Joint action: From perception-action links to shared representations

Janeen D. Loehr[1], Natalie Sebanz[1,2], and Günther Knoblich[1,2]

[1]Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour

[2]Central European University, Budapest, Hungary

In W. Prinz, M. Beisert, and A. Herwig (Eds.),

Tutorials In Action Science, MIT Press.

Address correspondence to:

Janeen Loehr
Department of Psychology
University of Saskatchewan
9 Campus Drive
Saskatoon, Saskatchewan, Canada
S7N 5A5
janeen.loehr@usask.ca

Imagine a pianist and a violinist performing a duet together. Each performer must plan and produce a long sequence of precisely timed actions, and each must monitor the perceptual consequences of those actions to ensure that the correct sequence of sounds is produced. Successful performance requires that the pianist and violinist perform these tasks jointly rather than individually: Each performer must time his or her actions to coincide with the other's, each must modify his or her action plans depending on the perception and prediction of the other's actions, and each must monitor not only the perceptual consequences of his or her own actions but also the combined outcome of the two performers' actions. Thus, the two performers' actions and perceptions are intimately linked (see Kiefer & Barsalou, this volume); one performer's actions and perceptions can hardly be understood without the other's.

It is not only in ensemble music performance that one person's actions and perceptions are linked with another's. Joint actions, in which two or more people coordinate their actions in space and time to bring about a change in the environment (Sebanz, Bekkering, & Knoblich, 2006), abound in daily life. Examples range from exchanging money with a cashier or cooking dinner with a friend to dancing, playing sports, and having conversations (Clark, 1996). In this chapter, we aim to elucidate the basic mechanisms of perception and action that underlie such joint actions, using music performance as a particularly rich example. We focus on two types of mechanisms. We begin with close links between perception and action, which support joint action by evoking similar actions or similar timing of actions in multiple individuals (Marsh, Richardson, & Schmidt, 2009). We then turn to shared representations of co-performers' tasks and perceptions. These representations support joint action by specifying how each individual's actions are planned, monitored, and controlled (Vesper, Butterfill, Knoblich, & Sebanz, 2010).

Close Links Between Perception and Action

Close links between perception and action lead to interpersonal coordination because two or more individuals process similar perceptual or motor information, which induces them to act in similar ways. We distinguish two related phenomena evoked by perception-action links: entrainment and perception-action matching. Entrainment leads people to align their actions in time (Schmidt & Richardson, 2008), even when their actions differ. For example, the pianist and violinist described above produce different actions, which are nevertheless nearly simultaneous. In contrast, perception-action matching induces people to perform similar actions (Brass & Heyes, 2005) or actions with similar perceptual consequences (Hommel, Müsseler, Aschersleben, & Prinz, 2001; Prinz, 1997), without necessarily being aligned in time. For example, people engaged in conversation tend to mimic each other's actions at variable temporal delays (Chartrand & Bargh, 1999). Both entrainment and perception-action matching can result in coordinated behavior regardless of whether or not the individuals involved plan to perform a joint action, as when audience members clap in unison during spontaneous applause (Néda, Ravasz, Brechet, Vicsek, & Barabási, 2000), or when people yawn upon seeing someone else do so (Provine, 1986). Coordination that results from these processes is therefore referred to as emergent coordination (Knoblich, Butterfill, & Sebanz, 2011), though these processes also support action alignment in planned coordination, as we will describe next.

Entrainment: Alignment in Time

One striking aspect of the duet music performance described above is the close temporal alignment between performers. Ensemble musicians typically produce nominally simultaneous tones with asynchronies of around 30 to 50 ms (Rasch, 1979), and asynchronies are even smaller for simple pieces performed on the same instrument (Loehr & Palmer, in press). Entrainment is one mechanism by which this precise temporal alignment may be achieved. Proponents of entrainment as an explanation for interpersonal coordination argue that the same mathematical principles that underlie coordination between physical systems (specifically, the nonlinear dynamics of coupled oscillation) also underlie coordination between people (see Pikovsky, Rosenblum, & Kurths, 2001, for an accessible introduction to the mathematics of coupled oscillation). By this account, cognitive representations of perceptions and actions are not necessary. Just as two pendulums suspended from the same beam will come to swing at the same rate because they can influence each other directly through the mechanical connection between them, so will two people's actions become aligned in time when they can influence each other directly through shared visual, auditory, or haptic information (Kelso, 2001; Schmidt, Carello, & Turvey, 1990; for a review, see Schmidt & Richardson, 2008).

Evidence of entrainment in joint action. The mathematical principles underlying coupled oscillations predict patterns of interpersonal coordination that arise when partners coordinate rhythmic movements based on visual perception of each other's movements. For example, people show two stable modes of coordination when they are asked to swing pendulums back and forth together: in-phase coordination, in which both pendulums are in the same position in the movement cycle (e.g., both swung forward) at the same time, and anti-phase coordination, in which the pendulums are in opposite positions at the same time (when one pendulum is swung forward, the other is swung backward; Schmidt et al., 1990). Anti-phase coordination is less stable than in-phase coordination, and because of this reduced stability, a sudden transition to in-phase coordination occurs as the rate at which participants swing their pendulums increases (Schmidt, Bienvenu, Fitzpatrick, & Amazeen, 1998). Each of these patterns is predicted based on the Haken-Kelso-Bunz model of coupled oscillations (Haken, Kelso, & Bunz, 1985); thus, when two people coordinate their rhythmic movements based on visual perception, the temporal alignment they produce reflects the dynamics of coupled oscillations.

The dynamics of coupled oscillation are also evident in coordination based on auditory perception. When people listen to music alone, their internal rhythms (e.g., sense of the musical beat, as sometimes manifest in tapping along with music) become entrained to the music (Large, 2008; Large & Palmer, 2002). The mathematical principles underlying coupled oscillation can explain, for example, how people track rate changes in expressive music performance, in which performers slow down or speed up to convey a particular musical interpretation (Large & Palmer, 2002). Recent work has shown that these same principles can explain how people adapt to rate changes that may arise when they play music together: Musicians are better able to coordinate their performances with auditory sequences that slow down than sequences that speed up, as predicted based on oscillator dynamics (Loehr, Large, & Palmer, 2011). Thus, the dynamics of coupled oscillation may underlie people's ability to maintain precise temporal coordination despite fluctuations in their co-performers' timing.

Entrainment based on auditory perception is also evident when people engage in conversation. Several studies have shown that the body movements of two people having a conversation become aligned in time even when they are not able to see each other. For example, two people's postural sway (automatic movements that keep the body in a stable posture) is more similar when they discuss cartoon pictures with each other than when each discusses them with a confederate. This is true whether the participants can see each other or not (Shockley, Santana, & Fowler, 2003). Shared postural sway is greater when conversation partners speak more quickly and produce words with similar stress patterns, suggesting that the coordination of articulatory actions may drive the entrainment (Shockley, Baker, Richardson, & Fowler, 2007). Entrainment may also explain why people's eye movements become temporally aligned when they converse with each other (Richardson, Dale, & Kirkham, 2007).

Top-down influences on entrainment. Entrainment can be modulated top-down by, for example, people's intentions to coordinate. When pairs of participants swing pendulums in sight of each other, but under instructions to maintain their own preferred rate, they show similar patterns of coordination as when they intend to coordinate with each other. However, this unintentional coordination is weaker than intentional coordination (Schmidt & O'Brien, 1997). Similarly, the intention to coordinate enhances entrainment in young children. They are more accurate at synchronizing their actions with a drum-beat, and can synchronize with a wider range of intervals, when they drum along with an experimenter than when they drum with a mechanical device (Kirschner & Tomasello, 2009). Together, these findings indicate that the intention to coordinate enhances entrainment, and suggest that entraining with a more adept partner may allow novices to discover new ways of performing.

Indirect benefits of entrainment. Entrainment may benefit joint action indirectly by moderating the relationship between co-actors as well as co-actors' cognitive processing. People who have moved in synchrony with each other report stronger feelings of liking, connectedness, and trust towards each other, and are more likely to cooperate with each other in economic games, than people who have moved asynchronously (Hove & Risen, 2009; Wiltermuth & Heath, 2009). Participants who move in synchrony with an experimenter remember more words spoken by the experimenter and better recognize the experimenter's face in surprise memory tests following the interaction (Macrae, Duffy, Miles, & Lawrence, 2008), suggesting that entraining with a partner may increase the attention allocated to that partner. Participants who move in synchrony with a partner also show increased perceptual sensitivity to object motion, as well as increased responsiveness to their partner's motion during a subsequent joint action, compared to pairs who have not moved in synchrony (Valdesolo, Ouyang, & DeSteno, 2010). Thus, joint performance may benefit from improvements in individual memory and perceptual processing that result from having been entrained with a partner.

In sum, entrainment can explain how people's actions become tightly coordinated in time during joint action. When people intend to synchronize their actions with each other, they display coordination patterns that follow the mathematical principles governing coupled oscillators. Entrainment also occurs when people engage in joint actions without the explicit goal to synchronize their movements, as evident in the coupling of body and eye movements between conversation partners and in the weaker temporal coordination that arises when people can simply see each other's movements. Entrainment may also benefit joint action indirectly, by increasing feelings of interpersonal affiliation and improving cognitive processing during and after the synchronized action.

Perception-Action Matching: Similarity in Goals and Movements

Whereas entrainment can explain why people align their actions in time, perception-action matching can explain why people produce similar actions or actions with similar perceptual consequences. Perception-action matching relies on the common representations that have been proposed to underlie perception and action (Hommel et al., 2001; Jeannerod, 1999; Prinz, 1997). These common representations allow people to match the actions they perceive onto their own action repertoires. This match can be based on movement similarity, as when observing someone dance activates a representation of dancing movements, or on similarity of actor-object relations, as when observing someone grasp an apple activates a representation of grasping round objects (Knoblich et al., 2011). Matching a perceived action or actor-object relation onto one's own action repertoire can induce in the perceiver a tendency to produce an action similar to the perceived action (Brass, Bekkering, & Prinz, 2001; Stürmer, Aschersleben, & Prinz, 2000) and can allow the perceiver to make predictions about how the perceived action will unfold (Sebanz & Knoblich, 2009; Wilson & Knoblich, 2005). Both of these processes can lead to coordination even in the absence of an intention to coordinate (i.e., emergent coordination; Knoblich et al., 2011). Both also support planned joint action, as we describe below.

Perception-action matching during observation. Much of the evidence for perception-action matching comes from studies that involve one individual observing another individual's action. Neurophysiological studies have established that in monkeys, groups of neurons discharge both when an action is observed and when it is executed (Rizzolatti, Fadiga, Gallese, & Fogassi, 1996), and in humans, similar brain regions are activated when an action is perceived and when it is executed (Rizzolatti & Craighero, 2004; Rizzolatti & Sinigaglia, 2010). The behavioral literature has in turn established that observing someone produce an action facilitates production of a similar action by the observer. For example, people are faster to perform a grasping action when they concurrently watch a video of a hand performing a grasping action compared to a spreading action (Stürmer et al., 2000). Similar facilitation effects occur with goal-directed actions; observing someone kick a ball facilitates foot responses and observing someone type on a keyboard facilitates finger responses (Bach & Tipper, 2007). Furthermore, perceived actions interfere with produced actions when the two do not correspond: People's vertical arm movements are more variable when they concurrently observe horizontal arm movements compared to when they observe vertical arm movements (Kilner, Paulignan, & Blakemore, 2003). Together, these findings indicate that observing an action leads to automatic activation of that action in the observer.

The effects of perceiving an action on production of that action can be modulated by the intentions of the perceiver or the actor. This is important for joint action because it often requires complementary rather than corresponding actions. In an investigation of how an observer's intentions modulate action facilitation, van Schie, van Waterschoot, and Bekkering (2008) instructed participants to either imitate an observed co-actor's grip or to produce a complementary grip (as if taking the object from the co-actor). On some trials, participants responded to a colour cue by producing one of the two grips, regardless of the interaction context. Participants were faster to produce the observed grip in the imitation context, but were faster to produce the complementary grip in the complementary action context. Thus, the observer's intention to produce a complementary action overrode the automatic activation of imitative actions. Likewise, the actor's intentions modulate the movements induced by action observation. For example, observing a person balancing on a foam roller does not induce imitative movements (body tilt in the same direction as the actor's) but rather induces compensatory movements (body tilt in the opposite direction to the actor's, as would be necessary for the actor to maintain balance; Sebanz & Shiffrar, 2007). Together, these studies on action observation suggest that perception-action matching may facilitate joint action by inducing similar or complementary action tendencies in people who perceive each other's actions.

Perception-action matching in joint action. Perception-action matching in joint action has mainly been investigated in settings that involve conversation. This research shows that perception-action matching leads people to produce similar behaviors even when such similarity is not an explicit goal of the interaction. For example, people tend to mimic the actions of their conversation partners: they are more likely to produce foot shaking movements when they converse with a partner who shakes his or her foot than when they converse with a partner who rubs his or her face (Chartrand & Bargh, 1999). Participants are not aware of this tendency to mimic; it is thought to occur because observing a partner's actions activates representations of those actions, which are then manifest as overt imitation. This manifestation of perception-action matching may benefit joint action indirectly by increasing interpersonal affiliation. People whose actions are mimicked during conversation report liking their conversation partner more (Chartrand & Bargh, 1999) and are more likely to help their partner (van Baaren, Holland, Kawakami, & van Knippenberg, 2004) than people whose actions have not been mimicked. Thus, perception-action matching may not only support joint action directly by facilitating action coordination, but may also benefit joint action indirectly by moderating the relationship between co-actors.

Action simulation. The findings reviewed so far indicate that perception-action matching can support joint action by inducing people to produce similar or complementary actions. Perception-action matching can also support joint action in a second way. Once a perceived action or actor-object relation is matched to an action in the perceiver's action repertoire, the perceiver can use his or her own motor system to generate predictions about how that action will unfold (Sebanz & Knoblich, 2009; Wilson & Knoblich, 2005). This process, referred to as action simulation, uses the same internal forward models that predict the sensory consequences of one's own actions (Miall & Wolpert, 1996) to predict the consequences of other people's actions (Wolpert, Doya, & Kawato, 2003). Action simulation typically occurs during action observation, but can also be triggered by knowledge about an upcoming action (Kilner et al., 2004).

Evidence that one's own action repertoire can be used to generate predictions about others' actions comes from studies showing that experience in producing an action enhances people's ability to predict the outcome of similar actions. Professional basketball players are better able to predict whether or not an observed basketball shot will be successful than are sports journalists, who have extensive experience observing, but not producing, basketball shots (Aglioti et al., 2008). Similarly, people are better able to predict the outcome of actions they themselves have produced than actions produced by others. For example, people are better able to predict where a dart will land when they watch videos of their own throwing movements than when they watch others' throwing movements (Knoblich & Flach, 2001), Thus, the more similar an action is to the observer's action repertoire, the more accurately its outcome can be predicted.

Action simulation in joint action. Action simulation can support joint action in several ways. First, simulation may allow people to make predictions about their co-actors' upcoming actions and modify their own actions accordingly. Becchio and colleagues asked participants to move an object from one location to another in two conditions (Becchio, Sartori, Bulgheroni, & Castiello, 2008). The end location was exactly the same in both conditions, except that it took the form of another person's hand in one condition and of a hand-shaped pad in the other. Participants produced different movement trajectories in the two conditions; the object was approached more slowly and with a smaller grip aperture, and was also placed in its end location more slowly, when the end location was another person's hand. Thus, participants handled the object in a way that made it easier for the other person to grasp it in the joint action condition. This suggests that participants may have simulated the other person's upcoming action (grasping the object) and modified their own action kinematics accordingly.

Second, action simulation guides attention during joint action. Welsh and colleagues asked pairs of participants to sit across from each other and detect targets that appeared on a screen in between them (Welsh et al., 2005). When people perform a target detection task alone, they are slower to detect targets that appear in the same location as an immediately preceding target, a phenomenon known as inhibition of return. Welsh and colleagues showed that inhibition of return also occurred when participants were asked to detect a target to which they had just watched their co-actor respond. This suggests that people simulated their partner's action, which led to similar inhibition of attention toward their partner's location as if they had responded to that location themselves. This effect also occurred when the participants could see the effect of the partner's response (i.e., the illumination of a response button) but not the partner's movements, indicating that knowledge of the partner's action was sufficient to induce simulation (Welsh, Lyons, et al., 2007). Flanagan and Johansson (2003) tracked participants' eye movements as they stacked a series of blocks and as they watched another person stack the blocks. When people move objects themselves, their eyes make predictive movements toward upcoming objects and locations rather than reactive movements following the motion of the objects. In the observation condition, participants' eye movements showed this same pattern of predictive eye movements, suggesting that the participants simulated the observed actor's movements and were thus able to attend to the objects they expected to be manipulated next. Thus, predictive action simulation may guide the allocation of attention during joint action.

Third, action simulation may support temporal coordination between multiple performers, in two ways. First, in tasks such as ensemble music performance that require performers to produce independent yet coordinated actions, each performer may run parallel simulations of their own and their partners' actions. In support of this hypothesis, Keller, Knoblich, and Repp (2007) showed that pianists were better able to synchronize one part of a duet with a second part that was recorded months earlier by themselves than by another performer. This suggests that they may have relied on internal models to predict the timing of their own duet part and the part they synchronized with, and that this prediction was most successful when the part they synchronized with closely matched their own (in this case, because it was their own earlier action). Converging evidence for this hypothesis comes from a study of duet performance which showed that pianists are better able to coordinate with partners to whom they are more similar in terms of preferred performance rate (Loehr & Palmer, in press).

Second, in tasks that require performers to produce events jointly rather than independently, co-performers may generate predictions about the temporal consequences of their combined actions rather than separate predictions for each person's actions. To test this hypothesis, pairs of participants were asked to track a moving target on a computer screen (Knoblich & Jordan, 2003). One participant controlled the key that decelerated the tracker, and the other controlled the key that accelerated it. After training during which participants had access to feedback concerning the timing of each other's keypresses, pairs of participants were just as good at performing the task as individual participants who controlled both keys. These findings suggest that with training, participants were able to generate predictions about the timing of their combined actions and use these predictions for effective coordination.

Finally, action simulation may be modulated by the relationships between interaction partners. Kourtis and colleagues asked groups of three participants to take turns lifting an object (Kourtis, Sebanz, & Knoblich, 2010). EEG was recorded from one participant, who gave the object to one member of the group (the interaction partner) on some trials, but never interacted with the other group member (the loner). Each group member lifted the object alone on some trials. Participants' anticipatory motor activity, reflecting simulation of another group member's upcoming action, was stronger when they expected their interaction partner to lift the object alone than when they expected the loner to lift the object alone. Thus, the degree to which participants simulated others' actions depended on their social relationship to the observed actor.

In sum, perception-action matching supports joint action in two ways. Matching an observed action onto one's own motor system can facilitate production of that same action or a complementary action, depending on the goals of the actor and the observer. This process can lead interaction partners to produce matching behaviors, as evident in non-conscious mimicry during conversation, and may indirectly benefit joint action by increasing interpersonal affiliation. Matching an observed action onto one's own motor system also allows predictions to be generated about unfolding actions. This predictive action simulation allows people to modify their actions in relations to others' actions, guides people's attention during joint action, and helps people predict when a partner will act in order to align the individual components of joint actions in time.

It is important to remember that close links between perception and action can lead to emergent coordination whether people plan to coordinate with each other or not. These links also support planned joint action such as music performance, in which two or more people intend to coordinate their behavior. However, planned joint action also requires mental representations and processes that go beyond the perception-action links discussed above. We turn to these processes next.

Shared Representations and Perceptions

Shared representations of the desired outcome of joint action, as well as each person's part in achieving that outcome, can support joint action even in the absence of direct perception of a partner's actions. These representations may specify the particular actions the co-actor must produce or the task rules under which the co-actor operates, and can be used to guide action planning, monitoring, and control (Vesper et al., 2010). People engaged in joint action may also incorporate their partners' perceptions into their representations of the task, by taking the partner's perspective or by inferring what the partner can and cannot perceive (Samson, Apperly, Braithwaite, Andrews, & Bodley Scott, 2010). This too can facilitate monitoring of and adaptation to a partner's actions. We describe each of these processes in turn.

Shared Task Representations

In ensemble performance of Western music, each performer's part is typically specified in advance in a musical score. The pianist and violinist in our opening example would have used the score to learn their individual parts and how they relate to each other to form a musical whole (Keller, 2008). They may also have used the score to learn the details of each other's parts. This musical example illustrates the representations that may be shared in planned joint action. Each person must minimally represent his own part in the joint action (the pianist's part) as well as the desired joint outcome (the musical piece), and there must be some awareness that the joint outcome cannot be achieved alone (Vesper et al., 2010). Each person may also represent his co-performer's part in the joint action (the violinist's part), though this is not always necessary for successful joint action. For example, the pianist may simply represent her own part in the musical performance and have the goal of remaining temporally coordinated with the violinist. However, people often form representations of each other's tasks during joint action when doing so is not necessary for, or is even detrimental to, performance (Sebanz, Knoblich, & Prinz, 2003; Tsai, Kuo, Jing, Hung, & Tzeng, 2006). These representations of a co-actor's task modulate each person's action planning, control, and monitoring.

Action planning. The first evidence that representations of a co-actor's task influence action planning came from studies that examined joint performance of a Simon task (Sebanz et al., 2003; Sebanz, Knoblich, & Prinz, 2005). When participants perform a Simon task alone, they are required to produce one of two action alternatives, such as a left or right button press, in response to a stimulus, such as the color of a ring on a finger. A stimulus feature that is irrelevant to the task, such as the pointing direction of the finger, induces a response conflict if it is spatially incompatible with the required response. For example, participants are slower to respond to a red ring with a right button press if the ring sits on a finger that points to the left (in the direction of the alternative response). In the joint Simon task, the two action alternatives are distributed between two people, and spatial incompatibility effects are still observed: Participants are slower to respond to a red ring with a right button press if the ring sits on a finger that points to the left, even when it is the co-actor rather than the participant himself who is responsible for left button presses. No such effects are observed when participants perform their half of the task alone (i.e., when participants are responsible for the right button press and no one is responsible for the left button press). Thus, participants in the joint Simon task form representations of their co-actor's task, which influence their own action planning (in this case, response selection).

Shared representations influence action planning even in the absence of direct perception of the co-actor's actions. The spatial compatibility effects observed in the joint Simon task also occur when participants merely believe that another person is responsible for the second action alternative (Ruys & Aarts, 2010; Tsai, Kuo, Hung, & Tzeng, 2008; but see Welsh, Higgins, Ray, & Weeks, 2007, for different findings). This suggests that participants' action planning is constrained by representations of their partners' tasks formed in advance of the joint action. Furthermore, there need not be overlap between the spatial features of the stimulus and the co-actors' responses for the partner's task to influence action planning. Rather, participants represent the arbitrary task rules that govern their co-actors' behavior (Atmaca, Sebanz, & Knoblich, 2011; Ramnani & Miall, 2004; Sebanz, Knoblich, & Prinz, 2005). For example, Ramnani and Miall (2004) trained pairs of participants to produce finger movements in response to arbitrary visual stimuli (shapes) whose color determined which participant should respond. The brain activity of one member of the pair was then measured using fMRI while he or she performed the task under the belief that the other member of the pair performed the task outside the scanner. Stimuli to which the co-actor was meant to respond yielded preparatory motor activity in the scanned participant, indicating that the arbitrary task rules under which the co-actor responded activated motor representations in the scanned participant. Thus, co-representation of arbitrary task rules influences action planning during joint action.

Action control and monitoring. Shared task representations also influence action control. Evidence for this comes from EEG recordings of participants' brain activity while they performed the Simon task described above (Sebanz, Knoblich, Prinz, & Wascher, 2006; Tsai et al., 2008). Recall that in the joint Simon task, participants respond to one stimulus color while the co-actor responds to the other stimulus color. Thus, the participant must not respond when it is the co-actor's turn to do so; trials of this type are referred to as no-go trials. Participants can also perform this task individually: they can respond to one stimulus color but not the other, with no co-actor involved. No-go trials can then be compared under conditions in which the stimulus refers to a co-actor's task (joint condition) and in which the stimulus refers to no-one's task (individual condition). EEG activity measured during no-go trials reveals a more pronounced positive event-related potential 300-500 ms after the stimulus in the joint condition compared to the individual condition (Sebanz, Knoblich, et al., 2006). This ERP component is thought to reflect action inhibition, which is stronger when it is the co-actor's turn to act than when it is no one's turn to act. Thus, this finding indicates that co-representation leads participants to recruit action control processes to ensure that they do not act when it is the co-actor's turn to do so.

Representations of a co-actor's task also govern action monitoring. Schuch and Tipper (2007) investigated action monitoring and control by having participants perform a stop signal task with a co-actor. In stop signal tasks, people are required to respond to a target as quickly as possible. On some trials, they must stop the response when a stop signal is presented shortly after the target. The timing of the stop signal is manipulated so that participants sometimes successfully stop the response (engaging inhibitory processes) and sometimes cannot stop the response (producing errors). When people perform this task alone, they are typically slower at responding to the target that immediately follows the stop signal, both when they inhibited the previous action and when they made an error. These same aftereffects are evident when participants respond to a target after observing a co-actor successfully inhibit their action or make an error. Thus, participants monitor their partners' actions for errors, which affect participants' performance just as their own errors do; similarly, a partner's action inhibition elicits similar inhibitory mechanisms in one's own actions.

Shared task representations can also influence multisensory integration. In a recent study, participants were asked to judge the location of a tactile stimulus that was presented on the top or bottom of a foam cube that they held with both hands (Heed, Habets, Sebanz, & Knoblich, 2010). When participants perform this task alone, visual distractors (lights) that are presented at the elevation opposite to the tactile stimulation (top versus bottom of the cube) reduce the speed and accuracy with which participants judge the location of the tactile stimulus, indicating interference between the two modalities (Spence, Pavani, & Driver, 2004). However, when participants performed the same task with a co-actor who responded to the location of the visual distractors, this cross-modal interference was reduced. In other words, participants were better able to ignore the distractors when another participant responded to them. This finding suggests that co-representation may facilitate rather than interfere with task performance when co-actors respond to stimuli from different sensory modalities. However, this facilitation only occurred when the co-actor sat within the participant's peripersonal space, indicating that co-representation effects may differ depending on the spatial relationship between co-actors (see Guagano, Rusconi, & Umiltà, 2010; Welsh, 2009, for related findings).

Social context. Finally, the tendency to share representations with a co-actor may be modulated by the social context of the interaction and the characteristics of the co-actors. Iani and colleagues found typical co-representation effects in the joint Simon task in a cooperative but not in a competitive context (Iani, Anelli, Nicoletti, Arcuri, & Rubichi, 2011). Hommel and colleagues found the joint Simon effect when the participants' co-actor was friendly and cooperative, but not when the co-actor was competitive and intimidating (Hommel, Colzato, &

van den Wildenberg, 2009). Kuhbandner, Pekrun, and Maier (2010) found co-representation effects after participants watched movies that induced a positive or neutral mood, but not after they watched movies that induced a negative mood. Thus, a negative relationship with the co-actor or a negative mood reduces people's tendency to co-represent their partners' actions. The ability to infer others' mental states may also influence co-representation. Ruys and Aarts (2010) used the 'mind in the eyes' test, in which participants judged the emotional state of pairs of eyes, to determine participants' ability to infer mental states. Those who were less able to infer others' mental states only showed co-representation in a competitive context, whereas those who were better able to infer others' mental states showed co-representation effects in both competitive and cooperative contexts. In line with this finding, neurological patients with impairments in mental state attribution do not show the joint Simon effect, unless they are explicitly instructed to pay attention to their co-actor (Humphreys & Bedford, 2011). However, Sebanz, Knoblich, Stumpf, and Prinz (2005) found similar co-representation effects in a group of high-functioning individuals with autism and a matched control group of adolescents and adults, suggesting that deficits in understanding others' mental states do not necessarily preclude co-representation.

In sum, shared representations of co-actors' tasks are formed even when they are not necessary for successful performance of a joint task, and despite the fact that they sometimes interfere with performance of individual tasks. Shared task representations can arise even without visual access to the co-actor's actions, suggesting that advance knowledge of a co-actor's task can become incorporated into one's own action plans. Shared task representations influence not only action planning but also action control and stimulus processing, though these effects may be modulated by the social context of the joint action. Overall, then, people engaged in joint actions take their co-actors' tasks into account when producing their own actions. Can the same be said for their co-actors' perceptions?

Shared Perceptions

Joint action is supported by people's ability to assess what their co-actors can and cannot perceive. In duet music performance, both performers attend to the overall sound to ensure that they produce a cohesive performance (Keller, 2008). The performers also rely on shared perceptual cues that both attend to at key moments for temporal coordination (Williamon & Davidson, 2002). In other situations, people engaged in joint action may have access to the same perceptual information but from different perspectives, such as when two people perform a task while sitting on opposite sides of a table. In this case, they may take each other's perspectives when performing the joint action by mentally rotating themselves into each other's orientation (Kessler & Thompson, 2010) or by switching from an egocentric to an allocentric perspective that allows for an agent-neutral view on shared visual stimuli (Bockler, Knoblich, & Sebanz, 2011). People engaged in joint action may also have access to different perceptual information, in which case they may infer what their co-actors can and cannot perceive. Sharing perceptions with a co-actor, whether through joint attention, perspective-taking, or inference, facilitates coordination by allowing people to modify their own actions according to what their partner perceives.

Shared perceptual cues. In music performance, performers can attend to both their own part in the performance and the jointly produced performance. Keller (2005) asked drummers to perform one rhythm while a computer simultaneously produced a different rhythm, in which the performer's rhythm was embedded. The drummers were able to subsequently reproduce the computer's rhythm, indicating that they attended to the aggregate rhythm in addition to their own embedded rhythm. However, musicians do not always have access to auditory feedback from each other's actions. In this case, visual cues may be used to guide attention. Williamon and Davidson (2002) recorded duet pianists' practice sessions and showed that with practice, the pianists increasingly relied on eye contact to guide coordination when one or both performers had not yet begun to play. Goebl and Palmer (2009) showed that visual information about a given performer's movements (the height to which fingers were raised above the keys) was exaggerated when the co-performer did not have access to auditory feedback from the performance. These findings suggest that different perceptual cues may be used to draw a co-performer's attention, and thus facilitate coordination, depending on the sensory information available to each performer.

In the visual domain, being able to assess what a partner can perceive facilitates performance of joint tasks. Brennan and colleagues asked pairs of participants to perform a joint visual search task (e.g., looking for the letter O in an array of Qs). They compared performance in a condition in which partners shared information about where each of them searched the display (through cursors indicating each person's gaze location) with a condition in which partners could only communicate verbally. Performance was faster when partners could perceive each other's gaze. In the shared gaze condition, pairs were better able to distribute the search task between themselves (Brennan, Chen, Dickinson, Neider, & Zelinsky, 2008), and they used gaze to communicate the target location instead of relying on verbal communication (Neider, Chen, Dickinson, Brennan, & Zelinsky, 2010). Thus, assessing the location of a partner's attention can facilitate coordination in joint action.

Perspective-taking. People engaged in joint action may share access to the same visual information but from different perspectives. Recent evidence suggests that people may take their partner's perspective into account in such situations. When one person mentally rotates an image while a partner looks at the same image, the person is slower to perform a small rotation (large rotation from the partner's perspective) and faster to perform a large rotation (small rotation from the partner's perspective) compared to when the partner is not looking at the image (Böckler, Knoblich, & Sebanz, 2011). This suggests that engaging in joint attention from different perspectives triggers a switch from an egocentric perspective to an allocentric, agent-neutral perspective. People may also compute another person's perspective when that perspective contains different perceptual information than their own. Samson and colleagues (Samson et al., 2010) asked participants to judge their own or another person's visual perspective (to indicate how many discs, hanging on a virtual wall, could be seen by themselves or a computer avatar). People were slower to judge their own perspective when the other person's perspective contained different information (a different number of discs) than their own perspective. Thus, participants computed the avatar's perspective even though doing so was detrimental to their own performance. This parallels the findings on task co-representation, which occurs even in circumstances where taking another's task into account interferes with one's own performance.

In sum, sharing perceptions with a co-performer facilitates coordination in several ways. Providing a jointly attended perceptual cue to a partner facilitates the temporal alignment of actions in tasks such as music performance. Being able to assess what a partner can perceive allows for more efficient communication and task distribution. When partners have different perspectives on the same visual information, they may automatically compute each other's perspective; computing another's perspective may occur even when co-actors have access to different perceptual information. Sharing perceptions and attention with a partner is thus an important component in understanding joint action.

Conclusions

The two musicians who performed a duet at the beginning of this chapter relied on several processes to achieve their joint performance goals. Shared representations and perceptions of their own, their partner's, and their combined tasks governed how the performers planned and monitored their own actions and monitored each other's actions. At the same time, close links between the two performers' actions and perceptions allowed them to precisely coordinate their actions in time and to make predictions about their partners' actions as they unfolded. Neither perception-action links nor shared representations alone can explain how the pianists performed their duet. For example, shared representations defining the actions each performer should produce are not sufficient to guarantee that those actions will be aligned in time. Likewise, coupling between perceptions and actions is not sufficient to specify who should perform which actions to bring about a joint outcome. Both of these elements are important for the musicians' successful performance; how processes related to perception-action coupling and shared representations work together during duet performance and other joint actions remains a question for future research. Nevertheless, it is clear that the actions and perceptions of the pianist cannot be fully understood without considering the actions and perceptions of the violinist. Given the common occurrence of joint action in everyday life, neither can any individual's actions and perceptions be fully understood without considering the actions and perceptions of those with whom he or she interacts.

References

Atmaca, S., Sebanz, N., & Knoblich, G. (2011). The joint Flanker effect: Sharing tasks with real and imagined co-actors. *Experimental Brain Research, 211,* 371-385.

Bach, P., & Tipper, S. P. (2007). Implicit action encoding influences personal-trait judgments. *Cognition, 102*, 151-178.

Becchio, C., Sartori, L., Bulgheroni, M., & Castiello, U. (2008). The case of Dr. Jekyll and Mr. Hyde: A kinematic study on social intention. *Consciousness and Cognition, 17*, 557-564.

Böckler, A., Knoblich, G., & Sebanz, N. (2011). Giving a helping hand: Effects of joint attention on mental rotation of body parts. *Experimental Brain Research, 211,* 531-545.

Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica, 106*, 3-22.

Brass, M., & Heyes, C. (2005). Imitation: Is cognitive neuroscience solving the correspondence problem? *Trends in Cognitive Science, 9*, 489-495.

Brennan, S. E., Chen, X., Dickinson, C. A., Neider, M. B., & Zelinsky, G. J. (2008). Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition, 106*, 1465-1477.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology, 76*, 893-910.

Clark, H. H. (1996). *Using language*. Cambridge, UK: Cambridge University Press.

Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action prediction. *Nature, 424*, 769-771.

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

Guagano, D., Rusconi, E., & Umiltà, C. A. (2010). Sharing a task or sharing space? On the effect of the confederate in action coding in a detection task. *Cognition, 114*.

Haken, H., Kelso, J. A. S., & Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological Cybernetics, 51*, 347-356.

Heed, T., Habets, B., Sebanz, N., & Knoblich, G. (2010). Others' actions reduce cross-modal integration in peripersonal space *Current Biology, 20*, 1345-1349.

Hommel, B., Colzato, L. S., & van den Wildenberg, W. P. M. (2009). How social are task representations? *Psychological Science, 20*, 794-798.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences, 24*, 849-937.

Hove, M. J., & Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases affiliation. *Social Cognition, 27*, 949-960.

Humphreys, G. W., & Bedford, J. (2011). The relations between joint action and theory of mind: A neuropsychological analysis. *Experimental Brain Research, 211*, 357-369.

Iani, C., Anelli, F., Nicoletti, R., Arcuri, L., & Rubichi, S. (2011). The role of group membership on the modulation of joint action. *Experimental Brain Research, 211*, 439-445.

Jeannerod, M. (1999). The 25th Bartlett lecture. To act or not to act: Perspective on the representation of actions. *The Quarterly Journal of Experimental Psychology, 52A*, 1-29.

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

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

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.

Kelso, J. A. S. (2001). Self-organizing dynamical systems. In N. J. Smesler & P. B. Baltes (Eds.), *International Encyclopaedia of Social and Behavioral Sciences* (pp. 13844-13850). Amsterdam: Pergamon.

Kessler, K., & Thomson, L. A. (2010). The embodied nature of spatial perspective taking: Embodied transformation versus sensorimotor interference. *Cognition, 114*, 72-88.

Kiefer, M., & Barsalou, L. W. (this volume). Grounding the human conceptual system in perception, action, and introspection.

Kilner, J. M., Paulignan, Y., & Blakemore, S. J. (2003). An interference effect of observed biological movement on action. *Current Biology, 13*, 522-525.

Kilner, J. M., Vargas, C., Duval, S., Blakemore, S.-J., & Sirigu, A. (2004). Motor activation prior to observation of a predicted movement. Nature Neuroscience, 7, 1299-1301.

Kirschner, S., & Tomasello, M. (2009). Joint drumming: Social context facilitates synchronization in preschool children. *Journal of Experimental Child Psychology, 102*, 299-314.

Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological research on joint action: Theory and data. *The Psychology of Learning and Motivation, 54*, 59-101.

Knoblich, G., & Flach, R. (2001). Predicting the effects of actions: Interactions of perception and action. *Psychological Science, 12*, 467-472.

Knoblich, G., & Jordan, J. S. (2003). Action coordination in groups and individuals: Learning anticipatory control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 1006-1016.

Kourtis, D., Sebanz, N., & Knoblich, G. (2010). Favouritism in the motor system: Social interaction modulates action simulation. *Biology Letters, 6*, 758-761.

Kuhbandner, C., Pekrun, R., & Maier, M. A. (2010). The role of positive and negative affect in the "mirroring" of other persons' actions. *Cognition & Emotion, 24*, 1182-1190.

Large, E. W. (2008). Resonating to musical rhythm: Theory and experiment. In S. Grondin (Ed.), *Psychology of Time* (pp. 189-232). Bingley, UK: Emerald Group Publishing Ltd.

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

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.

Loehr, J. D., & Palmer, C. (in press). Temporal coordination between performing musicians. *Quarterly Journal of Experimental Psychology*.

Macrae, C. N., Duffy, O. K., Miles, L. K., & Lawrence, J. (2008). A case of hand waving: Action synchrony and person perception. *Cognition, 109*, 152-156.

Marsh, K. L., Richardson, M. J., & Schmidt, R. C. (2009). Social connection through joint action and interpersonal coordination. *Topics in Cognitive Science, 1*, 320-339.

Miall, R. C., & Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural Networks, 9*.

Néda, Z., Ravasz, E., Brechet, Y., Vicsek, T., & Barabási, A.-L. (2000). The sound of many hands clapping. *Nature 403*, 849-850.

Neider, M. B., Chen, X., Dickinson, C. A., Brennan, S. E., & Zelinsky, G. J. (2010). Coordination spatial referencing using shared gaze. *Psychonomic Bulletin & Review, 17*, 718-724.

Pikovsky, A., Rosenblum, M., & Kurths, J. (2001). *Synchronization: A universal concept in nonlinear sciences*. Cambridge, UK: Cambridge University Press.

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology, 9*, 129-154.

Provine, R. R. (1986). Yawning as a stereotyped action pattern and releasing stimulus. *Ethology, 72*, 109-122.

Ramnani, N., & Miall, R. C. (2004). A system in the human brain for predicting the actions of others. *Nature Neuroscience, 7*, 85-90.
Rasch, R. A. (1979). Synchronization in performed ensemble music. *Acustica, 43*, 121-131.
Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The art of conversation is coordination: Common ground and the coupling of eye movements during dialogue. *Psychological Science, 18*, 407-413.
Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience, 27*, 169-192.
Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research, 3*, 131-141.
Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. *Nature Reviews Neuroscience, 11*, 264-274.
Ruys, K. I., & Aarts, H. (2010). When competition merges people's behavior: Interdependency activates shared action representations. *Journal of Experimental Social Psychology, 46*, 1130-1133.
Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley Scott, S. E. (2010). Seeing it their way: What other people see is calculated by low-level and early acting processes. *Journal of Experimental Psychology: Human Perception & Performance, 36*, 1255-1266.
Schmidt, R. C., Bienvenu, M., Fitzpatrick, P. A., & Amazeen, P. G. (1998). A comparison of intra- and interpersonal interlimb coordination: Coordination breakdowns and coupling strength. *Journal of Experimental Psychology: Human Perception & Performance, 24*, 884-900.
Schmidt, R. C., Carello, C., & Turvey, M. T. (1990). Phase transitions and critical fluctuations in the visual coordination of rhythmic movements between people. *Journal of Experimental Psychology: Human Perception & Performance, 16*, 227-247.
Schmidt, R. C., & O'Brien, B. (1997). Evaluating the dynamics of unintended interpersonal coordination. *Ecological Psychology, 9*, 189-206.
Schmidt, R. C., & Richardson, M. J. (2008). Dynamics of interpersonal coordination. In A. Fuchs & V. Jirsa (Eds.), *Coordination: Neural, behavioural, and social dynamics* (pp. 281-308). Heidelberg: Springer-Verlag.
Schuch, S., & Tipper, S. P. (2007). On observing another person's actions: Influences of observed inhibition and errors. *Perception & Psychophysics, 69*, 828-837.
Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences, 10*, 71-76.
Sebanz, N., & Knoblich, G. (2009). Prediction in joint action: What, when, and where. *Topics in Cognitive Science, 1*, 353-367.
Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: Just like one's own? *Cognition, 88*, B11-B21.
Sebanz, N., Knoblich, G., & Prinz, W. (2005). How two share a task: Corepresenting stimulus-response mappings. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 1234-1246.
Sebanz, N., Knoblich, G., Prinz, W., & Wascher, E. (2006). Twin peaks: An ERP study of action planning and control in coacting individuals. *Journal of Cognitive Neuroscience, 18*, 859-870.
Sebanz, N., Knoblich, G., Stumpf, L., & Prinz, W. (2005). Far from action-blind: Representation of others' actions in individuals with autism. *Cognitive Neuropsychology, 22*, 433-454.
Sebanz, N., & Shiffrar, M. (2007). Bodily bonds: Effects of social context on ideomotor movements. In Y. Rossetti, M. Kawato & P. Haggard (Eds.), *Sensorimotor foundations of higher cognition (Attention and performance XXII).* Oxford, UK: Oxford University Press.

Shockley, K., Baker, A. A., Richardson, M. J., & Fowler, C. A. (2007). Articulatory constraints on interpersonal postural coordination. *Journal of Experimental Psychology: Human Perception & Performance, 33*, 201-208.

Shockley, K., Santana, M.-V., & Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception & Performance, 29*, 326-332.

Spence, C., Pavani, F., & Driver, J. (2004). Spatial constraints on visual-tactile cross-modal distractor congruency effects. *Cognitive, Affective & Behavioral Neuroscience, 4*, 148-169.

Stürmer, B., Aschersleben, G., & Prinz, W. (2000). Correspondence effects with manual gestures and postures: A study of imitation. *Journal of Experimental Psychology: Human Perception & Performance, 26*, 1746-1759.

Tsai, C. C., Kuo, W. J., Hung, D. L., & Tzeng, O. J. L. (2008). Action co-representation is tuned to other humans. *Journal of Cognitive Neuroscience, 20*, 2015-2024.

Tsai, C. C., Kuo, W. J., Jing, J. T., Hung, D. L., & Tzeng, O. J. L. (2006). A common coding framework in self-other interaction: Evidence from joint action task. *Experimental Brain Research, 175*, 353-362.

Valdesolo, P., Ouyang, J., & DeSteno, D. (2010). The rhythm of joint action: Synchrony promotes cooperative ability. *Journal of Experimental Social Psychology, 46*, 693-695.

van Baaren, R. B., Holland, R. W., Kawakami, K., & van Knippenberg, A. (2004). Mimicry and prosocial behavior. *Psychological Science, 15*, 71-74.

van Schie, H. T., van Waterschoot, B. M., & Bekkering, H. (2008). Understanding action beyond imitation: Reversed compatibility effects of action observation in imitation and joint action. *Journal of Experimental Psychology: Human Perception & Performance, 34*, 1493-1500.

Vesper, C., Butterfill, S., Knoblich, G., & Sebanz, N. (2010). A minimal architecture for joint action. *Neural Networks, 23*, 998-1003.

Welsh, T. N. (2009). When 1+1=1: The unification of independent actors revealed through joint Simon effects in crossed and uncrossed effector conditions. *Human Movement Science, 28*, 726-737.

Welsh, T. N., Elliott, D., Anson, J. G., Dhillon, V., Weeks, D. J., Lyons, J. L., & Chua, R. (2005). Does Joe influence Fred's action? Inhibition of return across different nervous systems. *Neuroscience Letters, 385*, 99-104.

Welsh, T. N., Higgins, L., Ray, M., & Weeks, D. J. (2007). Seeing vs. believing: Is believing sufficient to activate the processes of response co-representation? *Human Movement Science, 26*, 853-866.

Welsh, T. N., Lyons, J., Weeks, D. J., Anson, J. G., Chua, R., Mendoza, J., & Elliott, D. (2007). Within- and between-nervous-system inhibition of return: Observation is as good as performance. *Psychonomic Bulletin & Review, 14*, 950-956.

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

Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin, 131*, 460-473.

Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. *Psychological Science, 20*, 1-5.

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