

Categorizing coordination from the perception of joint actions

Joseph M. Burling¹ · Hongjing Lu^{1,2}

Published online: 29 November 2017
© The Psychonomic Society, Inc. 2017

Abstract The ability to perceive others' actions and coordinate our own body movements accordingly is essential for humans to interact with the social world. However, it is still unclear how the visual system achieves the remarkable feat of identifying temporally coordinated joint actions between individuals. Specifically, do humans rely on certain visual features of coordinated movements to facilitate the detection of meaningful interactivity? To address this question, participants viewed short video sequences of two actors performing different joint actions, such as handshakes, high fives, etc. Temporal misalignments were introduced to shift one actor's movements forward or backward in time relative to the partner actor. Participants rated the degree of interactivity for the temporally shifted joint actions. The impact of temporal offsets on human interactivity ratings varied for different types of joint actions. Based on human rating distributions, we used a probabilistic cluster model to infer latent categories, each revealing shared characteristics of coordinated movements among sets of joint actions. Further analysis on the clustered structure suggested that global motion synchrony, spatial proximity between actors, and highly salient moments of interpersonal coordination are critical features that impact judgments of interactivity.

Keywords Biological motion · Joint action · Perception and action · Perceptual categorization and identification

✉ Hongjing Lu
hongjing@ucla.edu

¹ Department of Psychology, University of California, Los Angeles, CA 90095, USA

² Department of Statistics, University of California, Los Angeles, CA 90095, USA

Introduction

Our ability to perceive the actions of others and coordinate them with our own body movements is essential for interacting with the social world (Sebanz & Knoblich, 2009; Sebanz, Bekkering, & Knoblich, 2006). Human social interaction makes it possible to adapt and update our own actions in response to changes in other individuals' movements, their emotional state, and intentions (Miles, Nind, & Macrae, 2009; Richardson, Marsh, & Baron, 2007a). Hence, joint actions play an important role in strengthening interpersonal connections to achieve a common goal (Lakin, Jefferis, Cheng, & Chartrand, 2003; Michael, Sebanz, & Knoblich, 2016). In many social situations, joint actions often rely on executing bodily movements in a timely manner. Time-critical execution requires varying degrees of temporal coordination and precise, moment-to-moment control over our own limbs and body (Pezzulo, Donnerumma, & Dindo, 2013; Vesper, Schmitz, Safra, Sebanz, & Knoblich, 2016).

During passive viewing, however, inferring that two people form a coordinated unit requires that the human visual system detect a set of features indicating social cooperation (Bernieri, 1988; Richardson, Marsh, Isenhower, Goodman, & Schmidt, 2007b). Researchers have found important evidence that both bottom-up and top-down processes facilitate joint action identification. de la Rosa et al. (2013) have shown a connection between social interaction recognition performance and the velocity of specific body joints (e.g., the arms, feet, and hips), and Thurman and Lu (2014) found that participants can identify spatially-scrambled displays of human dancers as long as movement between individual joints and global bodies was congruent. When viewing point-light actions involving another person, these types of actions are shown to enhance recognition even when embedded in a noisy background (Manera et al., 2011; Neri, Luu, & Levi, 2006; Su, van Boxtel, & Lu, 2016).

The present study investigated a common set of visual features for inferring cooperation while observing point-light displays of joint actions. If inferences regarding cooperation depend on the perception of critical visual features, do groups of features emerge to form categories, helping to differentiate between types of joint actions? We manipulated the temporal alignment between actors and their movements to measure sensitivity to motion synchrony and cooperation. To gauge a common set of characteristics among ten different joint actions, we measured how human judgments of interactivity change as a function of temporal misalignment. Lastly, previous research has shown that recognition sensitivity varies according to different categories of individual actions (Dittrich, 1993; van Boxtel & Lu, 2011) and social interactions having multiple levels of categorization (de la Rosa et al., 2014). We use a statistical clustering model (latent dirichlet allocation) to categorize joint actions based on shifts in “interactivity” ratings given temporal misalignment of the actors’ body movements.

Method

Participants

A total of 55 undergraduate students ($M_{age} = 21.4$, $SD_{age} = 4.3$, 42 females) at the University of California, Los Angeles (UCLA) with normal or corrected-to-normal vision enrolled in the study. Participants gave informed consent as approved by the UCLA Institutional Review Board and were provided with course credit in exchange for their participation.

Stimuli and materials

Participants viewed short action sequences showing two actors performing interpersonal interactions. We used ten different joint actions obtained from the CMU motion capture database (<http://mocap.cs.cmu.edu>): *approach & high-five*, *greet & shake hands*, *passing an object*, *playing catch*, *chicken dancing*, *tug-of-war*, *arguing & gesturing*, *threatening*, *circular skipping*, and *salsa dancing*. The durations of the actions were in the range of 1–3.67 seconds. Action sequences were determined by first finding a *key frame* from the full-action file (e.g., locating the hand-contact point of a high-five, or when the partner twirls the other during salsa dancing), finding a start frame of the sequence (approaching each other to give the high-five or dance steps that build up to twirling), and then calculating the number of frames between start and key frame (length of the bar between frame 1 and the vertical red line in Figure 1). The same number of frames were added *after* the key frame to determine the end frame. The total duration was action specific since key interactive moments varied significantly in length from one action to another (e.g.,

momentary shaking of hands vs. an entire choreographed dance sequence).

For each trial, a participant viewed two actors engaged in a joint action in its original, synchronized form or with a manipulation of temporal offset (Figure 1). For an offset manipulation, one actor was selected as a reference actor for each joint action. The second actor was shifted backward in time, resulting in a lag behind the reference actor (negative offset values), or shifted forward in time ahead of the reference actor (positive offset values). We temporally shifted the second actor in seven, equally spaced increments within the range of the action’s start and end frames. Three offsets were negative (lagging), three positive (leading), and one centered at zero (synchronized actors in their original form). The top of Figure 2 shows the actor positions at different temporal offset positions for the action *approach & high-five*, with offset magnitudes (ms) under the bar plots.

Actors were presented as skeletal figures by connecting 17 body markers from each actor. Body marker coordinates were scaled with the BioMotion Toolbox (van Boxtel & Lu, 2013). The skeletal outlines were white lines with a thickness of 0.17° visual angle, superimposed with white dots as body joint markers (a diameter of 0.2°), on a black background. The viewing distance was 34.5 cm. A 60-Hz CRT monitor was used to display the stimuli. All joint actions and additional visual angle information are shown in Figure 2.

Procedure

After viewing two actors engaged in a joint action, either in its synchronized form or with a temporal offset, participants were asked to provide a rating judging “the degree to which the two actors appear to be interacting.” We used a 7-point scale, with 1 indicating “certainly not interactive” and 7 indicating “certainly interactive.” They were permitted up to 20 seconds to make a response, allowing ample time to rate the interaction. They were not provided with a definition of “interacting” to avoid biasing participants to specific visual indicators of interpersonal actions, nor were they explicitly informed of temporal misalignment between actors. The experiment included a total of 280 randomized trials based on 7 offsets, 10 joint actions, and 4 repetitions. After completing 50 trials, participants were allowed to take a short break up to 20 seconds before continuing with the experiment.

Results

We first conducted an analysis to observe changes in ratings due to joint action type, temporal offset, or the interaction between the two factors. We fit a linear mixed effects model with the participant as a random effect, adding temporal offset as a random factor that varies for each participant. The offset

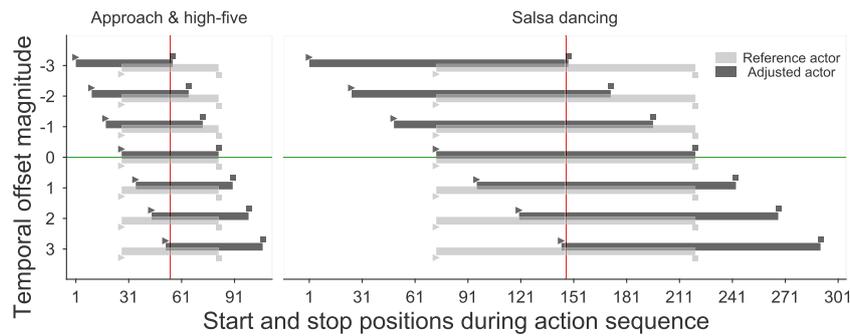


Figure 1. Example of temporal misalignment for two of the joint actions. If the offset is zero (green horizontal line), the reference and adjusted actor start on the same frame and are in sync. The start position for the adjusted actor is behind or ahead of the reference actor

and varies by temporal offset magnitude and action type. Both actors always cross the key action frame (red vertical line). The length of the horizontal bars indicates the total action sequence duration.

factor consisted of four levels, the synchronized action with no offset (level 0) and three other levels of absolute offset magnitude (levels 1–3, omitting indicators of leading or lagging offsets). We found a significant interaction between joint action and temporal offset using either an ANOVA method with

Satterthwaite approximations of the degrees of freedom ($F(27, 15290) = 8.5, p < 0.001$), or a likelihood-ratio test between models with and without the interaction term ($\chi^2(27) = 227.5, p < 0.001$). The significant interaction indicates that participant ratings vary according to temporal misalignment for some

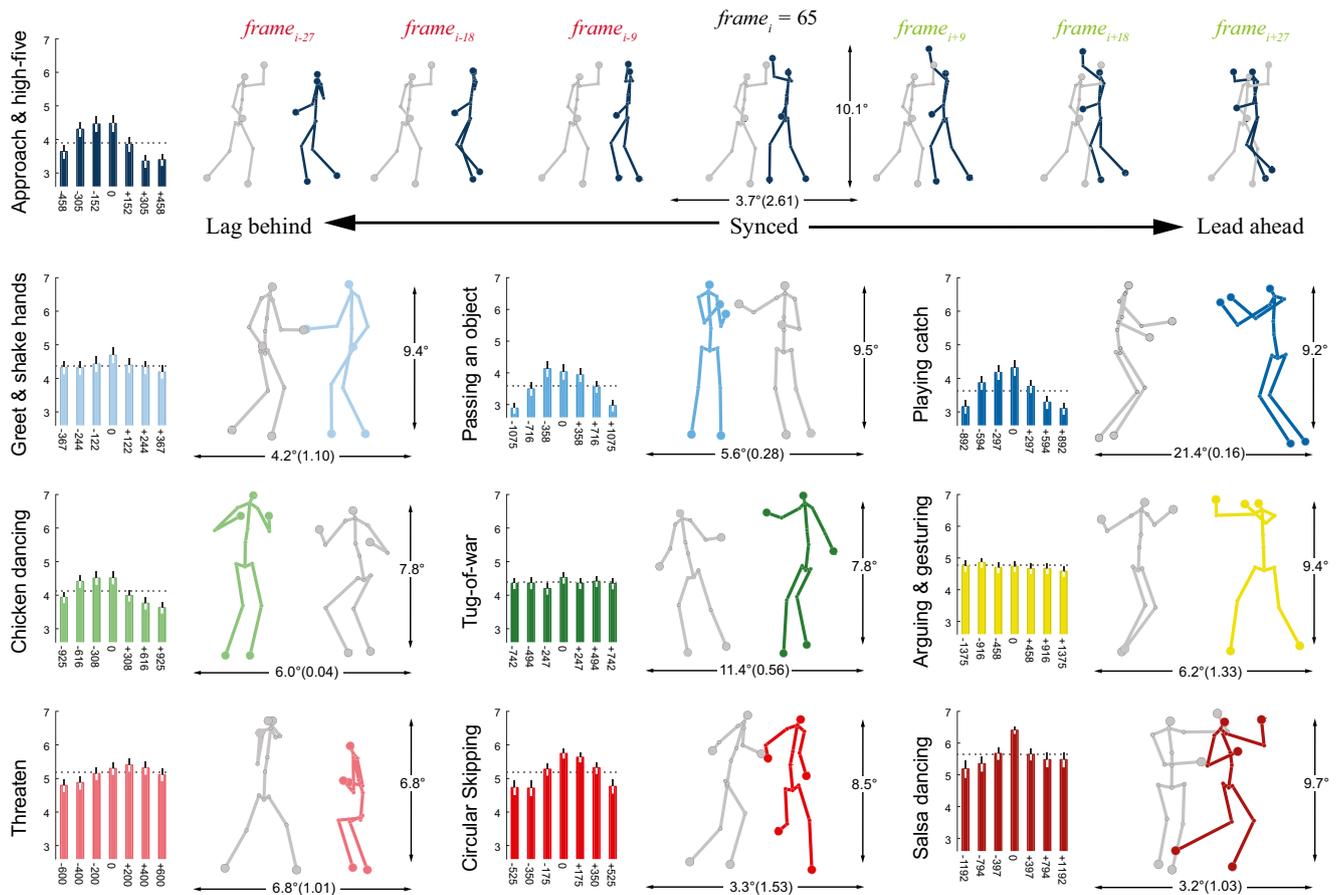


Figure 2. Joint action stimuli and temporal offset ratings. The top of the figure shows the action *high-five* at the 65th frame for the reference actor (gray) and relative position of the adjusted actor (blue) for each temporal offset. The bar plots show human ratings as a function of temporal offset, with the dashed horizontal line indicating the average rating across offsets. Standard errors of the mean are shown for each bar. The key frame of interactivity is

shown to the right of each bar plot. Vertical lines with arrows show the mean vertical visual angle between the maximum and minimum joints across frames. Horizontal lines show the mean (SD) horizontal visual angle of the action sequence. Sizes of skeletal lines and points are enlarged for illustration purposes. The color scheme used for actions (blues, greens, yellow, reds) indicates class membership (used for all figures).

subset of actions. We further examine the effect of temporal misalignment on each action by analyzing the shapes of the discrete interactivity rating distributions.

Shape of rating distributions

Some actions in Figure 2 resulted in flat rating distributions across offsets, indicating temporal misalignment had a negligible effect on interactivity ratings (e.g., *arguing* & *gesturing* and *tug-of-war*), whereas other actions showed a peak in the distribution at offset zero, indicating sensitivity to temporally-aligned actions. We developed a “peakedness” index specifically for the discrete, multivariate offset distributions that captures the change in sensitivity between truly coupled joint actions and temporally misaligned actions. The peakedness index was calculated by subtracting ratings at offset zero (unaltered, synchronous actions) from the mean of all other non-zero offsets. Because the zero-offset condition corresponds to the truly coupled, synchronized, and coordinated actions, a large rating difference between the zero-offset condition and other non-zero offset conditions reflects higher sensitivity to temporal coordination and signals interactivity between actors. As the peakedness index approaches zero, the shape of the rating distribution becomes flatter, indicating that interactivity ratings do not reflect participant discrimination among truly coupled joint actions and temporally misaligned actions. A second characteristic of the offset distribution is asymmetry, and reflects participants’ sensitivity to the directionality of temporal offsets. For example, the joint action *approach* & *high-five* was rated to be less interactive if the second actor was lagging behind the reference actor, but more interactive if shifted forward in time. We quantified asymmetry by subtracting mean ratings for the negative offset conditions from the mean of positive offset conditions (omitting the zero-offset condition). Note that the indices defined in the present study are different from the statistical moments of kurtosis and skewness for measuring peakedness and symmetry of a distribution. Our indices are calculated using the zero-offset condition as a reference point, as opposed to the mean and standard deviation of some random variable as in the statistical definitions.

We tested whether the peakedness and asymmetry indices for each joint action were significantly different from zero using Bonferroni adjusted *t*-tests. Six of the ten joint actions showed significant peakedness, with a unimodal rating distribution and decreased ratings for non-zero offset conditions (Figure 3). For these six joint actions, participants showed high sensitivity to temporally synchronized body movements between the two actors and rated actions with temporal misalignment as less interactive. The asymmetry indices showed a directionality effect of temporal offset for five of the ten joint actions. Figure 3 shows that *threaten* and *circular skipping*

actions yielded higher ratings when the shifted actor was ahead of the reference actor compared with lagging behind.

Categorization of joint actions

Our approach toward interactivity categorization analyzed the rating frequency distribution for each joint action and each temporal offset. Each of the 70 unique frequency distributions (10 actions \times 7 offsets) was obtained from counting the frequency of ratings 1 to 7 from 220 responses (4 repeated trials \times 55 participants). The top panel of Figure 4 shows the rating frequency distributions for four of the ten joint actions at each offset. We fit a generative clustering model to estimate probabilities of assigning each rating frequency distribution to a latent class. The clustering algorithm is a variant of the latent Dirichlet allocation model (LDA, also known as topic models), which can discover the number of clusters and allow for latent classes to be correlated with one another (Blei & Lafferty, 2007). The resulting latent classes represent categories of joint actions, in which the joint actions in one category tend to share similar rating distributions. The model assumes that each latent class is a weighted mixture of *all* rating distributions, instead of each distribution being exclusive to a specific class. For example, each of the 28 rating frequency distributions shown in Figure 4 may have a high probability for one class (and lower probabilities for the rest), whereas others may have high probabilities for different classes. We chose a parameter of 10 as the maximum number of latent classes, based on the conservative assumption that each joint action forms its own unique category. The clustering results showed that only 4 of 10 latent classes were needed to account for the majority of the variability (72%) in the rating frequency distributions (bottom left panel of Figure 4). The size of the nodes in Figure 4 (bottom right) show that some actions were highly representative of a specific class (e.g., *salsa dancing*), whereas others shared membership between classes (e.g., *arguing*).

To visualize the category similarity between joint actions, further analyses were conducted by examining the matrix containing the clustering probabilities of each joint action for each class. We first computed the Hellinger distance, suitable in this case for comparing similarity among probability distributions, and then applied multidimensional scaling (MDS) on the distance matrix to visualize the categories of joint actions (Figure 4, bottom right). We found lower ratings of interactivity but high sensitivity to synchronized coordination in joint actions such as *passing an object*, medium interactivity ratings and more tolerance to temporal misalignment in joint actions such as *tug-of-war*, and high interactivity ratings and high sensitivity to temporal coordination in joint actions such as *salsa dancing*. Thus, the interactive ratings of joint actions can be used to categorize joint actions according to two critical features: tolerance to temporal offset and the degree of interactivity involved in a joint action. Joint actions with a high

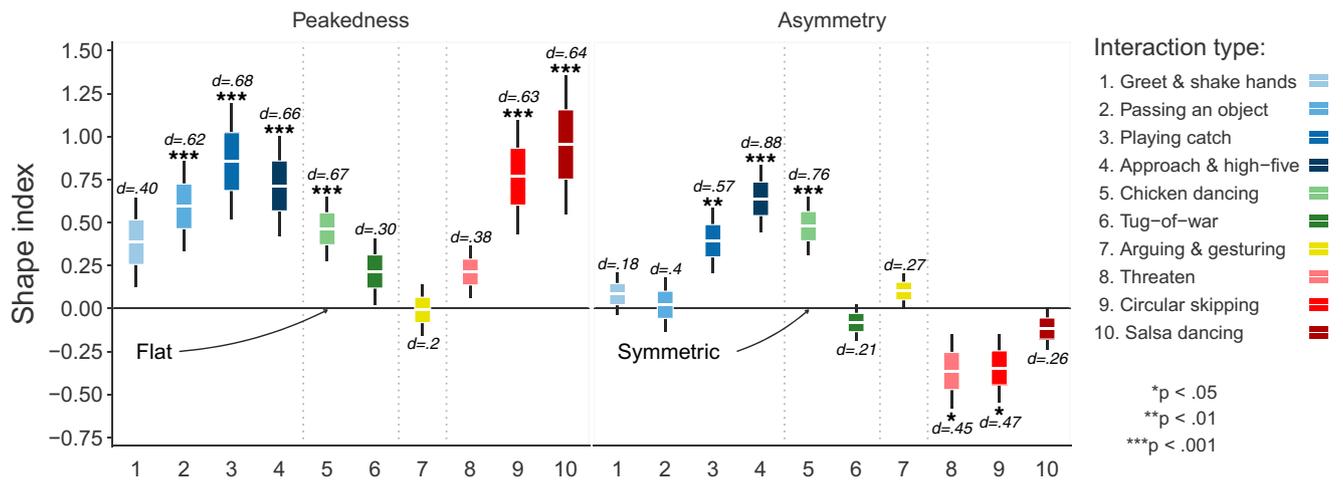


Figure 3. Shape indices of asymmetry and peakedness from human rating distributions for each joint action. Positive values for peakedness indicate higher ratings for zero offset (synced) conditions than those with temporal misalignment. Positive values for asymmetry indicate positive skewness from higher ratings for “lag” offsets over “lead” offsets. The

middle of the bars marks the average values of the indices, with bar length displaying the standard error of the mean. Lines extending from bars display the 95% confidence interval. Asterisks indicate significant *p* values after adjusting for multiple comparisons. Effect sizes (Cohen’s *d*) are shown above or below the bars.

probability value for a single class (indicated by node size) are more spatially distant from the center (e.g., *salsa dancing*),

and actions more likely to share multiple classes are near the center (with smaller nodes, e.g., *shake hands*).

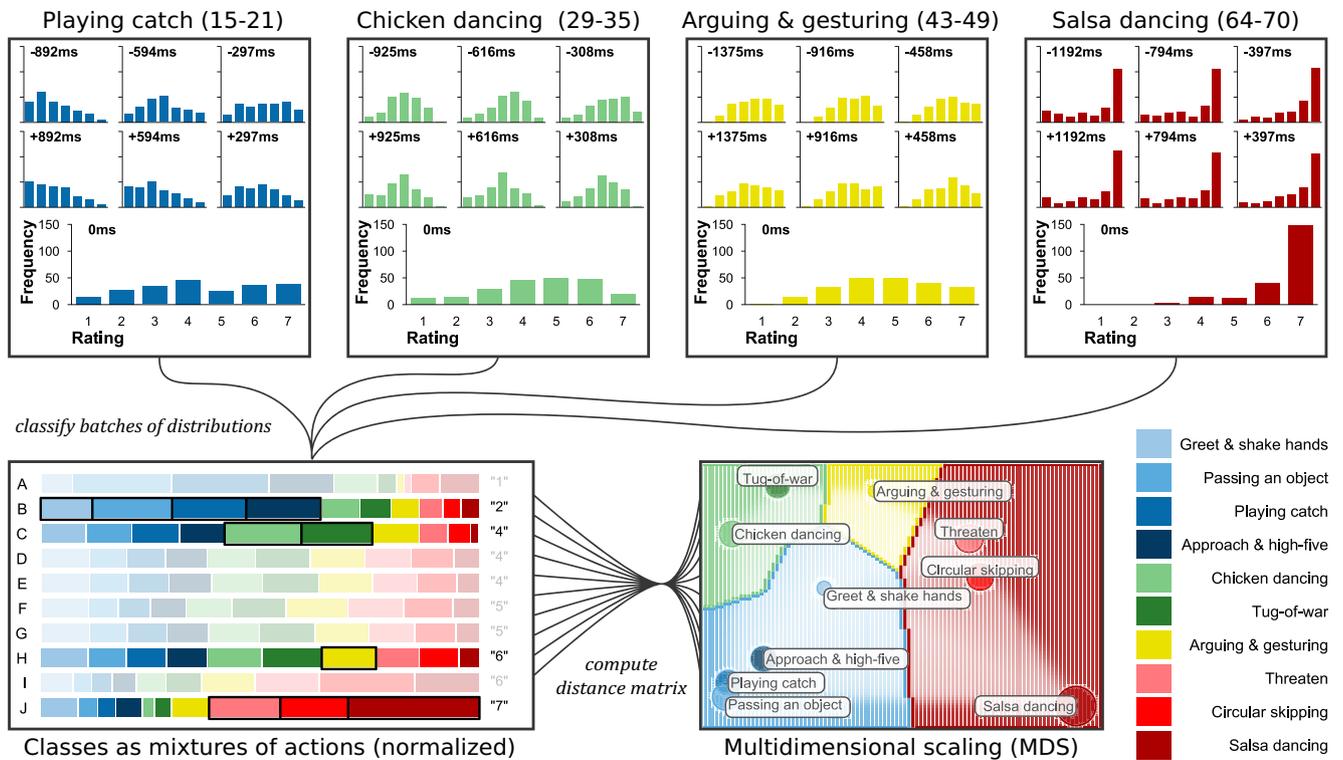


Figure 4. Classification approach and categorization results. *Top.* Examples of interactivity rating distributions for a few joint actions. The colors indicate different joint actions. *Bottom left.* Latent classes inferred by the probabilistic clustering model. Distinct categories are denoted by the letter on the left. The most likely interactivity rating for each category is shown on the right. Individual bar segments are joint actions, with the length proportional to its probability within a category. Black frames around segments indicate the most probable category

assignment for each action. *Bottom right.* Similarity among joint actions. Action locations were obtained by computing the Hellinger distance between posterior probabilities, and then applying multidimensional scaling (MDS) to the distance matrix. The vertical axis organizes joint actions from low (top) to high (bottom) offset sensitivity. The horizontal axis organizes them by average rating, from low (left) to high (right). Nodes are color coded by class, and the size indicates the probability of belonging to that class.

Discussion

In the current study, we tested the visual perception of social cooperation from the perspective of temporal misalignment having a graded influence on interactivity ratings, which may result in the use of graded features to guide these judgments. Although the clustering analysis makes no assumptions about the underlying nature of joint action categories, we can examine common properties among categories by comparing characteristics among joint actions given category membership. To gauge which visual features may play an important role in judging social cooperation from observed joint actions, we conjectured that the manipulation of temporal misalignment of joint action affects the reliability of detecting critical features, which results in systematic changes in rating distributions among actions in the same category. We specifically examined the potential role of three important visual features for identifying the interactivity of joint actions: global motion synchrony, spatial proximity, and salient moments of interpersonal coordination.

We observed that higher interactivity ratings were more likely to correspond to actions in which there was continuous, close contact between persons, particularly when bodies were physically connected, as in *salsa dancing* and *circular skipping*. Dyadic movements during these displays were dynamic, cyclical, and often intersecting or connected by locking arms or holding hands. Body movements of the two actors were well coordinated for most of the duration of the activity; thus temporal shifts of a few hundred milliseconds were easily detectable. Observers are tuned to global motion synchrony and are sensitive to subtle changes in temporal offset for these types of joint actions, in which local regions of oscillatory motion become anti-phasic. Spatial proximity between the two actors involved in joint actions is another important visual feature. The results shown in Table 1 indicate that, as spatial distance between actors increased, participants were less likely to give higher ratings of interactivity. This finding is consistent with the previous research (de la Rosa et al., 2013; Shu, Thurman, Chen, Zhu, & Lu, 2016). Spatial proximity may make it more difficult to detect the coordination between common body limbs, and thus decreases judgments of interpersonal cooperation. Similarly, for the joint actions that yielded above-average interactivity ratings despite large temporal

shifts, spatial proximity may be used as an immediate indicator of meaningful interaction (e.g., in *tug-of-war* or *arguing*), especially when subtle changes in coordination are difficult to detect. However, use of spatial proximity as a general cue for joint actions does not imply that observers are not influenced by task-specific cues. For instance, for actions that involved an object that was not shown in the stimulus (e.g., *passing an object*), interactivity may have been more difficult to judge without this critical piece of visual information that binds the two distant actors. The third critical feature indicates brief salient moments of interactivity during joint action. Events such as two persons approaching each other and shaking hands, or giving a high-five as they pass, are salient interactions lasting for a short duration. For these joint actions, a small window of opportunity is provided that clearly demonstrates synchrony and coordination, even though most of the movement leading up to and after the critical event does not appear to be well-coordinated between actors.

In addition to visual features signaling joint action activity, other high-level cognitive processes may also be involved. For example, we observed the temporal manipulations yielded asymmetrical rating averages between leading and lagging conditions. This may result from the distinct roles each actor plays for some of the actions, consistent with the literature on real-time action prediction (Graf et al., 2007) and causal-effect relations in actions (Peng, Thurman, & Lu, 2017). For example, the *threaten* stimulus displayed the attacker as the reference actor while shifting the defender actor. When the defender's body movements were shifted in time ahead of the attacker, the ratings were higher than the ones for lagging offsets, suggesting that people may use top-down cues of the defender's movement to predict the attacker's reaction. For other joint actions, the asymmetry effect may be due to bottom-up processing of unnatural body displacements when shifted in a negative or positive temporal direction. *Circular skipping*, for example, shows changes in skipping speed from beginning to end, and in this case, temporal misalignment can yield unnatural events, such as actors colliding with one another or actors showing noticeably different velocities in body displacement for negative offsets.

Joint action recognition may manifest itself at a later stage of processing after deciding that two individuals are engaged in social cooperation, in which previous experience, motor

Table 1. Average interactivity ratings shown with the average distance between actors (in horizontal visual angle units) by action category. As distance between actors increases, average rating within the corresponding category decreases.

Category	Joint actions	Avg. rating score (SD)	Avg. horz. angle (SD)
B	Greet & shake hands; approach & high-five; playing catch; passing an object	3.87 (1.74)	8.67 (1.47)
C	Tug-of-war; chicken dancing	4.25 (1.46)	8.27 (0.59)
H	Arguing & gesturing	4.76 (1.51)	6.24 (1.06)
J	Salsa dancing; circular skipping; threaten	5.34 (1.68)	4.08 (1.25)

repertoires, and goal-oriented inference may play a role in recognition and further interpretation of the nature of joint action (Casile & Giese, 2006). Joint actions that generated high ratings of interactivity, and were sensitive to temporal shifts in movement, likely had multiple visual features that observers could use when interpreting the degree of interpersonal coordination. Further work is necessary to examine how individuals judging interactivity weigh multiple competing visual features among a wide range of joint actions. This line of research can provide insight into how people evaluate other humans in terms of their potential to cooperate and help execute specific action goals.

References

- Bemieri, F. J. (1988). Coordinated movement and rapport in teacher-student interactions. *Journal of Nonverbal Behavior*, *12*(2), 120–138. doi:<https://doi.org/10.1007/BF00986930>
- Blei, D. M., & Lafferty, J. D. (2007). A correlated topic model of science. *The Annals of Applied Statistics*, *1*(1), 17–35.
- Casile, A., & Giese, M. A. (2006). Nonvisual motor training influences biological motion perception. *Current Biology*, *16*(1), 69–74. doi:<https://doi.org/10.1016/j.cub.2005.10.071>
- de la Rosa, S., Choudhery, R. N., Curio, C., Ullman, S., Assif, L., & Bühlhoff, H. H. (2014). Visual categorization of social interactions. *Visual Cognition*, *22*(9), 1233–1271. doi:<https://doi.org/10.1080/13506285.2014.991368>
- de la Rosa, S., Mieskes, S., Bühlhoff, H. H., & Curio, C. (2013). View dependencies in the visual recognition of social interactions. *Frontiers in Psychology*, *4*. doi:<https://doi.org/10.3389/fpsyg.2013.00752>
- Dittrich, W. H. (1993). Action categories and the perception of biological motion. *Perception*, *22*(1), 15–22. doi:<https://doi.org/10.1068/p220015>
- Graf, M., Reitzner, B., Corves, C., Casile, A., Giese, M., & Prinz, W. (2007). Predicting point-light actions in real-time. *NeuroImage*, *36*, T22–T32. doi:<https://doi.org/10.1016/j.neuroimage.2007.03.017>
- Lakin, J. L., Jefferis, V. E., Cheng, C. M., & Chartrand, T. L. (2003). The Chameleon Effect as Social Glue: Evidence for the evolutionary significance of nonconscious mimicry. *Journal of Nonverbal Behavior*, *27*(3), 145–162. doi:<https://doi.org/10.1023/A:1025389814290>
- Manera, V., Becchio, C., Schouten, B., Bara, B. G., & Verfaillie, K. (2011). Communicative interactions improve visual detection of biological motion. *PLOS ONE*, *6*(1), e14594. doi:<https://doi.org/10.1371/journal.pone.0014594>
- Michael, J., Sebanz, N., & Knoblich, G. (2016). Observing joint action: Coordination creates commitment. *Cognition*, *157*, 106–113. doi:<https://doi.org/10.1016/j.cognition.2016.08.024>
- Miles, L. K., Nind, L. K., & Macrae, C. N. (2009). The rhythm of rapport: Interpersonal synchrony and social perception. *Journal of Experimental Social Psychology*, *45*(3), 585–589. doi:<https://doi.org/10.1016/j.jesp.2009.02.002>
- Neri, P., Luu, J. Y., & Levi, D. M. (2006). Meaningful interactions can enhance visual discrimination of human agents. *Nature Neuroscience*, *9*(9), 1186–1192. doi:<https://doi.org/10.1038/nm1759>
- Peng, Y., Thuman, S. M., & Lu, H. (2017). Causal action: A fundamental constraint on perception and inference about body movements. *Psychological Science*, *28*(6), 798–807. doi:<https://doi.org/10.1177/0956797617697739>
- Pezzulo, G., Donnarumma, F., & Dindo, H. (2013). Human Sensorimotor Communication: A Theory of Signaling in Online Social Interactions. *PLOS ONE*, *8*(11), e79876. doi:<https://doi.org/10.1371/journal.pone.0079876>
- Richardson, M. J., Marsh, K. L., & Baron, R. M. (2007a). Judging and actualizing intrapersonal and interpersonal affordances. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(4), 845–859. doi:<https://doi.org/10.1037/0096-1523.33.4.845>
- Richardson, M. J., Marsh, K. L., Isenhower, R. W., Goodman, J. R. L., & Schmidt, R. C. (2007b). Rocking together: Dynamics of intentional and unintentional interpersonal coordination. *Human Movement Science*, *26*(6), 867–891. doi:<https://doi.org/10.1016/j.humov.2007.07.002>
- Sebanz, N., & Knoblich, G. (2009). Prediction in joint action: What, when, and where. *Topics in Cognitive Science*, *1*(2), 353–367. doi:<https://doi.org/10.1111/j.1756-8765.2009.01024.x>
- Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*, *10*(2), 70–76. doi:<https://doi.org/10.1016/j.tics.2005.12.009>
- Shu, T., Thuman, S. M., Chen, D., Zhu, S.-C., & Lu, H. (2016). Critical features of joint actions that signal human interaction. In A. Papafragou, D. Grodner, D. Mirman, & J. C. Trueswell (Eds.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 574–579). Austin, TX: Cognitive Science Society.
- Su, J., van Boxtel, J. J. A., & Lu, H. (2016). Social interactions receive priority to conscious perception. *PLOS ONE*, *11*(8), e0160468. doi:<https://doi.org/10.1371/journal.pone.0160468>
- Thuman, S. M., & Lu, H. (2014). Perception of social interactions for spatially scrambled biological motion. *PLOS ONE*, *9*(11), e112539. doi:<https://doi.org/10.1371/journal.pone.0112539>
- van Boxtel, J. J. A., & Lu, H. (2011). Visual search by action category. *Journal of Vision*, *11*(7), 19. doi:<https://doi.org/10.1167/11.7.19>
- van Boxtel, J. J. A., & Lu, H. (2013). A biological motion toolbox for reading, displaying, and manipulating motion capture data in research settings. *Journal of Vision*, *13*(12). doi:<https://doi.org/10.1167/13.12.7>
- 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. doi:<https://doi.org/10.1016/j.cognition.2016.05.002>