Rhythm in joint action: psychological and neurophysiological mechanisms for real-time interpersonal coordination

Peter E. Keller1, Giacomo Novembre1 and Michael J. Hove2

1The MARCS Institute, University of Western Sydney, Locked Bag 1797, Penrith, New South Wales 2751, Australia
2Harvard Medical School, Boston, MA, USA

Human interaction often requires simultaneous precision and flexibility in the coordination of rhythmic behaviour between individuals engaged in joint activity, for example, playing a musical duet or dancing with a partner. This review article addresses the psychological processes and brain mechanisms that enable such rhythmic interpersonal coordination. First, an overview is given of research on the cognitive-motor processes that enable individuals to represent joint action goals and to anticipate, attend and adapt to other’s actions in real time. Second, the neurophysiological mechanisms that underpin rhythmic interpersonal coordination are sought in studies of sensorimotor and cognitive processes that play a role in the representation and integration of self- and other-related actions within and between individuals’ brains. Finally, relationships between social–psychological factors and rhythmic interpersonal coordination are considered from two perspectives, one concerning how social-cognitive tendencies (e.g. empathy) affect coordination, and the other concerning how coordination affects interpersonal affiliation, trust and prosocial behaviour. Our review highlights musical ensemble performance as an ecologically valid yet readily controlled domain for investigating rhythm in joint action.

1. Introduction

Human social interaction often involves the coordination of rhythmic behaviour between two or more individuals. Some forms of rhythmic interpersonal coordination emerge spontaneously, as when audience applause becomes synchronized, or when conversation partners unintentionally align their postural sway or inadvertently fall into step while walking [1–3]. Other forms of rhythmic interpersonal coordination are intentional and pre-planned [4–6]. Musicians in small ensembles coordinate their actions to produce sounds that form cohesive auditory tapestries; orchestral musicians synchronize with the gestures of a conductor, whereas jazz soloists groove to the beat of a rhythm section; dancers coordinate their body movements with one another and the sounds of a musical accompaniment.

The current article addresses the psychological processes and brain mechanisms that underpin the ability to coordinate intentionally with others in rhythmic joint activity. Such activity can be viewed as a specific class of joint action, that is, human behaviour that involves multiple individuals coordinating their thoughts and movements in space and time, with the goal to communicate [7] or to effect a change in the environment [8].

We consider joint actions to be rhythmic if their goals necessitate producing specific patterns of relative timing between co-acting individuals’ movements, and if these prescribed temporal relationships require precision in the order of tens of milliseconds. Regularly timed movements facilitate this degree of precision, but temporal regularity does not imply rigidity in the context of rhythmic joint action. Movement timing must be flexible enough to allow rate modulations—in the order of hundreds of milliseconds—to accommodate changing communicative goals, situational demands and degrees of mutual cooperativity between
interaction partners. Rhythmic interpersonal coordination thus requires simultaneous precision and flexibility of movements at different timescales.

In the following, we review the growing literature on rhythmic joint action. First, an overview is given of cognitive-motor processes that enable rhythmic interpersonal coordination. Then, the underlying neurophysiological mechanisms are described and, finally, we discuss the role of social–psychological factors (e.g. empathy and prosociality). Our focus is mainly on research concerning interpersonal coordination in musical contexts, specifically musical ensemble performance, which can be seen as a quintessential example of rhythmic joint action.

2. Cognitive-motor processes in rhythmic interpersonal coordination

The simultaneous need for temporal precision and flexibility in rhythmic joint action challenges the cognitive-motor systems of interaction partners. These demands are showcased in musical ensemble performance. Ensemble musicians coordinate their body movements to produce synchronous sounds and interlocking patterns in which separate instrumental parts articulate different but complementary rhythms. The tempo (musical speed) is typically modulated to communicate information about musical structure and expressive intentions [9]. During such tempo changes, which present moving synchronization targets, experienced ensemble performers maintain coordination, but this ability differs between individuals [10].

Keller [11,12] proposed a theoretical framework addressing the factors that influence rhythmic interpersonal coordination (figure 1). According to the framework, which was formulated in the context of ensemble performance, temporally precise rhythmic interpersonal coordination requires three core cognitive-motor skills: anticipation, attention and adaptation. These cognitive-motor skills are influenced by: (i) the performer’s goals concerning the interaction, (ii) his or her knowledge about the music and familiarity with co-performers, (iii) the use of regulatory strategies to facilitate coordination, and (iv) social–psychological factors (e.g. personality). While rhythmic interpersonal coordination in ensembles thus draws on mental resources that support joint action more generally [4,8,13], it does so in a manner that is tailored to handling the real-time demands and aesthetic goals of music performance.

We begin this section by addressing the real-time cognitive-motor skills that enable rhythmic interpersonal coordination, and then discuss the role of shared goals and knowledge. Our review focuses on research that has employed naturalistic musical tasks involving ensemble performance, as well as controlled laboratory tasks requiring individuals to coordinate simple movements (e.g. finger taps) with auditory and visual pacing sequences generated by a computer or another individual.

(a) Real-time coordination skills

Here, we review research on cognitive-motor skills that allow individuals to control the dynamics of real-time interpersonal coordination by anticipating, attending and adapting to each other’s actions.

(i) Mutual temporal adaptation

Rhythmic joint action is scaffolded by mechanisms that enable basic interpersonal entrainment. Generally speaking, entrainment entails the coupling of independently controlled rhythmic processes [6,14,15]. Entrainment in social contexts requires co-acting individuals to perceive rhythmic regularity in each other’s actions, to produce rhythmic movement and to integrate sensory information and produced movements via perception–action links [16]. Behavioural evidence for such links comes from studies showing that interpersonal entrainment arises spontaneously between individuals performing simple repetitive movements (e.g. in rocking chairs) [1,3].

Rhythmic joint action, however, often requires the coordination of complex movement sequences with irregular patterns of timing. Musical ensemble performance, for example, is characterized by intentional and unintentional variations in event micro-timing and tempo, as well as systematic deviations from strict synchrony between parts played by different individuals (e.g. leader–follower relationships where one part lags behind another) [11,17]. Such discrepancies in interpersonal timing must be kept in check through continuous mutual temporal adaptation.

Mutual adaptive timing is supported by temporal error–correction mechanisms that enable internal timekeepers—oscillations of neural populations in co-performers’ brains—to remain entrained despite irregularities in movement timing [18–23]. One mechanism, phase correction, is an automatic process that adjusts the alignment of pulses generated by an internal timekeeper in one individual relative to a sequence of pulses generated by a timekeeper in another individual. Phase correction supports precision in basic interpersonal coordination at the millisecond timescale. Another mechanism, period correction, involves consciously controlled adjustments to the duration of timekeeper intervals, and thus allows the flexibility that is required for co-performers to accommodate tempo changes at longer timescales.

The use of temporal error correction varies across individuals [24]. People engaged in rhythmic interpersonal coordination may therefore be predisposed to adapt to one another by differing amounts. These individual differences have been explored in sensorimotor synchronization experiments employing computer-controlled virtual partners instantiated as auditory pacing sequences that implement varying degrees of error correction, and hence cooperativity. This work has revealed that, while human phase correction remains constant across a range of cooperative virtual partners,
phase correction can be increased—and additional period correction may be used—when confronted with uncooperative virtual partners [21]. These countermeasures are most likely effortful [25] and may have costs in attentionally demanding forms of rhythmic joint action such as ensemble performance.

Related research has shown that strategic modulations of temporal adaptation that affect leader–follower relationships can be used to deal with differences in rhythmic skills between interaction partners. When faced with unstable partners who are prone to tempo drift, for example, the optimal strategy involves adopting the role of leader by assuming responsibility for tempo and adapting less to the partner’s irregular timing [26]. Entrainment is nevertheless a powerful phenomenon, making it difficult not to be influenced by others’ action timing. Indeed, entrainment can occur even when individuals actively avoid it, for example, in multi-religious rituals where independent groups of musicians exert their identity by exhibiting tight within-group coordination while attempting to avoid coordination between groups [27].

The foregoing suggests that mutual temporal adaptation is the glue that binds together individuals engaged in rhythmic joint action. In musical contexts, mutual temporal adaptation may, additionally, contribute to ensemble cohesion by enhancing the similarity of co-performers’ playing styles. Research with experimental tasks requiring piano duet performance [17,28] and dyadic finger tapping [29,30] has demonstrated that compensatory adjustments associated with error correction lead to co-dependencies, whereby successive time intervals produced by two interacting individuals tend to be similar in duration. Furthermore, a recent study employing an experimental task in which paired musicians tapped in alternation with an isochronous auditory pacing signal (analogous to two jazz soloists coordinating complementary parts with a rhythm section) found similarities between successive asynchronies produced by alternating individuals’ taps relative to the pacing tones [31]. This mutual temporal assimilation may be a form of non-conscious behavioural mimicry (see §4b) that facilitates ensemble cohesion by making multiple individuals sound collectively as one.

(ii) Attention
Rhythmic joint action is a form of multi-tasking. To produce a cohesive ensemble sound, ensemble musicians pay attention to their own actions (high priority) and those of others (lower priority) while concurrently monitoring the overall integrated ensemble output. This form of divided attention, which has been termed ‘prioritized integrative attending’ [32], involves a mixture of selective attention to self or other, and joint attention to the relationship between self and other (see §4b). Prioritized integrative attending therefore assists individuals to integrate their own actions with others’ actions while maintaining autonomous control of their own movements [21,33]. In music performance, it facilitates ensemble cohesion by allowing co-performers to adjust their actions based on the online comparison of the ideal ensemble sound and incoming perceptual information about the actual sound [11].

Prioritized integrative attending is cognitively demanding to the extent that it involves the simultaneous segregation and integration of information from separate sources [32,34–36]. Research on multi-part musical rhythm perception and production suggests that the flexibility required for such segregation and integration is enabled by metric frameworks. These frameworks guide the dynamic allocation of attentional resources in accordance with cognitive-motor schemas that comprise hierarchically arranged levels of pulsation (beat subdivisions, beats and groupings of beats into bars) [37–39]. The entrainment of internal timekeepers to multiple periodicities associated with metric structure thus provides a hierarchical temporal scheme for modulating attention in a manner that is conducive to monitoring different levels of the musical texture simultaneously [32,40].

(iii) Anticipatory mechanisms
Anticipatory mechanisms facilitate precise rhythmic interpersonal coordination by allowing individuals to plan the timing of their own actions with reference to predictions about the future time course of others’ actions. Ensemble performers use anticipatory cognitive-motor mechanisms to plan the production of their own sounds and to generate online predictions about the upcoming sounds of co-performers [11]. It has been proposed that these predictions evolve along two routes [6]. On one route, automatic expectancies about events at short timescales (e.g. the next tone or endpoint of a conductor’s beat gesture) are triggered directly by the perception of sounds and body movements. The other route involves anticipating co-performers’ actions by activating memory representations of shared goals [12] (§2b). Activating these internal representations entails running action simulations that are experienced as auditory and motor imagery of sound sequences and related movements [41].

Temporal prediction abilities are thus constrained by the fidelity of action simulations and mental images. This hypothesis was supported in an experiment which found that pianists were able to synchronize better with recordings of their own performances than with others’ recordings, presumably because the match between simulated and actual timing was best when both were products of the same cognitive-motor system [42]. A subsequent study on the role of mental imagery revealed that individual differences in the coordination of sounds and body movements in piano duos were positively correlated with performance on a task assessing the vividness of anticipatory auditory imagery [10]. Moreover, studies targeting temporal prediction directly have shown that the precision of sensorimotor synchronization in real and virtual interpersonal coordination is positively correlated with temporal prediction abilities [43,44], which are, in turn, correlated with auditory imagery abilities [45].

It has been claimed that action simulation is driven by internal models instantiated in the central nervous system [41,46–48] (§3b). Two varieties of internal model represent associations between efferent motor commands and afferent sensory information [49,50]. 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 to the motor commands that produce them. It has been argued that rhythmic interpersonal coordination relies on separate forward and inverse models to simulate one’s own and others’ actions slightly in advance of their production [11,41,51,52]. On this account, the coupling of ‘own’ and ‘other’ internal models in a ‘joint’ internal model facilitates fluent interpersonal coordination by allowing potential errors in timing to be anticipated and corrected before they occur [41].

The sensorimotor transformations represented in internal models are acquired and strengthened through active
experience and observational learning [53]. Reliable and efficient internal models are therefore a hallmark of expertise. Accordingly, it has been claimed that experienced ensemble musicians develop hierarchically arranged internal models that are able to represent events at multiple timescales [47,54] and to simulate a co-performer’s playing style by calibrating to his or her action system [55]. Moreover, anticipatory mechanisms and adaptive timing may be linked in hierarchically arranged ‘joint’ internal models of one’s own and others’ actions that allow interpersonal timing errors to be simulated and corrected in advance [56]. Attention may regulate this process by modulating the gain of error correction and predictive simulation processes.

(b) Shared goals and knowledge

Complex forms of rhythmic joint action—such as those encountered in musical ensemble performance—require pre-planning. Ensemble musicians usually prepare for performance through collaborative group rehearsal geared towards establishing shared performance goals, that is, unified conceptions of the ideal ensemble sound [11,57]. Co-performers thus form memories of each other’s parts and the relationship between these parts [12,28]. While this process may be grounded in the automatic tendency for individuals engaged in joint action to represent each other’s tasks [4,8], the amount of time that ensemble musicians invest in rehearsal suggests that developing shared performance goals is effortless. Other forms of skilled joint action similarly require collaborative training, such as dance, or when a football team practices a set play, a rowing team practices changes in pace, or a flight crew practices operational sequences in a simulator.

Research on ensemble performance suggests that developing shared goals involves acquiring knowledge about the musical structure and the expressive intentions and playing styles of ensemble members [58,59]. Musical structure refers to the hierarchical patterning of pitch and rhythmic elements [60]: individual tones are concatenated into melodic motives and phrases, while rhythmic durations can be defined relative to the temporal units of an underlying metric framework. The way musical structure is rendered in performance is flavoured by micro-timing deviations and aesthetically motivated tempo variations that reflect an individual musician’s expressive intentions and idiosyncratic playing style [9].

The importance of knowledge about both musical structure and playing style is highlighted in a recent study of piano duos [54]. Pairs of unacquainted pianists came to the laboratory after rehearsing their co-performer’s part. These results suggest that knowledge of a co-performer’s part, in the absence of knowledge about their playing style, engenders predictions about expressive micro-timing variations that are based instead upon one’s own personal playing style, leading to a mismatch between predictions and actual events at short timescales. As knowledge about a co-performer’s stylistic idiosyncrasies is acquired, however, the individual learns—through the calibration of internal models—to simulate the other’s action style. By contrast, body sway coordination was high throughout repeats in the familiar condition, while it started out low and improved across repeats in the unfamiliar condition. This suggests that knowledge about the structure of a co-performer’s part facilitates predictions at longer timescales related to high-level metric units and musical phrases and reflected in ancillary body sway movements. Knowledge of musical structure and personal playing style may thus function together to promote ensemble cohesion at multiple timescales.

Shared goals ensure that ensemble musicians take each other’s actions into account during performance. Research on joint action outside the music domain has shown that individuals behave differently when performing a task alone or with a co-actor [61]. In tasks that require interpersonal coordination, individuals increase the salience and regularity of their movements, as a strategy to improve coordination [13]. Musicians likewise employ regulatory strategies to facilitate ensemble cohesion. For instance, one study found that expressive devices, such as tempo accelerations and decelerations, are dampened during ensemble performance relative to when a musician performs their part alone [62]. Furthermore, an ensemble leader may sharpen the contrast between rhythmic durations in order to communicate clear expressive intentions [63], as well as exaggerate sound-producing movements while simplifying ancillary movements such as head gestures [17,64].

3. Neurophysiological mechanisms

Experimental research on the neural mechanisms that underpin rhythmic joint action has addressed two kinds of questions. The first deals with neural networks: what brain areas are implicated in interpersonal coordination? The second deals with information processing: to what extent can socially coordinated behaviour be understood in terms of neural oscillations within and/or between multiple brains? These two questions are complementary in the sense that a network constitutes the biological means for processing specific information. Nevertheless, research on each topic has made unique contributions to understanding different aspects of the neurophysiological mechanisms underpinning rhythmic joint action.

The network perspective builds on the notion of shared neural resources for action production and action perception. The coupling of perception and action has been investigated in human and non-human neurophysiology [65], computational neuroscience [66] and cognitive psychology [67,68]. Perception—action coupling serves social interaction by supporting the efficient integration of actions produced by self and others [50]. In line with this, it has been shown that others’ actions evoke stronger motor responses in the observer if he or she intends to interact with, rather than imitate, a partner [69]. Furthermore, communicative social behaviour, such as hand gestures [70] and facial expressions [71], is reflected in the directional transfer of sensorimotor brain states from one individual to another.

The oscillatory perspective, on the other hand, has provided evidence that socially coordinated actions can be identified by electrophysiological neuromarkers. Studies using
electroencephalography (EEG) have revealed that oscillatory brain activity in the α-band (8–12 Hz) is associated with effective social coordination [72–74]. Socially interactive tasks requiring temporal coordination are also associated with large pools of neurons oscillating coherently across co-acting individuals’ brains. This phenomenon has been observed in dual-EEG studies employing diverse tasks requiring joint rhythmic behaviour, such as guitar duet performance [75–77], unconscious imitation of finger movements [78], explicit imitation of hand movements [79] and joint speech [80]. Dual-EEG set-ups provide a promising avenue for exploring the temporal dynamics of mechanisms supporting rhythmic joint action, though the complexity of such set-ups means that caution must be exercised to ensure that studies are conceptually and methodologically sound [81,82].

Taken together, this body of research suggests that a network composed of tightly linked sensory and motor brain areas is responsible for the efficient integration of self and other-related behaviour. While self–other integration occurs within a single individual’s brain, it may potentially lead to coupling across individuals by inducing compatible modulations of activity in sensorimotor regions [83]. Such co-modulation may be a prerequisite for simple forms of unintentional interpersonal coordination as well as complex varieties of intentional coordination that require joint planning.

In the next sections, we describe studies that used the musicians’ brain as a means to examine the role of sensorimotor networks in rhythmic joint action. These studies are informative about the neurophysiological underpinnings of the cognitive-motor skills that enable real-time interpersonal coordination (anticipation, adaptation and attention; §2a) and the representation of self and others in the context of rhythmic joint action.

(a) Representing self and other in the brain

Intentional forms of temporal coordination with another individual require the brain: (i) to simultaneously represent self-generated and other-related actions, and (ii) to integrate them in real time. These two requirements are associated with distinct processes that both rely upon the brain’s capacity to code action production (related to the self) and perception (mostly related to others) in a comparable way. Experience with a given action appears necessary for this common coding [68,84], which occurs through changes in brain connectivity owing to neuroplasticity [85]. Common coding ensures that, when skilled pianists hear sounds [86,87] or see movements [88] associated with a trained piece, they represent these effects in terms of the neural resources necessary to produce them. These representations carry information about the action’s temporal dynamics [89], which permits the observer to anticipate the action internally [90]. On this view, action simulation (§2a) involves accessing this temporal information by activating an action representation.

Although the networks of brain areas responding to generated and observed actions are partially shared [65], the neurophysiological mechanisms underpinning these responses are assumed to be distinct. A certain level of distinction, or segregation, between self and other is presumably important for coordinating with others while maintaining a certain degree of autonomy [5,33]. The self–other distinction has recently been explored in the context of joint rhythmic behaviour, including piano duet performance.

In a transcranial magnetic stimulation (TMS) experiment [91], pianists were asked to practice several piano pieces bimanually before coming to the laboratory. Then, in the laboratory, the pianists were required to perform only the right-hand part of each piece, while the complementary left-hand part was either not performed or produced by another pianist hidden behind a screen (this hidden pianist feigned playing while the participant actually heard a recording). It was hypothesized that bimanual learning of the piece would lead to a co-representation of the left-hand part, which would then be associated either with the self (when it was not performed) or with the other player (behind the screen). To test this hypothesis, TMS pulses were occasionally delivered over the right primary motor cortex to elicit motor evoked potentials (MEPs), which were recorded from a forearm muscle that would normally be used to play the left-hand part (cf. [92]). Differences in MEP amplitude suggested that distinct patterns of cortico-spinal excitability—inhibition and excitation—were associated with the representation of self and other, respectively (cf. [93–96]). Interestingly, the same result was observed when the pianists were told that the co-performer was playing behind the screen, but auditory feedback from the hidden co-performer was not provided. This suggests that these motor representations arise in response to the potential for interaction with another, and may be intrinsically social in nature.

In a parallel vein, an EEG study [97] explored the representation of self and other during rhythmic joint action using a similar piano performance task to the one described above [91]. Pairs of pianists practiced each other’s parts prior to the experiment, and then played the left- and right-hand part, respectively, while EEG was recorded. The experimenters altered the auditory feedback provided to the pianists to create occasional mismatches between their keystrokes and the produced tones. The mismatch either did or did not affect the harmony between the players’ parts, hence permitting the differentiation of processes related to monitoring one’s own performance and the joint action outcome. Analysis of evoked response potentials showed that altered auditory outcomes elicited similar early responses (a feedback-related negativity peaking around 250 ms after tone onset) regardless of whether they were associated with the self or the other. By contrast, a later response, i.e. a P300, was larger for self-related (altered) feedback, as well as for altered feedback that affected the joint outcome. These data provide further evidence for distinct representations of self and other, as well as their integration, in the context of rhythmic joint action.

(b) Neural bases of real-time coordination skills

The real-time coordination skills that enable rhythmic joint action are supported by multiple neurophysiological mechanisms. In accordance with the theoretical framework described in §2 (figure 1), we now review cognitive neuroscience research that has explored the mechanisms that underpin temporal anticipation, prioritized integrative attention and adaptive timing.

Advantages associated with the capacity for temporal anticipation during rhythmic joint action were a probable driving force in the evolution of the human brain. Although the brain has long been viewed as an anticipatory machine [98,99], the notion that the online prediction of others'
behaviour is achieved by internal models is relatively recent. On this account, internal models in the cerebellum [49] cohere with cortical areas, including those subserving motor simulation [100,101]. It is owing to this recruitment of the observer’s motor system that internal models are calibrated to the observer’s own action style [28,42,54].

A recent repetitive TMS study [102] examined the role of motor cortical regions in temporal prediction by using a musical task in which pianists were asked to adapt the performance of the right-hand part of a piano piece to a tempo change in the left-hand part (which was heard but not performed). It was assumed that temporal predictions generated by the motor simulation process would facilitate rapid and accurate tempo adaptation. In order to manipulate the degree of motor simulation, the experimental design ensured that pianists either had or had not practiced the left-hand part prior to the tempo adaptation task [91,92]. Practicing the left-hand part was intended to assist the development of internal models that represented the sensorimotor transformations involved in performing it. Time-locked repetitive TMS was delivered over the primary motor cortex to interfere with the motor simulation processes prior to the tempo change, and then the accuracy of tempo adaptation was measured. It was found that TMS (compared to sham stimulation) impaired temporal adaptation accuracy only when the part had been practiced (and, therefore, could be simulated). This indicates that motor simulation provides a functional resource for the real-time temporal coordination of self- and other-generated actions.

Brain imaging studies employing perceptual- and production-based tasks with rhythmic auditory stimuli support the view that motor (simulation) processes are involved in temporal prediction [53,103–105]. A recent functional magnetic resonance imaging (fMRI) study investigated these processes using a task that required musicians to tap along with auditory tone sequences containing gradual tempo changes [106]. As noted in §2a, high scores on behavioural measures of prediction in this task are associated with relatively good performance at real rhythmic interpersonal coordination [43]. In order to manipulate prediction abilities in the scanner, the synchronization task was performed under conditions of variable cognitive load: in one condition, participants tapped while merely watching a stream of novel objects; in a more difficult condition, they counted objects repeated after an intervening item. A parametric analysis of brain regions in which activity decreased with decreases in temporal prediction across these conditions revealed an extensive network that included cortico-cerebellar motor-related brain areas (precentral gyrus, supplementary motor area and cerebellum), the auditory cortex and adjacent temporal areas, as well as medial prefrontal cortical areas implicated in error monitoring and social-cognitive processes. Decreases in the activation of this network were accompanied by increased activity in cerebellar subregions involved in temporal tracking and error correction (rather than prediction), as well as in a fronto-parietal network, which may reflect the working memory and attentional demands of the secondary task.

The attentional demands of rhythmic joint action entail the concurrent monitoring of one’s own actions, others’ actions and the relationship between them. In musical contexts, this form of prioritized integrative attending (§2a) necessitates the simultaneous segregation and integration of auditory streams produced by different individuals. Studies of auditory attention in general have revealed the involvement of a broad network of fronto-parietal and temporal brain regions [107,108]. To ascertain which parts of this network are implicated in prioritized integrative attending to multi-part musical textures, Uhlig et al. [36] conducted an fMRI study in which expert pianists were presented with recordings of piano duets composed of a melody and an accompaniment. The parts were shifted in terms of their relative onset times (i.e. the melody led or lagged behind the accompaniment by 28 ms, on average) and pianists were asked to assess the leader–follower relationship between them. Results suggested that a fronto-parietal brain network comprising the dorsolateral prefrontal cortex, the inferior parietal lobe and the intraparietal sulcus is involved in regulating the balance between segregating a high-priority part while integrating information across parts. This work was extended [109] by comparing prioritized integrative attention in the context of a human performance of a piano duet (which contained temporal asynchronies between parts) and a synthetic rendition of the duet without asynchronies. Results suggested that the planum temporale (located on the superior temporal gyrus) is implicated in segregation based on asynchronies between parts (which can contribute to the perception of separate auditory streams), while the intraparietal sulcus is involved in the integration of parts.

Like the anticipatory and attentional processes that enable rhythmic joint action, adaptive timing is supported by distributed networks of cortical and subcortical brain regions. Brain imaging, stimulation and patient studies of sensorimotor synchronization with auditory pacing sequences converge on the idea that phase correction is implemented via a network including subregions of the cerebellum that are linked to motor and auditory cortical areas, while period correction recruits an additional corticothalamic network that includes the basal ganglia, prefrontal, medial frontal and parietal regions [110]. The network serving phase correction is specialized for the pre-attentive processing of microtiming information, whereas the period correction network handles attention-dependent processing at longer timescales associated with musical beat and metre [111].

Adaptive timing involves the modulation of neural oscillations in these networks. The dynamics of neural oscillations linked to beat and metre perception in acoustic signals has been studied using EEG and magnetoencephalography (MEG) [105,112–114]. This research has revealed beat-related modulations in oscillatory activity in relatively high MEG and EEG frequency bands (20–30 Hz β and 30–60 Hz γ), as well as evidence for hierarchical patterns of neural entrainment at the beat and longer periodicities in EEG steady-state evoked potentials. Individual differences in these patterns of neural entrainment are a potential source of differencing adaptive timing skills.

The neural correlates of dynamic cooperativity in adaptive timing have been investigated in fMRI studies in which musicians are required to synchronize with virtual partners that implement differing degrees of temporal error correction [25,26]. In one study [25], virtual partners were programmed to implement different amounts of phase correction, and hence to exhibit different degrees of cooperativity. Overly cooperative virtual partners who engaged in high amounts of phase correction (resulting in over-compensation for synchronization errors when combined with the participant’s
own phase correction) led to poor inter-agent synchronization and the activation of lateral prefrontal areas associated with executive functions and cognitive control. Optimally cooperative virtual partners who engaged in moderate amounts of phase correction (matching typical estimates of human phase correction) facilitated precise synchronization and led to the activation of cortical midline structures associated with socio-affective processes. This finding points to the neurophysiological mechanisms that may underlie the process by which synchronous group behaviour promotes social cohesion (§4b).

4. Social—psychological factors

In this section, we outline how social factors affect rhythmic interpersonal coordination, and in turn how interpersonal coordination affects social judgements and behaviour. The bidirectional links between interpersonal coordination and social factors are probably involved in the enjoyment of synchronizing with others and the evolution of music.

(a) Social-cognitive influences upon interpersonal coordination

Social—psychological factors affect rhythmic interpersonal coordination at multiple levels. Studies with musical ensembles have revealed that communication effectiveness during rehearsal is influenced by personality, pre-existing interpersonal relationships, and verbal and non-verbal communication styles [58,115]. Experimental work addressing interpersonal coordination has identified links between personality characteristics and the cognitive-motor skills involved in rhythmic interpersonal coordination.

In general, timing of interpersonal coordination is affected by social skills. For example, children with higher social skills, as assessed by their teachers, synchronized better with others in a dyadic drumming task [116]. This could stem from increased awareness of others in a social context. In another drumming task, children as young as 2.5 years old synchronized better with an adult’s movement, as compared with a mechanical drumming machine or an auditory metronome. This increased coordination during social interaction could stem from creating a shared representation of the joint action [117]. Representing the others’ action should improve the ability to anticipate and adapt to the timing of their movements.

Social orientation, or people’s general disposition to social situations, affects interpersonal coordination. A study of unintentional coordination revealed that prosocial-oriented individuals spontaneously synchronized arm movements with others more than pro-self-oriented individuals, whether their social/self-orientation reflected their pre-existing disposition or resulted from an experimental manipulation [118]. Social orientation and motivation are malleable and affect coordination. Relevant studies have found that interacting with a late-arriving partner reduced stepping synchronization, compared with interacting with a partner who arrived on time [119], and bodily synchrony decreased during arguments compared with affiliative conversations [120].

Aspects of personality such as social competence, empathy and locus of control have also been shown to relate to the specific cognitive-motor mechanisms involved in interpersonal coordination, such as temporal anticipation and adaptation. The concept of empathy—understanding others’ thoughts and feelings—has been linked to anticipatory mechanisms related to action simulation [121]. In the TMS studies of piano duos described in §3a,b, scores on the ‘perspective-taking’ subscale of an empathy questionnaire correlated positively with neurophysiological measures of representing the other’s part in their own motor system [91], as well as how much this ‘other-representation’ was relied upon for coordination [102]. The latter finding suggests that empathic predispositions may affect temporal predictions at the millisecond timescale. In support of this, scores on a perspective-taking questionnaire correlated with the degree that individuals predicted event micro-timing in a tempo-changing pacing sequence in a synchronized finger-tapping task [43].

A recent study addressed the relationship between locus of control (i.e. the degree to which life events are perceived to result from one’s own actions) and temporal adaptation (error correction) [26]. Individuals tapped along with sounds produced by virtual partners who varied in tempo stability and were prone to drift when endowed with high levels of period correction. Results indicated that individuals with an internal locus of control (who attribute the cause of events to their own actions) engaged in less phase correction than individuals with an external locus of control (who attribute events to external factors). This may reflect a difference in leader–follower strategy: ‘internal’ individuals stabilized the tempo of their own performance (at the expense of synchrony) and took a leader role, whereas ‘external’ individuals synchronized with their partner (at the expense of maintaining a steady tempo) and took a follower role. A complementary leader–follower relationship can improve rhythmic social coordination [122]. In a pendulum swinging experiment, pairs with dissimilar social competences (high and low competence) synchronized with each other more stably than pairs with similar social competence [123].

Social factors extend beyond the personality and social competence of the individual. Interpersonal coordination depends on the interaction partners’ social group membership. Miles et al. [124] manipulated participants’ group membership and showed that individuals synchronized most stably with someone from a different social group. The authors suggest that interpersonal synchrony could be a means to reduce group differences and lessen social distance. This strategy makes sense in light of the positive social outcomes of rhythmic interpersonal coordination reviewed next.

(b) Social consequences of interpersonal coordination

Interpersonal coordination impacts social variables and leads to greater social cohesion, cooperation and trust. Activities involving rhythmic interpersonal coordination, such as marching, military training, religious ritual, chanting, dance and music can create the feeling of expanding into the larger group, and have long been used to increase social bonding [125].

Over the past decades, theoretical and empirical links have been established between coordinated movement and social rapport (e.g. [126]). Early support came primarily from studies on non-rhythmic mimicry (e.g. [127,128]). However, rhythmic synchrony and precise timing coordination are considered critical. The synchrony–rapport connection was supported in an early study of mock student–teacher interactions: participants’ ratings of rapport correlated with outside observers’ ratings of their movement synchrony [129]. In the past few
years, experimental work has grown substantially on the social consequences of rhythmic coordination. Movement synchrony has been quantified and manipulated and has been shown to affect numerous social outcome measures.

Interpersonal synchrony leads to increased social cohesion. In a study by Hove & Risen [130], participants tapped their fingers with a visual pacing sequence, while seated next to an experimenter who tapped with the same or different pacing sequence (or did not tap in a control condition). After tapping in synchrony, participants rated the experimenter more likeable; and the degree of timing precision predicted subsequent likeability ratings. Ratings in the quiescent control condition were similar to the asynchrony condition, demonstrating that synchrony boosted affiliation, rather than asynchrony decreasing it [130].

In a series of studies, Marsh et al. [122] showed the importance of synchrony for promoting social cohesion. Interactions with more synchronous movement (in a pendulum swinging task) were rated as friendlier and more harmonious. In a rocking chair study, pairs of participants, who directed their attention to their partner, coordinated more and perceived a greater sense of ‘team-ness’ than pairs who stared straight ahead. These studies provide support for synchrony as the basis of sociality [122]. From this dynamic perspective, interpersonal coordination and its effects emerge spontaneously in social interaction (for reviews, see [122,131,132]), providing a foundation upon which complex and rewarding patterns of intentional coordination can be built in musical contexts.

In addition to social judgement, rhythmic interpersonal coordination boosts social behaviour, such as cooperation. After synchronizing movements in walking or musical interactions, participants cooperated more in group-economic exercises, even when requiring personal sacrifice [133,134]. This increased cooperation was thought to stem from synchrony’s ability to increase social attachment. The effects of interpersonal synchrony are powerful: prosocial effects extend both to partners and to non-participants [135], and can even spawn extreme cooperation in the case of destructive obedience [136]. But in general, synchrony has positive social effects such as cooperation and trust [137], and can improve empathy and prosocial behaviour in children [138,139].

The cooperative effects of interpersonal coordination appear to develop early [140]. Fourteen-month-old infants, who were bounced in synchrony with an experimenter, were more likely to help the experimenter pick up the ‘accidentally’ dropped pencils [141]. Four-year-old children showed increased cooperation after synchronizing in a joint-musical game; this cooperation was thought to stem from increased attention to shared goals [142].

The social consequences of interpersonal coordination stem from multiple factors. Interpersonal coordination can direct attention to the partner and create an ‘attentional union’ that augments perception of the other [143]. Moving in synchrony with a partner improved memory of what the partner said [143], whereas moving out-of-synchrony improved memory of self-produced words [144]. Thus, rhythmic coordination can direct attention and enhance the representation of the other. In turn, this could increase perception of similarity. Individuals who move in synchrony are judged to form an integrated social unit [145–147]. After a synchronous interaction, participants viewed their partner as more similar to themselves (and were more compassionate), thus suggesting that synchrony may lead individuals to perceive themselves as united [148].

When we move together, we attend to each other more and interpret the coordination as a marker of our similarity and shared goals.

The social consequences of interpersonal coordination can also stem from factors related to temporal adaptation and anticipation. After rocking in synchrony, pairs felt more connected, and in a subsequent joint action task they were better able to adapt to their partner, suggesting that synchrony hones the cognitive-motor skills needed for coordination [149]. Synchrony’s ability to tune adaptation and anticipation could enhance joint action success and promote social cohesion. Interactive coordination relies on representing or simulating the others’ action. Representations of self and other are highly aligned during synchronous coordination (§3a). This could decrease the distinction between self-produced and other produced action and drive affiliation [150], cf. [151,152]). Coordinated and contingent (but non-synchronous) movements, as in many musical contexts, also promote affiliation ([153], cf. [154]), and could stem from co-representing the others’ action. Neural alignment during coordinated behaviour might allow access to others’ states and a ‘sameness recognition’ [122,135], and thus provide an important grounding for social cognition and a scaffold for communication [155,156]. In conversation, movement synchrony may serve as a low-level aid to realign speakers when higher level communication breaks down [157]. At the level of the brain, rhythmic coordination can be used to promote neural alignment, which could improve joint action, communication, neural efficiency and thus social bonding [158].

Synchrony–prosocial links are undoubtedly supported by neural and neurochemical mechanisms, but little work has directly examined these mediators. An fMRI study suggested that the influence of synchrony on prosociality relates to caudate activation in the brain’s reward system [51]. Reward signals are mediated by release of dopamine and opioids [159], thus neurochemistry could influence the synchrony–affiliation links [160]. Opioids and endorphins have been implicated in social bonding, and endorphin levels (as measured by pain tolerance) appear higher after synchronous rowing (but not after antiphase joint rowing) [161,162]. A study of tango dancing varied music and social factors and found that moving with music decreased cortisol levels, whereas moving with a partner increased testosterone levels [163]. Oxytocin is another candidate involved in social bonding and affiliation and has been shown to increase after musical interactions [164]. Future work that varies social, musical and synchrony factors should be fruitful in understanding the role of neurochemistry in social coordination and affiliation.

In summary, many, perhaps complementary, features are involved in the prosocial consequences of rhythmic interpersonal coordination. Moving together in time has powerful social consequences and might have played a role in the evolution and enjoyment of music (e.g. [30,142,160,165,166]). Ancient people with a heritable tendency to enjoy synchrony would synchronize more, have closer social ties and leave more surviving offspring [167], thus offering an evolutionarily plausible account for the human pleasure of moving together in time.

5. Conclusion

Rhythmic joint action requires simultaneous temporal precision and flexibility in interpersonal coordination at multiple
timescales across different sensory modalities. Such coordination is supported by cognitive-motor skills that enable individuals to represent joint action goals and to anticipate, attend and adapt to other’s actions in real time. The neurophysiological mechanisms that underpin these cognitive-motor skills are shaped by experience to enable precise yet flexible interpersonal entrainment, as well as the representation and integration of information about self and other within and between individuals’ brains.

Individual differences in rhythmic interpersonal coordination can be accounted for by the interaction of an individual’s cognitive-motor skills with their knowledge and goals concerning the task, familiarity with co-actors, use of regulatory strategies and social-cognitive aspects of personality (e.g. empathy and locus of control). Furthermore, interpersonal coordination can have reciprocal effects upon social outcomes concerning interpersonal affiliation, trust and prosocial behaviour.

A leitmotif in our review is the notion that human interaction in musical contexts, such as ensemble performance, provides an ecologically valid yet readily controlled domain for investigating the psychological processes and neurophysiological mechanisms that underlie rhythmic joint action. Moreover, to the extent that musical group behaviour is a microcosm of human social interaction, this ancient form of communication may provide a portal for exploring the roots of human prosociality.

Acknowledgements. Many of the ideas described in this article originated from collaborative work conducted in the Max Planck Research Group on ‘Music Cognition and Action’ at the Max Planck Institute for Human Cognitive and Brain Sciences in Leipzig, Germany. The authors are grateful to all members of the group, and regular visitors (especially Bruno Repp), for years of stimulating discussions.

Funding statement. P.E.K. and G.N. were supported by the MARCS Institute while writing this article and M.J.H. was supported by NIH grant no. T32 MH16259-31.

References


82. Konvalinka I, Roepstorff A. 2012 The two-brain network while listening to newly acquired actions.
70. Fukai T, Hishida T, Tononi G. 2006 The dynamic contrast between excitatory and inhibitory plasticity.
68. Fujisawa S, Miller EK. 2006 The dynamic contrast between excitatory and inhibitory plasticity.