Cognition 180 (2018) 246–258

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Original Articles

Identifying others' informative intentions from movement kinematics

Luke McEllin*, Natalie Sebanz, Günther Knoblich

Central European University, Budapest, Hungary

ARTICLE INFO

Keywords:
Joint action
Teaching
Motor simulation
Action observation
Intention

ABSTRACT

Previous research has demonstrated that people can reliably distinguish between actions with different instrumental intentions on the basis of the kinematic signatures of these actions (Cavallo, Koul, Ansuini, Capozzi, & Becchio, 2016). It has also been demonstrated that different informative intentions result in distinct action kinematics (McEllin, Knoblich, & Sebanz, 2017). However, it is unknown whether people can discriminate between instrumental actions and actions performed with an informative intention, and between actions performed with different informative intentions, on the basis of kinematic cues produced in these actions. We addressed these questions using a visual discrimination paradigm in which participants were presented with point light animations of an actor playing a virtual xylophone. We systematically manipulated and amplified kinematic parameters that have been shown to reflect different informative intentions. We found that participants reliably used both spatial and temporal cues in order to discriminate between instrumental actions and actions performed with an informative intention, and between actions performed with different informative intentions. Our findings indicate that the informative cues produced in joint action and teaching go beyond serving a general informative purpose and can be used to infer specific informative intentions.

1. Introduction

People derive mental states such as intentions and expectations from observing the movements of others (Cavallo, Koul, Ansuini, Capozzi, & Becchio, 2016; Grèzes, Frith & Passingham, 2004). Using early movement kinematics of perceived actions, observers can discriminate between different instrumental intentions (Cavallo et al., 2016; Manera, Becchio, Cavallo, Sartori, & Castiello, 2011). In addition, informative intentions can also be reflected in kinematics. On the one hand, people acting together produce informative action modulations in order to support interpersonal coordination by facilitating spatial and temporal prediction (Pezzulo, Donnarumma & Dindo, 2013; Vesper & Richardson, 2014; Vesper, Schmitz, Safra, Sebanz, & Knoblich, 2016). On the other hand, parents and teachers modify their movements to support learning through demonstration by highlighting the structure of an action (Brand, Baldwin & Ashburn, 2002). These findings suggest that the same action can be modulated in different ways to convey different informative intentions to an observer.

But can observers actually identify informative intentions based on movement kinematics? The first aim of the present study was to investigate whether people can discriminate actions with informative intentions from actions without informative intentions using kinematic cues. The second aim was to investigate whether people are able to distinguish different interactive intentions based on kinematic cues. Specifically, we asked whether observers can tell whether perceived agents are intending to teach a co-actor or whether they intend to perform a coordinated joint action with a co-actor.

1.1. Perceiving intentions from actions

Much of the research on perception of individuals' intentions has focused on perception of instrumental actions. This research has demonstrated that humans have the ability to derive different mental states of an actor by observing the kinematics of their actions. For instance, people can recognize whether an actor intends to cooperate or compete (Manera et al., 2011), whether or not an actor has a false belief (Grèzes, et al., 2004) or even whether or not an actor has a deceptive intention (Runeson & Frykholm, 1983). Even though these actions are not intended to inform, people can still read mental states from them.

A recent study by Cavallo et al. (2016) demonstrated that people can discriminate observed actors' instrumental intentions based on early kinematic features of the action. In their study participants observed reach to grasp movements of actors intending to grasp a bottle in order to pour from it, or in order to drink. They found that kinematic features such as wrist height and grip aperture predicted how well an observer could discriminate between the two different underlying

* Corresponding author at: Department of Cognitive Science, Central European University, Oktober 6 Street 7, Budapest 1051, Hungary.
E-mail address: mcellin_luke@phd.ceu.edu (L. McEllin).

https://doi.org/10.1016/j.cognition.2018.08.001
Received 12 April 2018; Received in revised form 31 July 2018; Accepted 1 August 2018

intentions. Moreover, the accuracy of participants' discrimination between the two underlying intentions could be modulated by modifying kinematic parameters that predicted classification accuracy. In contrast to many earlier studies, Cavallo et al. (2016) were not only able to show that different intentions can be discriminated, but they could also quantify the contribution of different kinematic parameters to the accuracy of identifying a particular intention.

There is also evidence that movement kinematics carry information about social intentions. Becchio, Sartori, Bulgheroni and Castiello (2008) carried out a study in which participants were required to grasp an object to build a tower together with a co-actor, either with a cooperative intention (build the tower together) or a competitive intention (place the object at the bottom of the tower before the other participant). They showed that compared to competitive actions, cooperative actions had a larger trajectory, were slower, and displayed a smaller grip aperture. Another study by Manera et al. (2011) demonstrated that people could discriminate between cooperative and competitive intentions when perceiving reach to grasp movements. Moreover, participants could still discriminate between competitive and cooperative intentions when viewing point light displays of reach to grasp movements, demonstrating that dynamic kinematic cues were used to discriminate between different intentions.

Evidence obtained in sports experts indicates that identifying intentions from action kinematics taps into motor simulation. Aglioti, Cesari, Romani and Urgesi (2008) demonstrated that expert basketball players could predict the accuracy of a free throw on the basis of the player's kinematics, whereas expert watchers and novices could not. Similarly, Sebanz and Shiffrar (2009) found that expert basketball players could distinguish real passes from fake passes by observing another player's actions, both when the actions were shown in videos and when they were shown as point-light displays. In contrast, novice basketballers were not able to discriminate real and fake passes. These results imply that motor expertise can be a pre-condition for identifying intentions from an observed agent's kinematics.

In sum, previous research shows that movement kinematics provide a rich source of information that observers can use to make predictions about observed agents' intentions. Even when instrumental actions are not intended to inform the observer, they are nonetheless a rich source of information due to dedicated perceptual processing of kinematic cues (Becchio, Cavallo, Begliomini, Sartori, Feltrin, Castiello et al., 2012; Becchio, Manera, Sartori, Cavallo, Castiello, 2012) and people's ability to map observed actions onto their own motor repertoire (Ansuini, Cavallo, Bertone & Becchio, 2015; Rizzolatti and Sinigaglia, 2010).

1.2. Sensorimotor communication in joint action coordination and teaching

The kinematics of an action do not only provide cues to intention as a side effect of an actor's performance, but they can also reflect an actor's intention to inform another agent (Sperber & Wilson, 2004). Thus, action kinematics can be actively used as a channel of information for joint action coordination and communication. Pezzulo et al. (2013) coined the term 'sensorimotor communication' for this active use of kinematics to inform. Sensorimotor communication is special compared to other forms of communication in that communication is superimposed on performed instrumental actions. Specifically, actors make instrumental actions informative by modulating kinematic parameters so that the actions become more predictable and less ambiguous (Pezzulo, et al., 2013).

Sensorimotor communication is often observed in joint actions, where co-actors make their actions more informative in order to effectively achieve interpersonal coordination. In a study by Sacheli, Tidoni, Pavone, Aglioti, and Candidi (2013), two participants were instructed to grasp a bottle synchronously with either a power or a precision grip. Crucially, only the 'leader' knew which part of the bottle to grasp, while the 'follower' relied on the leader's actions to select the appropriate grip. Compared to followers, leaders reduced the velocity of their movements, and modulated wrist height and grip aperture. This made their movements more informative, communicating task relevant information to their joint action partner. It is also important to note that sensorimotor communication is only produced when informative cues are required, which is evidenced by findings demonstrating that actors no longer produce kinematic cues when their co-actor already has access to the information necessary to complete the joint task (Pezzulo & Dindo, 2011; Leibfried, Grau-Moya, & Braun, 2015).

Developmental research on imitation shows that sensorimotor communication also occurs in teaching contexts, with teachers adjusting their actions to make them more informative for the learner. Brand, et al. (2002) found that when mothers demonstrated actions to their children, their movements were more punctuated and pronounced, with a larger range of motion. This was labelled 'motionese' and has been shown to facilitate imitation of observed actions. Infants are more likely to imitate actions containing motionese, compared to actions without motionese (Koterba & Iverson, 2009). It has been proposed that motionese enhances understanding of the goal structure of the action by guiding attention to important parts of an action sequence (Nagai and Rohlfing, 2009). These studies can be taken as evidence that sensorimotor communication is important for teaching through demonstration.

Using a virtual xylophone playing task, McEllin, Knoblich and Sebanz (2017) directly compared sensorimotor communication in joint action and in teaching through demonstration. Participants who had been trained to play melodies on a virtual xylophone produced different kinematic cues when trying to play the melodies in synchrony with a novice, compared to when they were demonstrating melodies to a novice. Specifically, modulations of movement height were used to support both teaching and coordination, modulations of the acceleration phase (ascent) of a movement were used to support spatial prediction in joint action coordination, and modulations of the deceleration phase (descent) of a movement were used to support temporal prediction in joint action coordination. This indicates that different kinematic cues are produced to support different informative intentions. In joint action kinematic cues are optimized to make the communicator's action more predictable, whereas in teaching kinematic cues are optimized to orient the learner's attention.

1.3. Reading informative intentions from actions

The finding that communicators modulate the kinematics of their actions differentially in joint action and teaching contexts raises the question of whether the recipients of the communication can identify communicators' informative intentions from observing their movements. We first aimed to investigate whether the recipients of sensorimotor communication can distinguish instrumental actions that have an informative intention superimposed from regular instrumental actions. Given that actors differentially modulate kinematics for different informative intentions (coordination vs teaching), we further aimed to investigate whether people can distinguish different informative intentions based on the kinematics of observed actions. Finally, we aimed to investigate which types of kinematic cues make people perceive that an actor has a coordination intention or a teaching intention.

We used a task in which participants were presented with a point light-display of a mallet movement that corresponded to an actor playing simple melodies on a virtual xylophone. Participants were asked to categorize the displays as reflecting individual action, demonstration for teaching, or part of a coordinated joint action. The observed movements were synthesized so that they corresponded to fundamental movement laws. Maximum height and velocity profile of the movements were systematically varied because they had been identified as the main cues communicators used in coordination and teaching contexts in our previous study (McEllin et al., 2017). Artificially modulating kinematic parameters rather than using natural kinematics gave us full experimental control over the kinematic cues in the display.

Assuming that intentions can be rendered 'visible' based on the kinematic signatures of actions, we made the following predictions. First, we predicted that participants would be able to discriminate between actions without informative intention (individual) and actions with an informative intention (teaching and joint), on the basis of kinematic cues. More specifically, given that actions that are intended to improve joint action coordination have been shown to be slower with a larger maximum height (McEllin et al., 2017; Vesper and Richardson, 2014), we predicted that participants would use exaggerated movement height and duration in order to categorize actions as joint, compared to individual actions. Based on the finding that teaching actions are characterized by more exaggerated movements (larger maximum height or larger range of motion) (Brand et al., 2002; McEllin et al., 2017) and by a slower pace (Dunst, Gorman & Hamby, 2012), we predicted that participants would also use exaggerated height and duration in order to categorize actions as teaching actions, compared to individual actions. Second, we predicted that participants would be able to discriminate between different informative intentions (joint action versus teaching) on the basis of different kinematic cues.

2. Experiment 1a: Discriminating individual and joint actions

The aim of Experiment 1 was to investigate whether people can discriminate actions that are performed with an informative intent in the context of joint action coordination from regular instrumental actions. Based on previous findings on kinematics produced during joint action (Vesper & Richardson, 2014), we predicted that participants would infer from exaggerated movement height and duration that the observed movement reflected an intent to inform a task partner about movement goals in order to achieve joint action coordination.

2.1. Method

2.1.1. Participants

Using an online participant database (Sona systems, www.sona-systems.com), we recruited 20 participants (13 males, 7 females), with a mean age of 25.4 (SD = 4.3). All participants gave informed consent and were given 1500 Forint (approximately 5 Euros) worth of vouchers for their participation. This study was approved by the United Ethical Review Committee for Research in Psychology (EPKEB). Informed consent was obtained from all participants, and they were fully briefed and debriefed before and after the experiment.

2.1.2. Apparatus and stimuli

Using data from previous experiments (McEllin et al., 2017), we synthesized point light displays of sequences of mallet movements reflecting the playing of melodies on a xylophone (see Fig. 1). The virtual xylophone had ten projected keys, each 5 cm wide and 24 cm long, separated by a 4 cm gap. Participants were required to learn simple action sequences, by moving the xylophone mallet from key to key in order to play a melody. To derive realistic parameters for our synthesized movements we computed, from the participants' individual performances without informative intent, the mean trajectories for movements of one, two and three key distances, for left and right movements. This resulted in six movement primitives, movements of one, two, and three keys to the left and right, which could be configured to synthesize, with appropriate resting times on the keys (100 ms delay between movements), action sequences reflecting the playing of melodies. While synthesizing the action sequences we used a pseudo-random sequence of the movement primitives with the added constraint that there had to be a direction change at least every two movements. This served to ensure that the mallet did not move off the xylophone displayed. Twenty unique six element action sequences were synthesized (see Appendix A).

We artificially modulated the kinematic parameters of movement height and duration, by increasing (exaggeration) or decreasing (suppression) these parameters by 25% relative to the movement height and duration of the mean trajectory (the transformation was applied to all samples of the trajectory). For each kinematic parameter this resulted in three levels of modulation, (suppressed, original, exaggerated). Every action sequence was subject to each level of both height modulation and duration modulation, resulting in nine different height-duration combinations, for every action sequence.

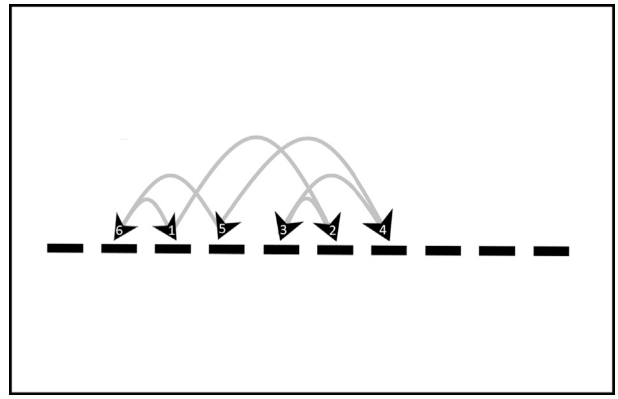

Fig. 1. Graphical depiction of the stimulus for one trial. Numbers represent the movement order for the action sequence, and arrows depict movement direction and end position of each movement.

The data were animated using MATLAB psychophysics toolbox. A lateral view of the xylophone was represented by ten blue rectangles (96 × 15 dimensions), arranged horizontally, and separated with a (72 pixel) gap. These dimensions were proportionate to the dimensions of the original xylophone. The xylophone mallet was represented by a green circle, which moved in accordance with the motion data. Please see Fig. 1 for a sketch of an example of one trial. The vertical and horizontal motion data were transformed into pixels and scaled down to fit within the dimensions of the animated xylophone. Data were presented at a rate of 60 HZ, with a frame of data being sampled and presented every 16 ms. Responses were recorded using a custom designed button box.

2.1.3. Procedure

Participants were told that they would complete a task in which they would have to decide whether a xylophone sequence played showed a participant playing alone (individual) or a participant playing together with another participant (joint). Participants were then provided with information about the individual condition and the joint condition from the previous set of experiments (McEllin et al., 2017). We described the individual condition as a task in which the observed participant played a xylophone sequence alone. We described the joint condition as a task in which the observed participant played the action sequence together in synchrony with an unknown participant who did not know the sequence.

Participants were told that half of the action sequences they were about to observe were from the individual condition, and half of the action sequences were from the joint condition. They were also told that participants played the exact same action sequences in both conditions. Participants were then familiarized with the current stimuli, being shown a frame depicting the xylophone and the mallet. They were told that for each trial the data from one of the two conditions would be reanimated, with the green circle representing the mallet head. We then had participants complete two practice trials, in order to further familiarize them with the kinematic displays and the decision they were asked to make.

In each trial of the main experiment participants were presented with a 500 ms fixation cross, followed by an animation of one of the action sequences. The duration of the action sequences ranged from

2460 to 4100 ms. Then participants were presented with a prompt screen which instructed them to indicate whether the action sequence they just watched had been played individually or as part of a coordinated joint action, by responding on a button box.

Each participant completed 180 trials judging 20 different action sequences for each height-duration modulation. The order of action sequences with different height-duration modulations was fully randomized. Whether participants categorized an action as individual or joint with a left or right button press was counterbalanced across participants.

2.1.4. Design

This experiment had a 3 × 3 within participant design, with the factors height modulation (suppressed, baseline, exaggerated) and duration modulation (suppressed, baseline, exaggerated). Our dependent variables were percentage of trials judged as joint (% Joint).

2.2. Results

A 3 × 3 within-participants ANOVA with the factors height modulation and duration modulation revealed a significant main effect of height modulation, $F(2,19) = 72.89$, $p < .001$, $\eta p2 = .79$ (see left panel of Fig. 2). The percentage of joint choices was significantly larger for exaggerated height than for original height and for suppressed height. Moreover, the percentage of joint choices for original height was significantly larger than for suppressed height. There was no significant main effect of duration, and no interaction between height and duration (all $p > .05$).

3. Experiment 1b: Discriminating individual and teaching actions

Experiment 1a demonstrated that participants use movement height as a cue to discriminate individual actions from actions performed with an informative intent in the context of joint action. This provides evidence that people can use kinematic cues to distinguish actions performed with an informative intent from actions performed without an informative intent. Another type of social interaction where actors modulate the kinematics of their movements to inform their co-actors is teaching. Here, the modulations serve to enhance attention to learning relevant information (Brand et al., 2002; McEllin et al., 2017). Experiment 1b asked whether people can discriminate between actions performed with the intention to teach and non-informative instrumental actions on the basis of kinematic cues. We predicted that participants would use movement height to discriminate actions performed with teaching intentions from regular instrumental actions, given the evidence for exaggeration of spatial parameters in teaching (Brand et al., 2002; McEllin et al., 2017). It is also possible that participants would use longer movement duration as an indication of a teaching intention, given that demonstrations for novice learners tend to be slower paced (Dunst, et al., 2012).

3.1. Method

3.1.1. Participants

Using an online participant database (Sona systems, www.sona-systems.com), we recruited 20 participants (12 males, 8 females), with a mean age of 23.7 ($SD = 3.5$). All participants gave informed consent and were given 1500 Forint (approximately 5 Euros) worth of vouchers for their participation.

3.1.2. Apparatus and stimuli

Like in Experiment 1a, participants were presented with point-light displays of artificially generated six-element xylophone sequences, in which height and duration were modulated.

3.1.3. Procedure and design

The procedure was like Experiment 1a, except that participants were asked to decide whether the animated action sequence showed an individual playing alone, or an individual teaching a learner. They were provided with information about the individual condition and the teaching condition from the previous set of experiments (McEllin et al., 2017). They were told that half of the action sequences were from an individual playing alone and half of the action sequences were from an individual teaching. The teaching condition was described as a task in which the observed participant was required to demonstrate the action sequence to an unknown student who was required to watch and reproduce what was observed. Like in Experiment 1a, whether participants categorized individual and teaching actions with a left or right button press was counterbalanced across participants. The design was the same as Experiment 1a but with the percentage of trials judged as teaching (% Teaching) as the dependent variable.

3.2. Results

The 3 × 3 within-subjects ANOVA with height modulation and duration modulation as factors (see right panel of Fig. 2) revealed a main effect of height, $F(2,19) = 19.92$, $p < .001$, $\eta p2 = .51$, with the percentage of teaching choices being significantly larger for

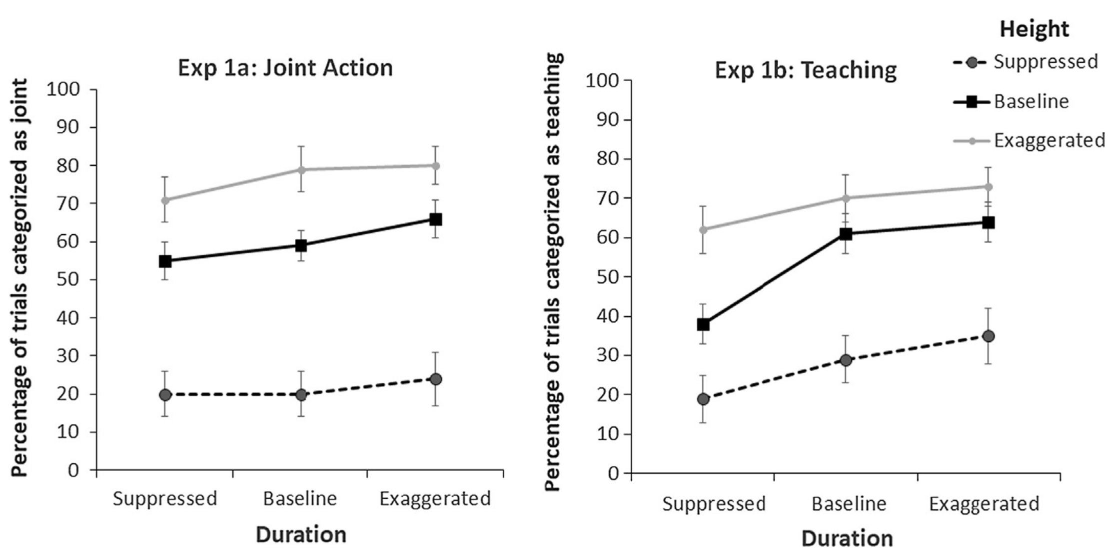

Fig. 2. Interaction between Height and Duration for Experiments 1a and 1b. Error bars represent ± 1 SEM.

exaggerated height than for original height and for suppressed height, and percentage of teaching choices for original height significantly larger than for suppressed height (all pairwise comparisons < .05). We also found a main effect of duration, $F(2,19) = 5.05$, $p = .011$, $\eta p2 = .21$, with the percentage of teaching choices being significantly larger for original duration than suppressed duration. Moreover, there was an interaction between height and duration, $F(4, 19) = 3.19$, $p = .018$, $\eta p2 = .144$. There was a lower percentage of teaching choices for movements with a lower movement height and shorter duration.

4. Discussion Experiment 1a and 1b

Taken together, the results from Experiment 1a and 1b demonstrate that people are sensitive to sensorimotor communication and can infer informative intentions using low level kinematic cues. Exaggerated movement height made participants more likely to judge actions as joint rather than individual (Experiment 1) and as reflecting the intention to teach (Experiment 2). Longer movement duration did not increase judgments of actions as joint rather than individual. This was unexpected given our earlier findings where participants acting in a joint coordination context moved more slowly than when acting alone (McEllin et al., 2017). It could be that movement height was a dominant cue in the present task, leaving open the question whether in the absence of height modulations people would use action duration to discriminate between actions performed with the intention to engage in coordinated joint action and individual actions. Movement duration had some effect on judgments of teaching intentions, with faster actions being judged unlikely to reflect a teaching intention.

Although participants were informed that half of the trials were individual trials and half of the trials were joint/teaching, participants seemed to be slightly biased towards categorizing trials as joint or teaching actions. One possibility could be that this reflects a more general bias towards perceiving social relations given minimal cues to interaction (Heider & Simmel, 1944). However, this bias cannot explain the observed results, as it does not imply a systematic effect of particular movement cues on judgments.

5. Experiment 2

Experiment 2 aimed to investigate whether people can discriminate between different types of informative intentions based on the different types of kinematic cues produced in these contexts. Specifically, we asked whether people can discriminate between actions performed with the intention to coordinate in a joint action and actions performed with the intention to teach. We did not make specific predictions for how participants would use height cues, given that an exaggerated movement height is observed in both joint action coordination and teaching (McEllin et al.2017; Vesper & Richardson, 2014) and that participants used maximum height to identify both the intention to teach and the intention to coordinate in Experiment 1a and 1b.

For duration, prior findings motivate two opposing predictions. On the one hand, longer durations may increase judgments of a teaching intention, given that demonstration often entails slower movements (Dunst, et al., 2012), and given the findings of Experiment 1b where longer duration served as a cue towards teaching. On the other hand, we found in an earlier study measuring the kinematics involved in producing the xylophone melodies (McEllin et al., 2017) that performing the actions in a joint context with a partner resulted in slower movements while demonstrating the actions to an observer did not reliably lead to a slowing down. This predicts that exaggerated movement duration would serve as a cue to joint action.

5.1. Method

5.1.1. Participants

Using an online participant database (Sona systems, www.sona-systems.com), we recruited 20 participants (10 males, 10 females), with a mean age of 24.7 ($SD = 4.7$). All participants gave informed consent and were given 1500 Forint (approximately 5 Euros) worth of vouchers for their participation.

5.1.2. Apparatus and stimuli

Like in Experiment 1a and 1b, participants were presented with point-light displays of artificially generated six-element xylophone sequences, in which height and duration were modulated.

5.1.3. Procedure

The procedure was like Experiment 1a and 1b, but participants were asked to discriminate between actions performed with the intention to coordinate in a joint action, and actions performed with the intention to teach. They were provided with information about the joint condition and the teaching condition before the experiment started and were told that half of what they observed were joint actions and half were teaching actions. Like in Experiment 1a and 1b, whether participants categorized joint and teaching actions with a left or right button press was counterbalanced across participants. The design was the same as in Experiment 1a, that is, the dependent variable was the percentage of trials judged as joint (% Joint).

5.2. Results

The 3 × 3 within-subjects ANOVA with height modulation and duration modulation as factors (see Fig. 3) showed a significant main effect of height, $F(2,19) = 3.49$, $p = .041$, $\eta p2 = .155$, with percentage of joint choice increasing as a function of height modulation. There was no significant main effect of duration and no significant interaction between height and duration.

5.3. Discussion

The results of Experiment 2 demonstrate that people can use kinematic cues in order to discriminate between actions performed with different informative intentions. Unexpectedly however, we found that participants used movement height, but not duration to discriminate between actions performed with the intention to coordinate in a joint action and actions performed with a teaching intention. These findings imply that exaggerated movement height is more likely interpreted as an attempt to achieve interpersonal coordination during joint action than to serve teaching purposes. Note, however, that the effect of height

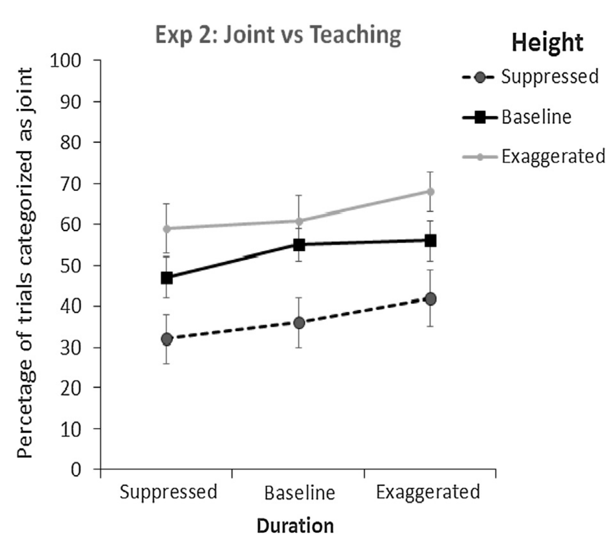

Fig. 3. Interaction between Height and Duration for Experiment 2. Error bars represent ± 1 SEM.

in Experiment 2 is considerably smaller than in the previous two experiments, implying higher uncertainty in discriminating the underlying intentions of the movement.

6. Experiment 3a

The results of the previous experiments could be taken to suggest that participants hardly use timing cues to discriminate between actions performed with different intentions. However, the role of action duration may have been underestimated because the height modulation may have dominated participants' judgments. Furthermore, McEllin et al. (2017) demonstrated that participants differentially modulate their movement velocity during different movement phases depending on what they intended to inform a joint action partner about. Specifically, we found that ascent velocity of the mallet (the movement speed from the xylophone up to the maximum height) was modulated when participants were informing a partner about spatial locations. Descent velocity (the movement speed from the maximum height to the target key) was modulated when participants were informing a partner about the timing of their movements. Thus, subtle changes in velocity parameters may also be used to discriminate between different informative intentions.

In the ensuing four experiments (3a–4b), we aimed to further investigate whether timing cues can be used to discriminate between informative intentions and instrumental intentions, and between different informative intentions. In Experiment 3a, we asked whether in the absence of height modulations people use action duration to discriminate between actions performed with the intention to engage in coordinated joint action, and individual instrumental actions. We manipulated the duration of the up stroke and down stroke of the mallet resulting in different ratios of ascent and descent velocity. Because ascent velocity has been shown to support spatial prediction, we hypothesized that exaggerated duration of the upstroke (a slow-down in ascent velocity relative to descent velocity) would be used to identify the intention to engage in a coordinated joint action. Because descent velocity has been shown to support temporal prediction, we predicted that exaggerated duration of the down stroke (a slow-down in descent velocity relative to ascent velocity) would also be used in order to identify an intention to engage in coordinated joint action.

6.1. Method

6.1.1. Participants

Using an online participant database (Sona systems, www.sona-systems.com), we recruited 20 participants (14 males, 6 females), with a mean age of 24.6 (*SD* = 5.6). All participants gave informed consent and were given 1500 Forint (approximately 5 Euros) worth of vouchers for their participation.

6.1.2. Apparatus and stimuli

Apparatus and Stimuli were the same as in Experiment 1a and 1b, except for how the kinematic parameters were modulated. We used the same action sequences as in Experiment 1a and 1b. Ascent duration was increased by a factor of 30%, 60% and 90% (see Appendix B). Descent duration was kept constant. We then combined the ascent and descent durations and normalized them so that the duration matched the original overall duration, thus increasing the ascent duration relative to the descent duration. We did the same for the duration of the descent phase of the movements, increasing the descent duration by 30%, 60% and 90% (see Appendix B), and then normalizing the overall duration to increase the proportion of the descent duration, relative to the ascent duration. We also had an individual baseline in which we never modulated the ascent or descent duration. We created each of these seven ascent-to-descent ratios for each of the twenty action sequences that we used in the previous experiments. Again, these action sequences were animated as point-light displays.

To dissociate effects of overall duration from effects of specific ascent and descent modulations, we also manipulated overall action duration. Every action sequence had a randomly modulated duration, which ranged from the original duration (3280 ms) to double the original duration (6560 ms).

6.1.3. Procedure and design

The procedure was the same as experiment 1a, but with 140 trials instead of 180 trials. The baseline duration and each ascent and descent duration modulation of the twenty action sequences were presented in a random order. Like in Experiment 1a, participants were provided with information about the individual condition and the joint condition, before being instructed to decide whether each of the observed action sequences was an individual action or performed with the intention to coordinate in a joint action.

We added the baseline as a level of both the ascent and descent factors, in order to compare each of the modulated actions to the unmodulated actions. We also created a factor of speed, by performing a median split based on movement duration in order to split the stimuli into slow and fast actions. This resulted in a design with a 2 × 4 within-participant comparison for ascent exaggeration (baseline, 30%, 60% and 90%) and speed (slow, fast) and a 2 × 4 within-participant comparison for descent exaggeration (baseline, 30%, 60%, 90%) and speed (slow, fast). Percentage of trials judged as joint (% joint) was the dependent variable.

6.2. Results

6.2.1. Ascent exaggeration

We carried out a 2 × 4 ANOVA for ascent exaggeration (left panel of Fig. 4) with the factors speed (fast, slow) and exaggeration (baseline, 30%, 60%, 90%). This analysis yielded a main effect of speed, $F(1,19) = 128.5$, $p < .001$, $\eta p2 = .87$ (Fig. 6), and a main effect of exaggeration, $F(3,19) = 4.7$, $p = .005$, $\eta p2 = .2$ (Fig. 7). The percentage of trials judged joint was higher for slow movements than for fast movements, and percentage of trials judged as joint increased as a function of ascent exaggeration. However, there was no interaction between speed and exaggeration, $F(3,19) = 2.1$, $p = .11$, $\eta p2 = .1$.

6.2.2. Descent exaggeration

We carried out a 2 × 4 ANOVA for descent exaggeration (left panel of Fig. 5) with the factors of speed (slow, fast) and exaggeration (baseline, 30%, 60%, 90%). The ANOVA revealed a main effect of speed, $F(1,19) = 120.09$, $p < .001$, $\eta p2 = .86$ (Fig. 6), and exaggeration, $F(3,19) = 2.98$, $p = .039$, $\eta p2 = .136$. The percentage of trials judged joint was higher for slow movements than for fast movements, and the percentage of trials judged as joint increased as a function of descent exaggeration (Fig. 7). There was no interaction between speed and exaggeration, $F(3,19) = .25$, $p = .86$, $\eta p2 = .01$.

7. Experiment 3b

This experiment aimed to investigate whether people use information from the velocity profile of an observed action in order to infer teaching intentions. In particular, we investigated whether participants can decide whether an observed action was performed alone or whether it was performed with the intention to teach a learner, on the basis of the speed and ascent and descent ratio of that action. Because the results from Experiment 1b indicate that action duration serves as a cue to teaching, we predicted that participants would mostly rely on overall movement speed in order to discriminate between individual movements and movements performed with the intention to teach.

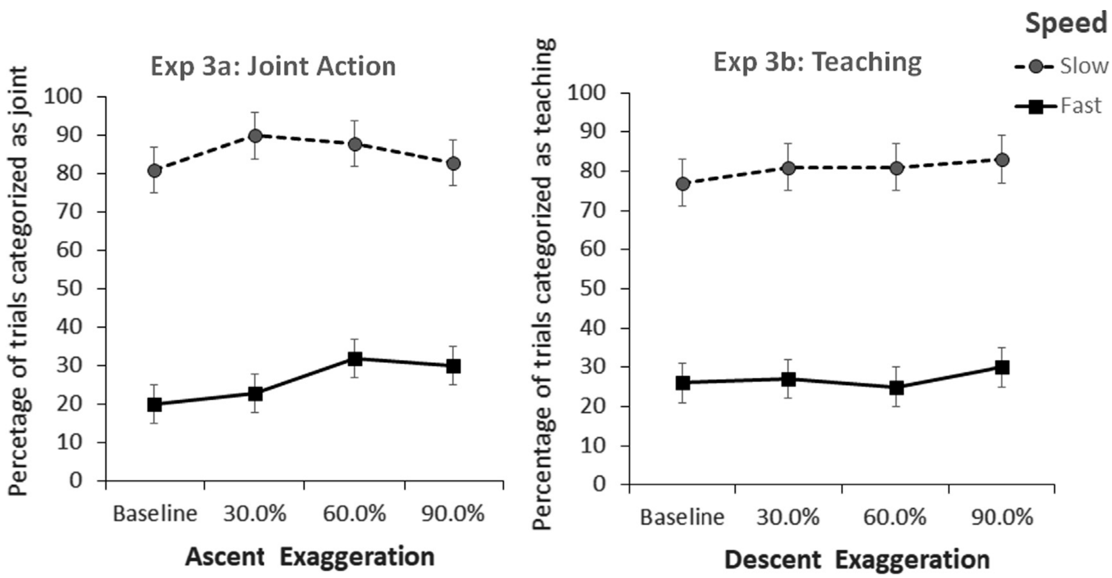

Fig. 4. Interaction between Ascent Exaggeration and Speed for Experiment 3a and 3b. Error bars represent ± SEM.

7.1. Method

7.1.1. Participants

Using an online participant database (Sona systems, www.sona-systems.com), we recruited 20 participants (11 males, 9 females), with a mean age of 22.8 (*SD* = 2.7). All participants gave informed consent and received 1500 Forint (approximately 5 Euros) worth of vouchers for their participation.

7.1.2. Apparatus and stimuli

Like in Experiment 3a, participants were presented with point-light displays of artificially generated six-element xylophone sequences, in which ascent and descent duration and overall duration were modulated.

7.1.3. Procedure

This was the same as in Experiment 3a, except that participants were familiarized with the individual condition and the teaching condition, and then instructed to decide whether the observed action was an individual action or a teaching action.

7.1.4. Design

This was the same as in Experiment 3a, except that percentage of teaching choices was the dependent variable.

7.2. Results

7.2.1. Ascent exaggeration

Like in Experiment 3a, we carried out a 2 × 4 ANOVA for ascent exaggeration (right panel of Fig. 4) with the factors of speed and exaggeration. It revealed a main effect of speed, $F(1,19) = 65.85$, $p < .001$, $\eta p2 = .78$ (Fig. 6), but no main effect of exaggeration, $F(3,19) = .82$, $p = .49$ (Fig. 7), $\eta p2 = .04$. The percentage of trials judged as teaching was higher for slow movements compared to fast movements. There was no interaction between speed and exaggeration, $F(3,19) = .33$, $p = .81$, $\eta p2 = .02$.

7.2.2. Descent exaggeration

For descent exaggeration, we carried out a 2 × 4 ANOVA with speed and exaggeration as within-participant factors (right panel of Fig. 5). It revealed a main effect of speed, $F(1,19) = 60.23$, $p < .001$,

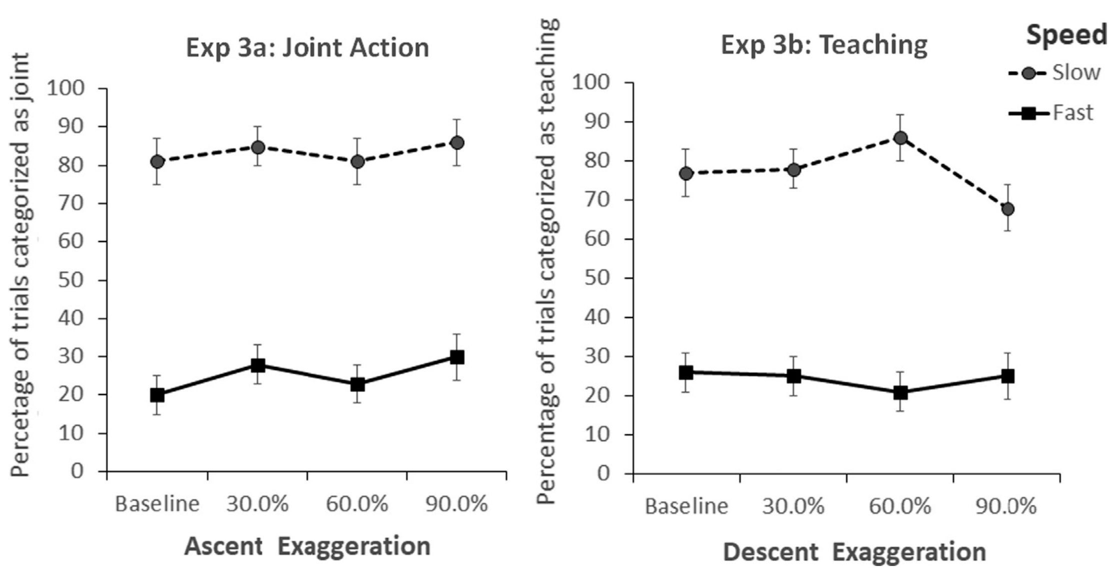

Fig. 5. Interaction between Descent Exaggeration and Speed for Experiment 3a and 3b. Error bars represent ± 1 SEM.

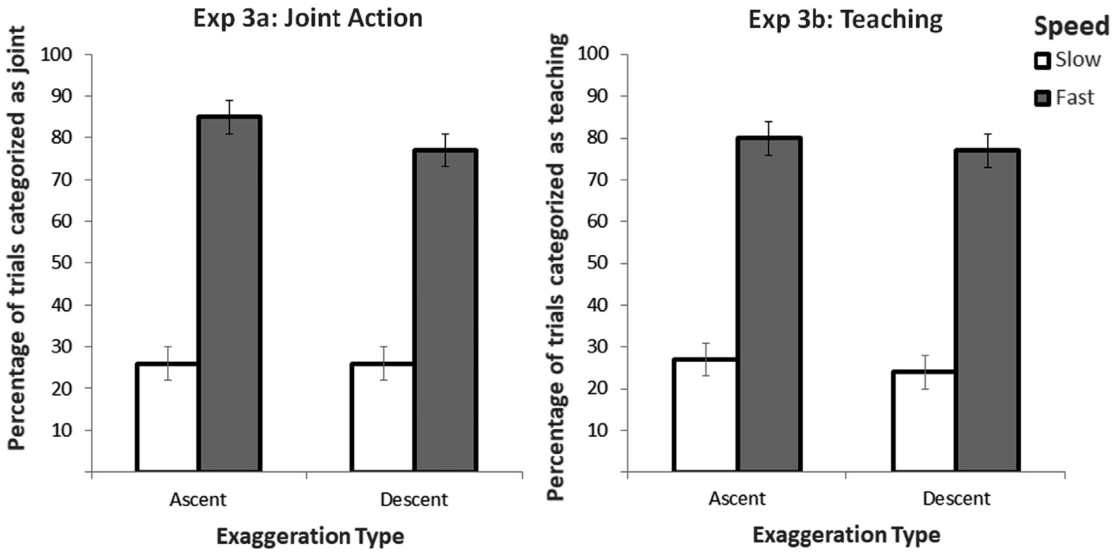

Fig. 6. Main effect of Speed for both Ascent and Descent modulations. Error bars represent ± 1 SEM.

$\eta p2 = .76$ (Fig. 6), but no effect of exaggeration, $F(3,19) = 1.65$, $p = .188$, $\eta p2 = .08$ (Fig. 7). The percentage of trials judged as teaching was higher for slow movements compared to fast movements. However, there was a significant interaction between speed and exaggeration, $F(3,19) = 3.17$, $p = .031$, $\eta p2 = .14$, with the percentage of teaching choices being lower for 90% descent exaggeration during slow speed.

8. Discussion Experiment 3a and 3b

The results from Experiment 3a and 3b show that people can use temporal cues to detect an actor's informative intentions (both coordination and teaching) and suggest that in our first two experiments, the duration modulation had not been salient enough for the participants.

Interestingly, we also found that participants used the relative length of the ascent phase and descent phase of the movements in order to discriminate between individual actions and joint actions (Experiment 3a), while they did not use this information to discriminate between individual actions and teaching actions (Experiment 3b). This may provide some indication that people are more sensitive to temporal cues in actions performed with an intention to coordinate, compared to actions performed with an intention to teach. However, an experiment comparing these two types of intention would be needed in order to provide conclusive evidence.

9. Experiment 4a

Experiment 4a aimed to investigate whether people can use temporal cues in order to discriminate between actions performed with different informative intentions (joint action and teaching). Considering that people have been shown to modulate the ratio of ascent to descent velocity in order to enhance spatial and temporal prediction in joint action (McEllin et al., 2018; Sacheli, et al., 2013), we predicted that participants would categorize movements with larger ratios of ascent to descent velocity as joint actions, as they understand the role this informative modulation plays in spatial and temporal prediction.

9.1. Method

9.1.1. Participants

Using an online participant database (Sona systems, www.sona-

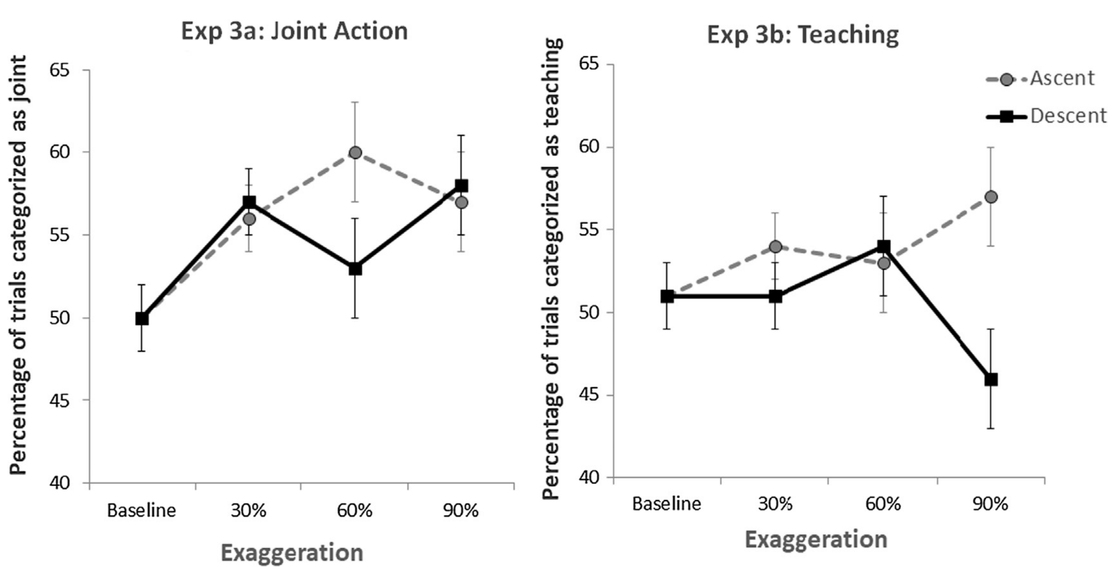

Fig. 7. Main effects of Exaggeration for both Ascent and Descent Modulations. Error bars represent ± 1 SEM.

systems.com), we recruited 20 participants (13 males, 7 females), with a mean age of 23.4 (SD = 5.2). All participants gave informed consent and received 1500 Forint (approximately 5 Euros) worth of vouchers for their participation.

9.1.2. Apparatus and stimuli

Like in Experiment 3a and 3b, participants were presented with point-light displays of artificially generated six-element xylophone sequences, in which ascent and descent duration and overall duration were modulated.

9.1.3. Procedure

This was the same as experiment 3a and 3b, except that participants were provided with information about the teaching condition and the joint condition, and then instructed to categorize the observed actions as being performed with an intention to coordinate or as being performed with the intention to teach.

9.1.4. Design

This was the same as Experiment 3a; percentage of trials judged as joint was the dependent variable.

9.2. Results

9.2.1. Ascent exaggeration

We carried out a 2 × 4 within-participant ANOVA and found neither a significant main effect of speed, $F(1,19) = .18$, $p = .67$, $\eta p2 = .01$ (Fig. 8), nor a significant main effect of exaggeration, $F(3,19) = .77$, $p = .52$, $\eta p2 = .04$. The interaction between exaggeration and speed was also not significant, $F(3,19) = 1.5$, $p = .22$, $\eta p2 = .08$.

9.2.2. Descent exaggeration

A 2 × 4 within-participant ANOVA revealed a significant main effect of exaggeration, $F(3,19) = 3.73$, $p = .016$, $\eta p2 = .16$ (Fig. 8), but no main effect of speed, $F(1,19) = .34$, $p = .56$, $\eta p2 = .02$. Percentage of trials judged as joint increased as a function of descent exaggeration. There was no significant interaction between exaggeration and speed, $F(3,19) = 1.61$, $p = .2$, $\eta p2 = .08$.

10. Experiment 4b

Experiment 4a provided first evidence that participants can use the ratio between the ascent and descent duration to discriminate between actions performed with an intention to coordinate and actions performed with a teaching intention. Experiment 4b served to replicate this finding and to determine whether discrimination becomes more reliable when overall speed does not vary.

10.1. Method

10.1.1. Participants

Using an online participant database (Sona systems, www.sona-systems.com), we recruited 20 participants (9 males, 11 females), with a mean age of 21.6 (SD = 1.8). All participants gave informed consent and received 1500 Forint (approximately 5 Euros) worth of vouchers for their participation.

10.1.2. Apparatus and stimuli

Like in Experiment 4a, participants were presented with point-light displays of artificially generated six-element xylophone sequences, in which ascent and descent duration modulated. However overall duration was not modulated.

10.1.3. Procedure

This was the same as Experiment 4a.

10.1.4. Design

This was the same as Experiment 4a.

10.2. Results

10.2.1. Ascent exaggeration

We carried out a one-way ANOVA with the factor exaggeration.

We found a significant main effect of exaggeration, $F(3,19) = 7.01$, $p < .001$, $\eta p2 = .27$. The percentage of trials judged as joint increased as a function of ascent exaggeration (Fig. 8).

10.2.2. Descent exaggeration

We carried out a one-way ANOVA that revealed a significant main effect of exaggeration, $F(3,19) = 6.69$, $p = .001$, $\eta p2 = .26$. The

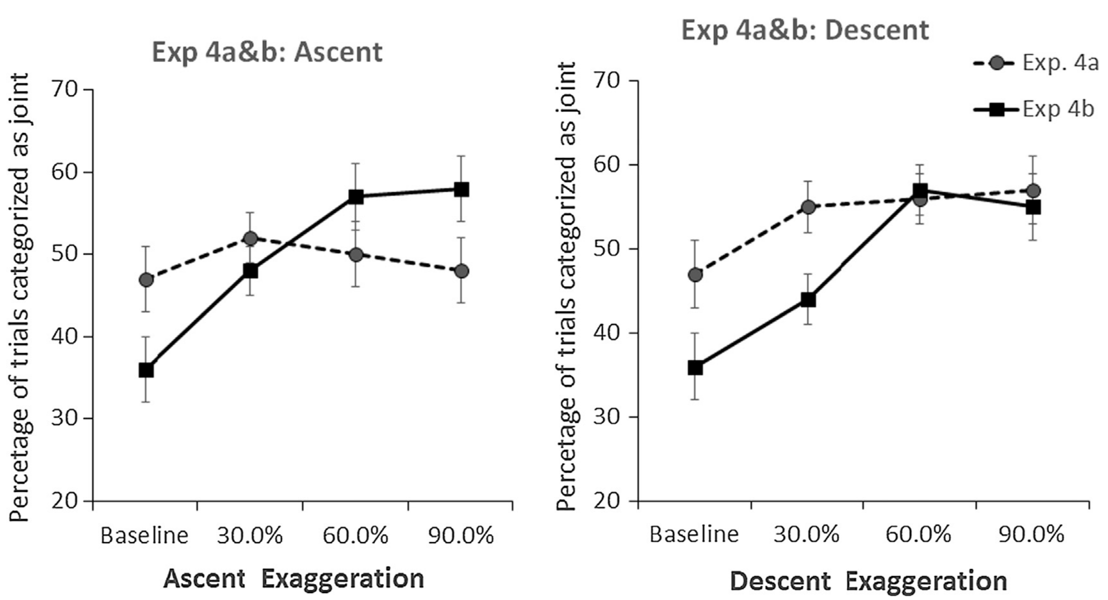

Fig. 8. Interaction between Exaggeration (Ascent: Left, Descent: Right) and Experiment (4a: Speed factor present. 4b: Speed factor absent). Error bars represent ± 1 SEM.

percentage of trials judged as joint increased as a function of descent exaggeration (Fig. 8).

11. Discussion Experiment 4a and 4b

Experiment 4a did not provide any evidence that participants use movement speed in order to discriminate between actions performed with the intention to coordinate and actions performed with the intention to teach. However, Experiment 4b provided evidence that people use exaggerated ascent and descent durations in order to discriminate between joint actions and teaching actions. These findings demonstrate that people can use specific information about the velocity profile in order to discriminate between different types of informative intentions.

Interestingly, we found that participants only used ascent exaggeration in the absence of any overall speed cues, whereas participants can use descent exaggeration regardless of whether speed cues are present or not. This could indicate that with regards to coordinated joint actions, people have stronger expectations about the descent phase of a movement, compared to the ascent phase. Considering that ascent velocity is typically used to inform a task partner about spatial movement parameters and descent velocity is used to inform a task partner about movement timing, it may be the case that the temporal requirements of joint action were more salient for the participants than the spatial requirements.

12. General discussion

We aimed to investigate whether people can discriminate between actions performed with informative intentions and purely instrumental individual actions and whether they can discriminate between actions with different informative purposes such as the intention to perform a coordinated joint action coordination and the intention to teach through demonstration.

Regarding the first aim, previous research has demonstrated that people can detect the instrumental and social intentions of an actor on the basis of kinematic signatures (Cavallo et al., 2016; Manera et al., 2011). We extend this research by demonstrating that people can also detect an actor's informative intentions as expressed through sensorimotor communication. Our findings demonstrate that people use different movement cues in order to distinguish instrumental actions performed with an informative intention from individual instrumental actions without informative intention. Actions that systematically deviate from the easiest way of individually performing an effective instrumental action are understood as fulfilling some informative purpose (Pezzulo et al., 2013; Pezzulo et al., 2018). Our findings challenge theories of social cognition suggesting that movement cues alone are not sufficient for detecting intentions beyond motor intentions (Jacob & Jeannerod, 2005). Minimally, the findings demonstrate that there are some instances where informative intentions are derived from the kinematics of an observed movement.

An important goal for future research is to quantify the accuracy with which informative intentions can be identified. In the present study, we exaggerated natural movement kinematics to be able to specify and dissociate the contribution of different movement parameters. This approach allowed us to measure participants' tendency to attribute particular informative intentions as a function of exaggeration of particular movement cues, while their judgments were not right or wrong. Exposing participants to actual kinematics from teaching and joint action coordination contexts can contribute to the understanding of the efficiency of sensorimotor communication in these contexts.

Theories of communication assert that in order for communication to succeed, one needs to explicitly recognize an interaction partner's 'communicative intention' (Sperber & Wilson, 2004). Our results could be taken to suggest that sensorimotor communication can be sufficient for making communicative intentions explicit. However, it is possible that participants merely derived informative intentions, which specify the kind of information to be transmitted rather than making the actor's communicative intention explicit. As understanding informative intentions seems to be sufficient in many joint action and teaching contexts, it may be the case that only very large deviations from optimal performance elicit explicit attributions of communicative intentions (such as when the observer sees the other waving the mallet to grab attention). Further research could aim to identify kinematic parameters that discriminate between actions produced in a context requiring the detection of communicative intentions and actions produced in a context requiring only the identification of informative intentions.

The second aim of the present study was to investigate whether people can discriminate between actions performed with the intention to inform a co-actor in a joint action, and actions performed with the intention to inform a student in a teaching context. We found that participants reliably used movement height in order to discriminate between actions performed with an intention to coordinate in a joint action and actions performed with the intention to teach. This is somewhat surprising given that modulations of the spatial trajectory serve not only to enhance spatial and temporal prediction in joint action, but also to highlight the structure of an action in teaching (McEllin et al., 2017). It could be that people expect modulations of the spatial trajectory of one's movement to be more crucial for spatial and temporal prediction than for highlighting the structure of an action because they have more experience with exaggerating the spatial trajectory of their movements when being engaged in coordinated joint actions compared to when teaching through demonstration.

Furthermore, we found that people reliably categorize movements with elongated ascent or descent ratios as joint actions rather than teaching actions. This suggests that participants perceive an elongated ascent or descent phase of a movement as an informative action modulation when trying to coordinate in a joint action, but not when teaching. Indeed, there is evidence that people elongate the ascent phase of their movement in order to make themselves spatially predictable, that they and elongate the descent phase of their movement in order to make themselves temporally predictable (McEllin et al., 2017; Sacheli et al., 2013). Considering this, it is likely that participants perceived the longer ascent phase as an action modulation which provides actors with more time to make spatial predictions about their co-actor's targets. Likewise, longer relative descent phases could be perceived as modulations which provide actors with more time to make temporal predictions about their co-actor's movements.

It is possible that the ability to detect informative intentions arises through participants' simulation of performing the observed actions. Indeed, Becchio, Sartori & Castiello (2010) proposed that the ability to understand an actor's intentions through observing their early kinematics relies on motor simulation, mapping the kinematics onto their own motor repertoire in order to predict how the action will unfold. This is supported by evidence showing stronger activation of mirroring networks when observing cooperative and competitive actions, compared to individual actions (Becchio, Cavallo, et al., 2012; Becchio, Manera, et al., 2012). The same mechanism could be employed in order to detect sensorimotor communication. Given that people deviate from optimality in order to send informative signals in social contexts (Pezzulo et al., 2013), they may also understand that an action is informative when it systematically deviates from the easiest way of performing the action in ways that they themselves would use to signal an informative intent to an observer.

Given that we used a musical task, one may wonder whether the observed actions were actually in our participants' motor repertoire. Although participants may not have direct experience with playing a xylophone, it is likely that they have experience with instrumentally similar actions (e.g., moving to a sequence of locations in a particular order), so that they could understand the actions of the xylophone players using motor simulation. Moreover, we found in our earlier study using the xylophone task that non-musicians reliably produced informative movement cues (McEllin et al., 2017), suggesting that the observed actions were indeed within our observers' repertoire. One discrepancy between the kinematics produced in performing the task and the kinematics used to infer informative intentions is that participants relied less on timing cues when judging observed actions. It is possible that limited expertise with the observed actions made it easier to detect deviations in movement height than subtle deviations in timing.

A way to further test the role of motor simulation would be to investigate performance on the present task in a population of people who lack motor experience of producing informative cues in social interactions. A recent study demonstrated that those with autistic spectrum conditions (ASC) are less likely to produce sensorimotor communication in coordination contexts (Curioni, Minio-Paluello, Sacheli, Candidi & Aglioti, 2017). Thus, for the current task one could predict that individuals with ASC would not be able to reliably discriminate between actions produced with an informative intention and non-informative actions on the basis of kinematic cues, due to the lack of experience in producing sensorimotor communication. In a similar vein, experts should be more sensitive to detect kinematic cues signalling different informative intentions in their domains of expertise. If expectations of sensorimotor communication are driven by experience in producing kinematic cues, increasing expertise should increase sensitivity to these cues.

A further issue that would be interesting to address is to determine when kinematic modulations stop to act as informative cues and become noise. In Experiment 3b, we unexpectedly found that participants were less likely to categorize slow movements with very long descent phases as teaching actions. One explanation for this finding could be that participants perceive very long descent phases as reflecting hesitance or uncertainty, rather than perceiving them as being informative in a teaching context. Alternatively, it could be that very large kinematic exaggerations are interpreted as mistakes or bad performance rather than signalling an informative intent in a teaching context. This may have occurred specifically for teaching actions but not joint actions because participants expect modulating ascent and descent ratio to be useful to achieve interpersonal coordination, but not for teaching (as evidenced by Experiment 3a, 4a and 4b). More generally, this finding suggests that although deviating from individual efficiency by exaggerating kinematic parameters allows one to provide useful and informative cues in social interactions such as coordination and teaching, there may be a threshold at which kinematic exaggeration is no longer informative, and actually makes the performed action more ambiguous and harder to predict. Further research should investigate at what point deviation from optimality actually begins to violate the process of mapping observed actions onto our own motor system rather than facilitating it.

Acknowledgments

This research was supported by the European Research Council under the European Union's Seventh Framework Program (FP7/2007-2013)/ERC (European Research Council) grant agreement no. 609819, SOMICS, and by ERC grant agreement no. 616072, JAXPERTISE (https://erc.europa.eu/).

Appendix A

(1)	−1	−2	2	3	−3	1
(2)	−1	3	1	−2	−3	2
(3)	2	−1	−2	3	1	−3
(4)	−1	2	−3	−2	1	3
(5)	−3	3	−2	−1	1	2
(6)	1	−3	−1	2	−2	3
(7)	2	−2	−3	1	3	−1
(8)	−2	2	−3	−1	1	3
(9)	−2	2	−1	3	1	−3
(10)	2	1	−3	3	−1	−2
(11)	1	−3	3	2	−2	−1
(12)	2	−3	−2	1	3	−1
(13)	−3	1	3	−2	−1	2
(14)	3	1	−2	−3	2	−1
(15)	2	1	−1	−2	3	−3
(16)	3	−2	2	−1	−3	1
(17)	−1	3	1	−3	−2	2
(18)	3	1	−1	−3	2	−2
19)	−3	1	3	−1	2	−2
20)	−2	−1	3	−3	1	2

List of twenty randomly generated action sequences. Positive numbers represent rightward movements, and negative numbers represent leftward movements.

Appendix B

Tables showing the kinematic parameters of our stimuli, for each of our experiments.

Height modulation			
	1 Key	2 Keys	3 Keys
Suppressed	3.04	3.89	4.58
Baseline	4.02	5.18	6.1

Exaggerated	5.06	6.48	7.63

Table showing height values (cm) for each key distance, for Experiment 1 & 2.

Duration modulation			
	1 Key	2 Keys	3 Keys
Suppressed	342	399	458
Baseline	456	532	610
Exaggerated	570	665	763

Table showing duration values (ms) for each key distance, for Experiment 1 & 2.

Ascent modulation				
		1 Key	2 Keys	3 Keys
Baseline	Ascent	220	257	295
	Descent	236	275	315
30%	Ascent	252	295	338
	Descent	204	237	272
60%	Ascent	284	332	381
	Descent	172	200	229
90%	Ascent	316	369	424
	Descent	140	163	186

Table showing ascent duration values (ms) for Experiment 3 & 4.

Descent modulation				
		1 Key	2 Keys	3 Keys
Baseline	Ascent	220	257	295
	Descent	236	275	315
30%	Ascent	183	214	246
	Descent	273	318	364
60%	Ascent	146	171	197
	Descent	310	361	413
90%	Ascent	110	129	148
	Descent	346	403	462

Table showing descent duration values (ms) for Experiment 3 & 4.

References

Aglioti, S. M., Cesari, P., Romani, M., & Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nature Neuroscience, 11*(9), 1109–1116.

Ansuini, C., Cavallo, A., Bertone, C., & Becchio, C. (2015). Intentions in the brain: The unveiling of Mister Hyde. *The Neuroscientist, 21*(2), 126–135.

Becchio, C., Cavallo, A., Begliomini, C., Sartori, L., Feltrin, G., & Castiello, U. (2012). Social grasping: From mirroring to mentalizing. *Neuroimage, 61*(1), 240–248.

Becchio, C., Manera, V., Sartori, L., Cavallo, A., & Castiello, U. (2012). Grasping intentions: From thought experiments to empirical evidence. *Frontiers in Human Neuroscience, 6*, 117.

Becchio, C., Sartori, L., Bulgheroni, M., & Castiello, U. (2008). Both your intention and mine are reflected in the kinematics of my reach-to-grasp movement. *Cognition, 106*(2), 894–912.

Becchio, C., Sartori, L., & Castiello, U. (2010). Toward you: The social side of actions. *Current Directions in Psychological Science, 19*(3), 183–188.

Brand, R. J., Baldwin, D. A., & Ashburn, L. A. (2002). Evidence for 'motionese': Modifications in mothers' infant-directed action. *Developmental Science, 5*(1), 72–83.

Cavallo, A., Koul, A., Ansuini, C., Capozzi, F., & Becchio, C. (2016). Decoding intentions from movement kinematics. *Scientific Reports, 6*.

Curioni, A., Minio-Paluello, I., Sacheli, L. M., Candidi, M., & Aglioti, S. M. (2017). Autistic traits affect interpersonal motor coordination by modulating strategic use of role-based behavior. *Molecular Autism, 8*(1), 23.

Dunst, C. J., Gorman, E., & Hamby, D. W. (2012). Child-directed motionese with infants and toddlers with and without hearing impairments. *Center for Early Literacy Learning, 5*(8).

Grèzes, J., Frith, C., & Passingham, R. E. (2004). Brain mechanisms for inferring deceit in the actions of others. *Journal of Neuroscience, 24*(24), 5500–5505.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology, 57*(2), 243–259.

Jacob, P., & Jeannerod, M. (2005). The motor theory of social cognition: A critique. *Trends in Cognitive Sciences, 9*(1), 21–25.

Koterba, E. A., & Iverson, J. M. (2009). Investigating motionese: The effect of infant-directed action on infants' attention and object exploration. *Infant Behavior and Development, 32*(4), 437–444.

Leibfried, F., Grau-Moya, J., & Braun, D. A. (2015). Signaling equilibria in sensorimotor interactions. *Cognition, 141*, 73–86 ISO 690.

Manera, V., Becchio, C., Cavallo, A., Sartori, L., & Castiello, U. (2011). Cooperation or competition? Discriminating between social intentions by observing prehensile movements. *Experimental Brain Research, 211*(3), 547–556.

McEllin, L., Knoblich, G., & Sebanz, N. (2017). Distinct Kinematic Markers of Demonstration and Joint Action Coordination? Evidence From Virtual Xylophone Playing. *Journal of Experimental Psychology. Human Perception and Performance*.

McEllin, L., Sebanz, N., & Knoblich, G. (2018). Identifying informative intentions from kinematics. https://doi.org/10.7910/DVN/H8OT2U, Harvard Dataverse, V1, UNF:6:cy6tFoF2/sPpcYosOQPS3Q = =.

Nagai, Y., & Rohlfing, K. J. (2009). Computational analysis of motionese toward scaffolding robot action learning. *IEEE Transactions on Autonomous Mental Development, 1*(1), 44–54.

Pezzulo, G., & Dindo, H. (2011). What should I do next? Using shared representations to solve interaction problems. *Experimental Brain Research, 211*(3–4), 613–630.

Pezzulo, G., Donnarumma, F., & Dindo, H. (2013). Human sensorimotor communication: A theory of signaling in online social interactions. *PLoS One, 8*(11), e79876.

Pezzulo, G., Donnarumma, F., Dindo, H., D'Ausilio, A., Konvalinka, I., & Castelfranchi, C. (2018). The body talks: Sensorimotor communication and its brain and kinematic

signatures. *Physics of Life Reviews.*.
Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. *Nature Reviews Neuroscience, 11*(4), 264.
Runeson, S., & Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General, 112*(4), 585.
Sacheli, L. M., Tidoni, E., Pavone, E. F., Aglioti, S. M., & Candidi, M. (2013). Kinematics fingerprints of leader and follower role-taking during cooperative joint actions. *Experimental Brain Research, 226*(4), 473–486.
Sebanz, N., & Shiffrar, M. (2009). Detecting deception in a bluffing body: The role of expertise. *Psychonomic Bulletin & Review, 16*(1), 170–175.
Sperber, D., & Wilson, D. (2004). Relevance theory. *Handbook of pragmatics* (pp. 607–632). Oxford: Blackwell.
Vesper, C., & Richardson, M. J. (2014). Strategic communication and behavioral coupling in asymmetric joint action. *Experimental Brain Research, 232*(9), 2945–2956.
Vesper, C., Schmitz, L., Safra, L., Sebanz, N., & Knoblich, G. (2016). The role of shared visual information for joint action coordination. *Cognition, 153*, 118–123.