C. L. (2010). Compatibility of motion facilitates visuomotor synchronization. *Journal of Experimental Psychology: Human Perception and Performance, 36*, 1525-1534.

Janata, P., Tomic, S. T.,& Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *Journal of Experimental Psychology: General, 141*, 54–75.

Keil, C. (1995). The theory of participatory discrepancies: A progress report. *Ethnomusicology, 39*, 1-19.

Keller, P. E. (2001). Attentional resource allocation in musical ensemble performance. *Psychology of Music, 29*, 20-38.

Keller, P. E. (2008). Joint action in music performance. In F. Morganti, F., A. Carassa, & G. Riva (Eds.), *Enacting intersubjectivity: A cognitive and social perspective on the study of interactions* (pp. 205-221). Amsterdam: IOS Press.

Keller, P. E. (2012a). Mental imagery in music performance: underlying mechanisms and potential benefits. *Annals of the New York Academy of Sciences, 1252*, 206–213.

Keller, P. E. (2012b). What movement force reveals about cognitive processes in music performance. In A. Mornell (Ed.), *Art in motion II* (pp. 115-153). Vienna/Frankfurt: Peter Lang.

Keller, P. E., & Appel, M. (2010). Individual differences, auditory imagery, and the coordination of body movements and sounds in musical ensembles. *Music Perception, 28*, 27-46.

Keller, P. E., & Burnham, D. K. (2005). Musical meter in attention to multipart rhythm. *Music Perception, 22*, 629-661.

Keller, P.E., & Repp, B.H. (2004). When two limbs are weaker than one: Sensorimotor syncopation with alternating hands. *Quarterly Journal of Experimental Psychology, 57A*, 1085-1101.

Keller, P. E., & Repp, B. H. (2005). Staying offbeat: Sensorimotor syncopation with structured and unstructured auditory sequences.*Psychological Research, 69*, 292-309.

Keller, P. E., Knoblich, G., & Repp, B. H. (2007). Pianists duet better when they play with themselves: On the possible role of action simulation in synchronization. *Consciousness and Cognition, 16*, 102-111.

Keller, P. E., Weber, A., & Engel, A. (2011). Practice makes too perfect: Fluctuations in loudness indicate spontaneity in musical improvisation. *Music Perception, 29*, 107–112.

King, E. C., & Ginsborg, J. (2011). Gestures and glances: Interactions in ensemble rehearsal. In

A. Gritten & E. King (Eds.), *New perspectives on music and gesture* (pp. 177–201). Aldershot: Ashgate Press.

Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological research on joint action: Theory and data. In B. Ross (Ed.), The psychology of learning and motivation (pp. 59-101). Burlington: Academic Press.

Konvalinka, I., Vuust, P., Roepstorff, A., & Frith, C. D. (2010). Follow you, follow me: Continuous mutual prediction and adaptation in joint tapping. *Quarterly Journal of Experimental Psychology, 63*, 2220-2230.

Large, E. W. (2008). Resonating to musical rhythm: Theory and experiment. In Simon Grondin (Ed.), *The psychology of time* (pp. 189-231). West Yorkshire: Emerald.

Large, E. W., & Palmer, C. (2002). Perceiving temporal regularity in music. *Cognitive Science, 26*, 1-37.

Large, E. W., Fink, P., & Kelso, J. A. S. (2002). Tracking simple and complex sequences. *Psychological Research, 66*, 3–17.

Lee, D., &Schögler, B. (2009). Tau in musical expression. In S. Malloch & C. Trevarthen, (Eds.), *Communicative musicality: Exploring the basis of human companionship* (pp. 83-104). Oxford: Oxford University Press.

Lehmann, A. C., & Ericsson, K. A. (1998). Preparation of a public piano performance: The relation between practice and performance. *Musicae Scientiae, 2*, 69-94.

Leman, M. (2007). *Embodied music cognition and mediation technology*. Cambridge, MA: MIT Press.

Loehr, J. D., & Palmer, C. (2011). Temporal coordination between performing musicians. *Quarterly Journal of Experimental Psychology, 64*, 2153-2167.

Loehr, J. D., Large, E. W., & Palmer, C. (2011). Temporal coordination and adaptation to rate change in music performance. *Journal of Experimental Psychology: Human Perception and Performance, 37*, 1292-1309.

Lucas, G., Clayton, M.,&Leante, L. (2011). Inter-group entrainment in Afro-Brazilian Congado ritual. *Empirical Musicology Review,6*, 75-102.

Luck, G., & Toiviainen, P. (2006). Ensemble musicians' synchronization with conductors' gestures: an automated feature-extraction analysis. *Music Perception, 24*, 189-200.

Luck, G., Saarikallio, S., Burger, B., Thompson, M. R., & Toiviainen, P. (2010). Effects of the Big Five and musical genre on music-induced movement. *Journal of Research in Personality, 44*, 714–720.

MacLellan, C. R. (2011). Understanding your band, orchestra, and choir students. personality

similarities and differences and what they mean for you. *Music Educators Journal, 97*, 37-42.

Madison, G., & Merker, B. (2005). Timing of action during and after synchronization with linearly changing intervals. *Music Perception, 22*, 441-459.

Maduell, M., & Wing, A. M. (2007). The dynamics of ensemble: The case for flamenco. *Psychology of Music, 35*, 591–627.

Marchini, M., Papiotis, P., & Maestre, E. (2012). Timing synchronization in string quartet performance: a preliminary study. In *Proceedings of the International Workshop on Computer Music Modeling and Retrieval* (CMMR12). London, UK.

Mason, J. A. (1960). Comparison of solo and ensemble performances with reference to Pythagorean, Just, and Equal-Tempered intonations. *Journal of Research in Music Education, 8*, 31–38.

Merker, B., Madison, G., & Eckerdal, P. (2009). On the role and origin of isochrony in human rhythmic entrainment. *Cortex, 45*, 4-17.

Moore, G., & Chen, J. (2010). Timings and interactions of skilled musicians. *Biological Cybernetics, 103*, 401-414.

Moran, N. (forthcoming). Social co-regulation and communication in North Indian duos.InM.R.L. Clayton, B. Dueck,& L. Leante (Eds.), *Experiencing meaning in music performance*.

Morrison, S. J., & Fyk, J. (2002). Intonation. In R. Parncutt & G. McPherson (Eds.), *Thescience and psychology of music performance: Creative strategies for teaching and learning* (pp. 183–197). New York: Oxford University Press.

Leman, M., & Naveda, L. (2010). The spatiotemporal representation of dance and music gestures using Topological Gesture Analysis (TGA). *Music Perception, 28*, 93–111.

Novembre, G., Ticini, L.F., Schütz-Bosbach, S., & Keller, P.E. (2012). Distinguishing self and other in joint action. Evidence from a musical paradigm. *Cerebral Cortex*.doi: 10.1093/cercor/bhr364

Nowicki, L., Prinz, W., Grosjean, M., Repp, B. H., & Keller, P. E. (in press). Mutual adaptive timing in interpersonal action synchronization. *Psychomusicology: Music, Mind, and Brain.*

Nusseck, M., & Wanderley, M. M. (2009). Music and motion – How music-related ancillary body movements contribute to the experience of music. *Music Perception*, *26*, 335–353.

Oullier, O., & Kelso, J. A. S. (2009). Coordination from the perspective of social coordination dynamics. In R. Meyers (Ed.), *Encyclopedia of complexity and systems science* (pp. 8198–8212). Berlin: Springer-Verlag.

Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition, 107*, 179-217.

Palmer, C. (1989). Mapping musical thought to musical performance. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 331–346.

Palmer, C. (1997). Music performance. *Annual Review of Psychology, 48*, 115–138.

Papiotis, P., Marchini, M., Maestre, E., & Perez, A. (2011). Measuring ensemble synchrony through violin performance parameters: a preliminary progress report. In *Proceedings of the 4th International ICST Conference on Intelligent Technologies for Interactive Entertainment* (INTETAIN 2011). Genoa, Italy.

Papiotis, P., Marchini, M., & Maestre, E. (2012), Computational analysis of solo versus ensemble performance in string quartets: Dynamics and intonation. In *Proceedings of the 12th International Conference of Music Perception and Cognition* (ICMPC12). Thessaloniki, Greece.

Pecenka, N., & Keller, P. E. (2009). Auditory pitch imagery and its relationship to musical synchronization. *Annals of the New York Academy of Sciences, 1169*, 282-286.

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

Phillips-Silver, J., & Keller, P.E. (2012). Searching for roots of entrainment and joint action in early musical interactions.*Frontiers in Human Neuroscience, 6:26*. doi: 10.3389/fnhum.2012.00026

Povel, D.J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception, 2*, 411-440.

Price, H. E., & Byo, J. L. (2002). Rehearsing and conducting. In R. Parncutt & G. E. McPherson. (Eds.), *The science and psychology of musical performance: Creative strategies for teaching and learning* (pp. 335–351). Oxford: Oxford University Press.

Prögler, J. A. (1995). Searching for swing: Participatory discrepancies in the jazz rhythm section. *Ethnomusicology, 39*, 21–54.

Rankin, S. K., Large, E. W., & Fink. P. W. (2009). Fractal tempo fluctuation and pulse prediction. *Music Perception, 26*, 401–413.

Rasch, R. A. (1988). Timing and synchronization in ensemble performance. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition*(pp. 70-90). Oxford, UK: Clarendon Press.

Repp, B. H. (1992). Diversity and commonality in music performance: An analysis of timing microstructure in Schumann's "Träumerei". *Journal of the Acoustical Society of America, 92*, 2546–2568.

Repp, B. H. (1996). Patterns of note onset asynchronies in expressive piano performance. *Journal of the Acoustical Society of America, 100*, 3917–3932.

Repp, B. H. (2000). Pattern typicality and dimensional interactions in pianists' imitation of expressive timing and dynamics. *Music Perception, 18*, 173–211.

Repp, B. H. (2006a). Musical synchronization. In E. Altenmüller, M. Wiesendanger, & J. Kesselring (Eds.), *Music, motor control, and the brain* (pp. 55–76). Oxford, UK: Oxford University Press.

Repp, B. H. (2006b). Rate limits of sensorimotor synchronization. *Advances in Cognitive Psychology, 2*, 163-181.

Repp, B. H., & Keller, P. E. (2004). Adaptation to tempo changes in sensorimotor synchronization: Effects of intention, attention, and awareness. *Quarterly Journal of Experimental Psychology, 57A*, 499-521.

Repp, B. H., & Keller, P. E. (2008). Sensorimotor synchronization with adaptively timed sequences. *Human Movement Science, 27*, 423-456.

Repp, B. H., & Keller, P. E. (2010). Self versus other in piano performance: Detectability of timing perturbations depends on personal playing style. *Experimental Brain Research, 202*, 101-110.

Repp, B. H., Keller, P. E., & Jacoby, N. (2012). Quantifying phase correction in sensorimotor synchronization: Empirical comparison of three paradigms. *Acta Psychologica, 139*, 281-290.

Repp, B. H., London, J., & Keller, P. E. (2005). Production and synchronization of uneven rhythms at fast tempi. *Music Perception, 23*, 61–78.

Resnicow, J. E., Salovey, P., & Repp, B. H. (2004). Is recognition of emotion in music performance an aspect of emotional intelligence? *Music Perception, 22,* 145-158.

Schmidt, R. C., & Richardson, M. J. (2008). Dynamics of interpersonal coordination. A. Fuchs & V. Jirsa (Eds.), *Coordination: Neural, behavioural and social dynamics*(pp. 281-308). Heidelberg: Springer-Verlag.

Schögler, B. (1999-2000). Studying temporal co-ordination in jazz duets. *Musicae Scientiae, Special Issue 1999-2000*, 75-91.

Schulze, H.-H., Cordes, A., & Vorberg, D. (2005). Keeping synchrony while tempo changes: Accelerando and ritardando. *Music Perception, 22*, 461-477.

Seashore, C. E. (1938). *Psychology of music*. New York, London: McGraw-Hill Book Company.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in Cognitive Sciences, 10*, 70-76.

Semjen, A., Schulze, H.-H., & Vorberg, D. (2000). Timing precision in continuation and synchronization tapping. *Psychological Research, 63*, 137-147.

Shaffer, L. H. (1981). Performances of Chopin, Bach, and Bartok: Studies in motor programming. *Cognitive Psychology, 13*, 326–376.

Shaffer, L. H. (1984). Timing in solo and duet piano performances. *Quarterly Journal of Experimental Psychology, 36A*, 577-595.

Smith, H., & Dean, R.T. (1997). *Improvisation, Hypermedia and the arts since 1945*. London: Harwood Academic.

Snyder, J. S., Hannon, E. E., Large, E. W., & Christiansen, M. H. (2006). Synchronization and continuation tapping to complex meter. *Music Perception, 24*, 135-146.

Sundberg, J. (2006).The KTH synthesis of singing. *Advances in Cognitive Psychology, 2*, 131-143.

Ternström, S. (2003).Choir acoustics: An overview of scientific research published to date. *International Journal of Research in Choral Singing, 1*, 3-11.

Ternström, S., & Karna, D. R. (2002). Choir. In R. Parncutt & G. McPherson (Eds.), *The science and psychology of music performance: Creative strategies for teaching and learning* (pp. 269-283). New York: Oxford University Press.

Todd N. P. M. (1985). A model of expressive timing in tonal music. *Music Perception, 3*, 33–58.

Todd, N. P. M., Cousins, R., & Lee, C. S. (2007). The contribution of anthropometric factors to individual differences in the perception of rhythm. *Empirical Musicology Review, 2*, 1-13.

Toiviainen, P., Luck, G., & Thompson, M. (2010). Embodied meter: hierarchical eigenmodes in music-induced movement. *Music Perception, 28*, 59-70.

Trusheim, W. H. (1993). Audiation and mental imagery: Implications for artistic performance. *TheQuarterly Journal of Music Teaching and Learning, 2*, 139-147.

Uhlig, M.,Schroeder, T., & Keller, P. E. (2012). Knowing too much or too little: The effects of familiarity with a co-performer's part on interpersonal coordination in piano duos. In *Proceedings of the 12th International Conference of Music Perception and Cognition* (ICMPC12). Thessaloniki, Greece.

Ullén, F., Forsman, L., Blom, Ö., Karabanov, A., & Madison, G. (2008). Intelligence and variability in a simple timing task share neural substrates in the prefrontal white matter. *Journal of Neuroscience, 28*, 4238-4243.

van Noorden, L., & Moelants, D. (1999). Resonance in the perception of musical pulse. *Journal of New Music Research, 28*, 43-66.

Varni, G., Volpe, G., & Camurri, A. (2010). A system for real-time multimodal analysis of nonverbal affective social interaction in user-centric media. *IEEE Transactions on Multimedia, 12*, 576 – 590.

Vesper, C., Butterfill, S., Knoblich, G., &Sebanz, N. (2010). A minimal architecture for joint

action. *Neural Networks, 23*, 998-1003.

Vorberg, D., & Wing, A. M. (1996). Modeling variability and dependence in timing. In H. Heuer & S. W. Keele (Eds.), *Handbook of perception and action: Motor skills*. San Diego, CA: Academic Press.

Williamon, A., & Davidson, J. (2002). Exploring co-performer communication. *Musicae Scientiae, 6*, 53-72.

Wöllner, C. & Cañal-Bruland, R. (2010). Keeping an eye on the violinist: motor experts show superior timing consistency in a visual perception task. *Psychological Research, 74*, 579-585.

Wöllner, C., Deconinck, F. J. A., Parkinson, J., Hove, M. J., & Keller, P. E. (in press). The perception of prototypical motion: Synchronization is enhanced with quantitatively morphed gestures of musical conductors. *Journal of Experimental Psychology: Human Perception & Performance*.

Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society B, 358, 593–602*.