Title
Joint Improvisation; Perception of Togetherness in Contemporary Dance Performance

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors
Speranza, Trinidad Belén
Flores Bravo, Ivonnia M
Ramenzoni, Verónica

Publication Date
2024

Peer reviewed

Joint Improvisation; Perception of Togetherness in Contemporary Dance Performance

Speranza, Trinidad B. (trinidadsperanza@uca.edu.ar)
National Scientific and Technical Research Council; Godoy Cruz 2290, C1425 CABA, Argentina
Centro de Investigaciones en Psicología y Psicopedagogía, Universidad Católica Argentina; Alicia Moreau de Justo 1800, C1107 CABA, Argentina

Flores Bravo, Ivonnia M. (ivonniaflores@uca.edu.ar)
National Scientific and Technical Research Council; Godoy Cruz 2290, C1425 CABA, Argentina
Centro de Investigaciones en Psicología y Psicopedagogía, Universidad Católica Argentina; Alicia Moreau de Justo 1800, C1107 CABA, Argentina

Ramenzoni, Verónica C. (veronicaramenzoni@uca.edu.ar)
National Scientific and Technical Research Council; Godoy Cruz 2290, C1425 CABA, Argentina
Centro de Investigaciones en Psicología y Psicopedagogía, Universidad Católica Argentina; Alicia Moreau de Justo 1800, C1107 CABA, Argentina

Abstract

Joint improvisation is central to how we navigate the social world, engage and maintain social interactions, and perceive interactions between other people. This project investigates people's ability to distinguish between joint and individual actions (contemporary joint vs. solo dance improvisation) and the information they use to make this determination. In Experiment 1, participants were asked to identify whether two people were improvising dance movements together or alone. Experiment 2 explored how much people's decision-making relies on information about the dancers' facial expressions and gaze direction. Overall, results showed we can accurately identify improvised joint actions, even when the actors' faces and gaze direction are occluded.

Keywords: joint improvisation; joint action; eye gaze; dance; perception.

Introduction

Imagine you are in a dance club and spot someone attractive dancing a short distance away. The idea might cross your mind to introduce yourself or even invite them to dance. You watch them dance for a bit and change your mind; it looks like they are already dancing with someone else. How did you make this determination? What in the person's behavior made you believe they were not dancing alone? We encounter situations like this every day, though we rarely pay attention to the decision-making process behind them. This ability is critical to navigating social landscapes, from the mundane decision to walk between people who seem to be walking separately to the potentially embarrassing decision to approach people who might not be alone at a bar.

Joint actions, unlike single or individual actions, are defined by a common goal that can only be achieved through cooperative interaction. Individual's actions within a joint task are linked and occur in response to the past and future actions of other people (Knoblich, 2011; Knoblich & Sebanz, 2006; Wilson & Knoblich, 2005; Wiltermuth & Heath, 2009). Though joint actions might involve a planning component and online monitoring, they often emerge spontaneously between people (Ramenzoni et al., 2012; Riley et al., 2011). In particular, improvised joint actions are characterized by their creative and spontaneous nature. They rely on the dynamic and spontaneous exchange of information between agents (Brinck, 2016). The agents' receptivity and responsiveness to each other's behavior and the opportunities it offers for creation give rise to unique and unplanned phenomena (Noy et al., 2011; Saint-Germier et al., 2021).

Research on joint improvising has approached the phenomena from both ends. Bottom-up studies have identified the emergence of coordination between agents and dependency on task constraints as the signatures of jointness (see Riley et al., 2011). Conversely, top-down approaches have focused on the cognitive abilities that make them possible (see Sebanz & Knoblich, 2021). For instance, Noy and collaborators (2011) investigated how expert and novice improvisers performed a classic theater exercise called the mirror game. They assessed how improvisers synchronized their movements, how well they imitated each other, and how they distributed roles within the interaction. They found that participants co-create complex, novel, and synchronized movements even without a designated leader. Hart and colleagues (2014) used the mirror game to establish that individuality and togetherness were essential factors in joint improvisational movement. While performers' movements coordinate over time, they also maintain a degree of individuality that allows each performer to express their preferences and style.

Perception of joint improvised actions has received less attention despite an increased interest in understanding the aesthetic experience of observing dance improvisation or listening to live music jamming. A recent study by McEllin and colleagues (2020) investigated whether the type of synchronization the agents exhibit affects the ability to

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

perceive performances as coordinated and the aesthetic experience that arises from it. Participants watched two moving dots on a screen and were told that these reflected the hand movements of two performers engaged in joint improvisation. The authors defined interval-based synchronization as the time alignment of movements to reach their endpoint simultaneously and velocity-based synchronization as the continuous alignment of movements throughout the trial. Performances exhibiting velocity-based synchrony were perceived as more coordinated than those exhibiting interval-based synchrony. In addition, performers were rated as more alike and beautiful in the velocity-based condition, providing observers with a more robust aesthetic experience. These findings suggest that the degree of coordination observed can be essential to how joint improvisation is experienced (McEllin et al., 2020).

Though improvising is fundamental to everyday activities and cultural endeavors, it is more readily recognized in contemporary arts such as jazz jamming or contemporary music. Dance provides a unique medium for the study of joint improvisation. It naturally unites people, creating a sense of togetherness while preserving all the necessary components of a joint task: dancers share a common goal, the roles between them are balanced, performance relies on responsiveness and adaptability, and it allows for the emergence of synergies between partners.

Studies seeking to understand the underlying cognitive mechanisms that support action understanding have investigated dancers from early on. Calvo-Merino and collaborators (2005) influential study of ballet and capoeira dancers provided early evidence of the importance of mirror neurons for action understanding; it indicated that the same neuronal system is activated when people observe actions that they performed and similar actions performed by other people. A follow-up study on expert dancers reached similar conclusions but applied it specifically to the activations that occur when simulating a previously learned choreographic sequence (Cross et al., 2006). These findings provided one of the premises upon which joint action theory was developed: that the perception of self-produced actions can allow individuals to distinguish their own actions from those of others (Repp & Knoblich, 2004). Loula and collaborators (2005) further demonstrated that this distinction can be drawn based solely on kinematic information. Participants can recognize themselves and their best friends in point-light displays of dancing, jumping, or boxing scenes. The self-other distinction plays a role in the aesthetic experience of observers. People’s aesthetic evaluations of dance change depending on their own physical ability to reproduce the movements they watch, with preferences skewing towards movements perceived to be difficult to perform (Cross et al., 2011).

Dance has also contributed to our understanding of another central component of joint actions: shared intentionality (Vesper et al., 2010). The ability to share intentions with other agents toward attaining a common goal has been hypothesized to lay the core of cultural cognition and evolution (Tomasello et al., 2005). Shared intentionality has been found to promote cooperation and synchronicity between dancers. Reddish and colleagues (2013) found that when dancing, dancers who observed synchrony and shared a goal to produce synchrony, received immediate feedback which led to successful cooperation among partners. Perceiving synchrony between performers is also a significant predictor of aesthetic appreciation in spectators of live dance performances. Performances are experienced as more likable when performers share a common goal to perform and coordinate their actions in time (Vicary et al., 2017). Shared intentionality and coordination are thus linked within joint actions, though their relationship is not one of causality. Interpersonal coordination might emerge spontaneously as a byproduct of the joint action and aid co-actors during task performance (Vesper & Sebanz, 2016).

Coordination within an improvised dance performance can be easily perceived in dance styles that require physical contact between partners, such as tango or salsa. Other dance styles, such as the twist, rely on the soft-assembly between partners and information communicated through head movements, eye contact, and facial expressions to build interpersonal coordination. Eye gaze plays a dual role in conveying and receiving information; the direction of a person's gaze serves as a clear indicator of their focus, whether it is directed towards another individual, a noteworthy environmental stimulus, or the intended direction of their movement (Bishop et al., 2019). Studies on performers established that eye-gaze is critical to the distribution of turn-taking during dance improvisations (Evola, et al., 2015) and that musical partners that share information through eye-gaze develop feelings of engagement and creative collaboration (Bishop et al., 2019; Evola & Skubisz, 2019). More importantly, access to information about gaze direction (i.e., where people look during an interaction) allows observers to infer shared intentionality (Böckler et al., 2011; Huang et al., 2015; Stephenson et al., 2021).

The current project investigates the ability to distinguish between spontaneous joint and solo actions and the information people use in decision-making. To assess this ability, we developed a stimulus set of contemporary dyadic dance performances and asked novice participants to determine whether they were watching a joint or solo performance. In addition, we explored whether their decisions varied depending on the tempo of the movements observed (Experiment 1) and the availability of information about the dancers’ facial features and gaze direction (Experiment 2).

Methods and Results

Experiment 1

This experiment aimed to test whether participants could identify if two people in a video were improvising a dance together or alone. We expected that participants would be able to make this decision successfully above chance. Based on McEllin's findings (2020), we also expected that detection ability would vary depending on the tempo of the performance. The present project is pre-registered on the Open Science Framework: https://osf.io/ah869

Participants 23 females (*mean* age =25.22 yrs; *SD* age = 4.61 yrs) participated in this experiment. Participants were recruited through advertisements on social media groups of college students. All participants were recruited at the Centro de Investigaciones en Psicología y Psicopedagogía de la Universidad Católica Argentina. Participants provided written informed consent prior to the beginning of the study; forms were approved by the ethics review board of the Fundación Favaloro Bioethics committee (approval number CBE 940/21). Participants received no monetary or other compensation for participating in the study. Inclusion criteria were 18-to-35-years-old, limited or no dance experience, and normal or corrected-to-normal vision. Dance experience was qualified by the completion of a questionnaire that assessed prior experience performing and watching dance. No participant had formal training in contemporary dance. When asked to evaluate their ability as a dancer on a 1- to 5 scale (1 = awful; 2 = bad; 3 = intermediate; 4 = good; 5 = very good), participants scored themselves with a mean rating of 2.8 (SD =0.5). To quantify experience with dance observation, the mean number of professional dance performances (or theatre/opera performances with some dance element) attended in the last year by participants was 1.9 (SD =3.01). Both measures were based on Cross et al., 2011.

Materials and Procedure The stimulus set featured female dancers improvising contemporary dance movements. Dancers were informed about the goal and hypothesis of the study and gave their authorization to use their image both in the experiments and in the works to be published. Four dancers with similar levels of expertise and anthropometrics (mean of height = 154 cm) were recruited. Dancers trained within a contemporary dance company (12 members) and perform weekly together. They were instructed to use the vocabulary taught by the company to improvise a range of movements varying in complexity, speed, and amplitude. They were recorded in profile dancing alone and in dyads to the same music piece at its original (steady beat; 120 bpm), a slowed down (85 bpm), and a speeded-up tempo (155 bpm). All music pieces were 1 minute long and extracted from the same song: "Buenos Aires'' by Nathy Peluso, an Argentinian artist (see original here). Dancers stood at a zero-point set by the researcher and moved within a fixed quadrant (2.74 m high x 2.52 m wide). In the together condition, they were 2.5 m apart from each other, and in the solo condition, they occupied the same space. Dancers were prevented from seeing each other dance when not performing. The videos were recorded on a 12-megapixel camera from a fixed tripod. The tripod was placed at the midpoint between dancers at 4 m away from them to record the whole scene. Dyads first performed the original version of the song, followed by the slowed-down and the sped-up versions (counterbalanced). To control for concurrent experience dancing together biasing solo performance, one dyad was first recorded dancing together and the other dancing solo. For couple number A, the videos of them dancing together were filmed first, and for couple number B, the videos of them dancing together were filmed first.

Stimuli were extracted in post-processing from the recorded masters and manipulated to show both dancers on the screen (see Figure 1) in 7 sec. clips (see sample here). Videos (812 pixels in height and 536 pixels in width) were presented in full color in the center of the screen. A 3 cm black separation between dancers was introduced in the together performance clips (1 for each tempo, 6 total). Dancing solo clips of the two participants were created by combining the two individual performances (1 for each tempo, 6 total). To double the total amount of stimuli, the original position of the dancers was flipped so that on one trial, a dancer appeared to the left and, in an identical trial, appeared to the right. Overall, the stimuli set comprised 12 videos of together improvising and 12 videos of solo improvising. Music was removed in post-processing. All videos were edited in the CapCut video editor software (version 2.6.0.; ByteDance).

To obtain a measure of the degree of coordination between dancers, a time series of the pixel change from frame to frame was obtained. FlowAnalyzer software (Barbosa & Vatikiotis-Bateson, 2016) was used to extract motion from 2D video sequences through Optical Flow Analysis (OFA). OFA is based on a computer vision algorithm called Optical Flow that computes pixel displacements between consecutive frames in the video. FlowAnalyzer allows the definition of rectangular regions of interest (ROIs) and disinterest (RODs). All velocity vectors within each region of interest are combined into a single measure. The ROIs of our stimuli are presented in Figure 1. Cross-correlation analyses were performed on the resulting data on Python with a time window of 500 msec and yielded the following average values (see Table 1; see Figure 3).

Figure 1: Stimuli presentation and ROIs selection on FlowAnalyzer.

Table 1: Average Cross-Correlations Coefficient.

	Together Improvising		Solo Improvising	
	Dyad A	Dyad B	Dyad A	Dyad B
85 bpm	.221	.263	.173	.107
120 bpm	.183	.214	.170	.098
155 bpm	.235	.167	.126	.092

The experimental task consisted of a force-choice paradigm and was presented on Psychopy on a 17'' Lenovo Laptop positioned 50 cm away from the participant. On each trial, participants saw a clip and were asked to respond by pressing the left arrow on the keyboard if they thought the two dancers were dancing separately and the right arrow if they thought they were dancing together. Participants completed 48 experimental trials (2 blocks of fully randomized trials) (see Figure 2). Trials ended after 7 seconds or when the participant made a decision.

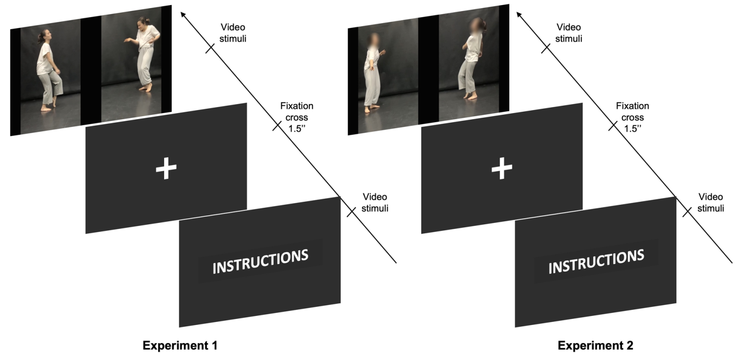

Figure 2: Trial layout of Experiment 1 and Experiment 2.

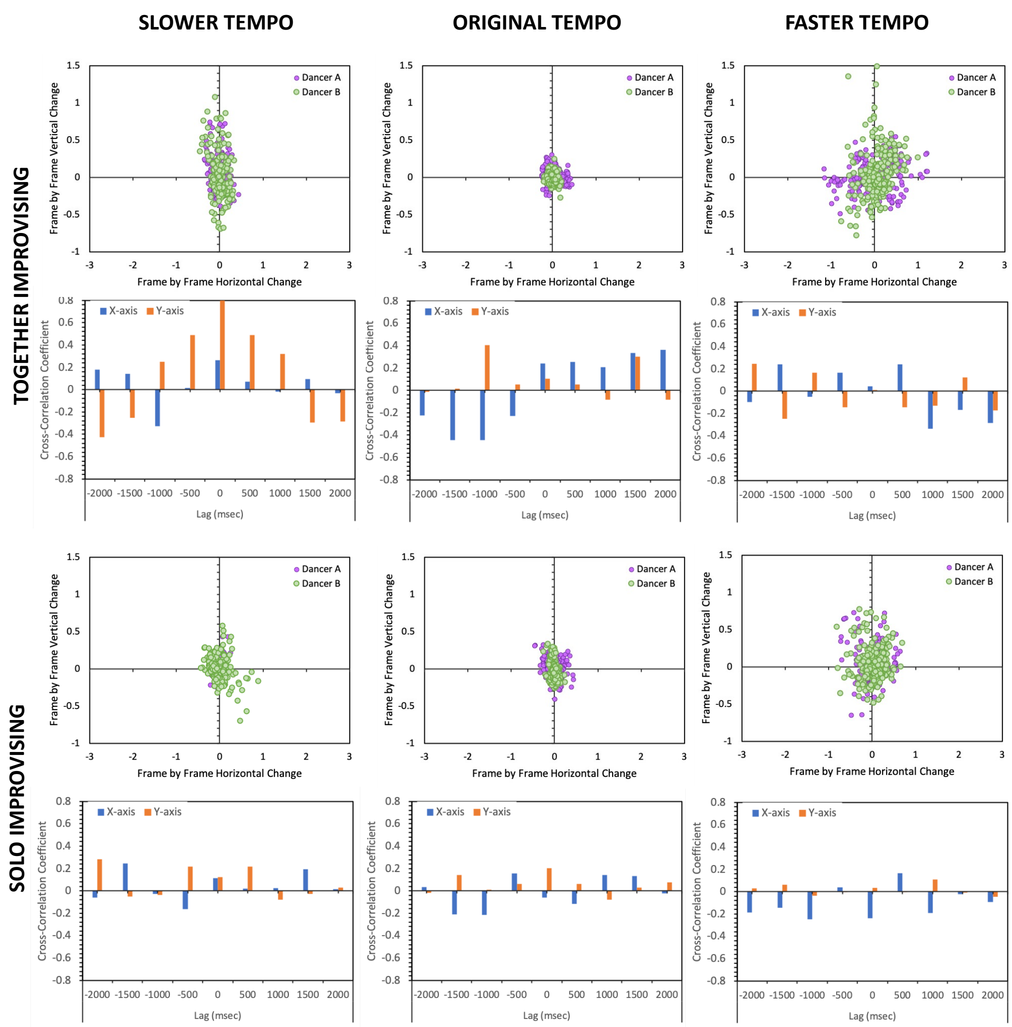

Figure 3: Data obtained from Dyad B for both conditions and the three tempos. Scatter plots show the change in frames in the X and Y direction for both dancers. Histograms show correlation coefficients between dancers calculated for the change in pixels in the X and Y dimensions in time windows of 500 msecs.

Results Participants' accuracy was submitted to a repeated measures ANOVA with condition (solo improvising vs. together improvising) and tempo (120, 155, and 85) as within factors. Results showed a significant main effect for condition, $F(1, 22) = 15.46$, $p<.001$, $\eta_p^2 =.41$. Participants were more accurate when watching the together improvising condition ($M= 0.79$, $SD= 0.18$) in comparison to the solo improvising condition ($M= 0.67$, $SD= 0.13$). No significant effect was found for tempo, $F(2, 44) = 0.21$, $p = 0.81$, $\eta_p^2 = 0.009$ nor for the condition and tempo interaction, $F(2, 44) = 0.27$, $p = 0.75$, $\eta_p^2 = 0.012$.

Participants' Reaction Times (RT) were submitted to a repeated measures ANOVA with condition and tempo as within factors. Results showed a significant main effect for condition, $F(1, 22) = 17.00$, $p < .001$, $\eta_p^2 = .44$. Participants were faster in responding to the together improvising condition ($M= 4.04$ sec., $SD= 0.95$ sec.) in comparison to the solo improvising condition ($M= 4.61$ sec., $SD= 1.00$ sec.). The interaction between condition and tempo was also significant, $F(2, 44) = 3.68$, $p < 0.033$, $\eta_p^2 = .14$. No significant effect was found for tempo, $F(2, 44) = 1.17$, $p = 0.32$, $\eta_p^2 =0.051$.

Tukey HSD post hoc comparisons indicated that for stimuli for the 85 bpm tempo, participants' decision-making was significantly faster for the together improvising condition (together improvising: $M = 3.98$ sec., $SD = 1.03$ sec., solo improvising: $M = 4.87$ sec., $SD = 1.20$ sec.) [$t(22) = 5.26$, $p<.001$].

Experiment 2

This experiment aimed to test whether participants could identify if two people in a video were improvising a dance together or alone when no facial information was made available (i.e., gaze direction, facial expressions). We expected that detection rates would drop compared to Experiment 1 and effect sizes would decrease.

Participants 23 females (*mean* age = 23.58 yrs.; *SD* age = 3.30 yrs.) participated in this experiment. The mean rating of their ability as dancers was 2.52 ($SD = 1.04$), and the mean number of professional dance performances attended in the last year was 1.58 ($SD = 4.12$).

Materials and Procedure The procedure and materials are similar to those followed in Experiment 1. The only difference was that the dancer's faces were blurred throughout the video clips (see Figure 2 and sample here**)**.

Results Accuracy was submitted to a repeated measures ANOVA with condition and tempo as within factors. Results showed a significant main effect for condition, $F(1, 22) = 4.94$, $p = 0.04$, $\eta_p^2 = .19$. Participants were more accurate when detecting the together improvising condition ($M= 0.64$, $SD= 0.12$) compared to the individual improvising condition ($M= 0.56$, $SD= 0.11$). No significant effect was found for tempo, $F(2, 44) = 1.17$, $p = 0.319$, $\eta_p^2 = 0.051$, nor for the condition-tempo interaction, $F(2, 44) = 0.761$, $p = 0.473$, $\eta_p^2 = 0.033$.

Participants' RTs were submitted to a repeated measures ANOVA with condition and tempo as within factors. Results showed a significant main effect for tempo, $F(2, 44) = 4.21$, $p < 0.02$, $\eta_p^2 = .16$. Participants were faster in responding for the 155 tempo stimuli ($M= 3.94$ sec., $SD= 1.30$ sec.) compared to the 120 tempo ($M= 4.20$ sec., $SD= 1.36$ sec.), and the 85 tempo ($M= 3.97$ sec., $SD= 1.39$ sec.). The interaction between condition and tempo was also significant, $F(2, 44) = 3.67$, $p< 0.034$, $\eta_p^2 = .14$. Tukey HSD post hoc comparisons indicated that for the together improvising condition participants decision-making was significantly faster at the 155 bpm ($M = 3.86$ sec., $SD = 1.27$ sec.) compared to the 120 bmp tempo ($M = 4.47$ sec., $SD = 1.55$ sec.) [$t(22) = 4.26$, $p < .004$]. No significant main effect was found for condition, $F(1, 22) = 3.93$, p = 0.06, $\eta_p^2 =0.152$.

General Discussion

The current study investigated the ability to detect whether two people are improvising together and how this ability is affected by available information about the actor's gaze direction and facial expressions. As expected, results showed that participants can perceive when two individuals improvise together. Participants' accuracy and reaction times varied depending on whether they watched a dyad dancing together or dancing solo. They were more accurate, and their ability to make that decision was faster when watching the together than solo performance.

These findings have several implications for our understanding of how we perceive social actions. In line with previous work (Riley et al., 2011), analyses showed that dancers' performances were more coordinated when they danced together compared to when they danced solo. Thus, participants were likely sensitive to the differences in synchronization between the solo and the together performances.

The results of Experiment 2 replicated the findings of Experiment 1. Detection rates were overall lower, suggesting that participants use gaze direction and facial expressions as cues to the joint nature of a performance. However, they can still make a successful determination in its absence. This finding is consistent with those observed using the mirror paradigm (Hart et al., 2014); observers likely take into the similarity in the type of body movements produced by the dancers in their decision-making. Our results also support the notion that during the observation of synchronized actions, people formulate beliefs about the source of interpersonal synchrony (Bernieri et al., 1994). Synchrony can signal to observers that agents might be acting together (Lakens & Stel, 2011). Though participants cannot rely on prior knowledge to establish shared intentionality between actors, synchrony provides a background over which it can be projected. Further studies with perhaps more robust baseline measures of coordination are required to determine what type of coordination (e.g., gaze vs. head movements or torso movements) provides more reliable information to participants.

Coordination values were also sensitive to changes in the tempo of the music, with the slower tempo allowing dancers to coordinate their movements to a higher degree than faster tempos. Changes in velocity-related synchrony are consistent with McEllin et al.'s findings (2020); this suggests that slower movements at similar speeds might provide information about the togetherness of the performance. In Experiment 2, the tempo of the clips did not affect detection success, but it did impact the speed at which decisions were made with faster tempos, leading to faster reaction times. Participants' decision-making took, on average, around 4.5 seconds, suggesting that, at a minimum, several beats of movement were observed before deciding. Decisions were faster when watching a together performance; to some extent, this might be due to the nature of the task. Future studies should include a second task where participants are asked to determine whether the performances were recorded independently or jointly. These findings would likely be replicated if participants had access to the sound accompanying the videos. However, this is an open question; musical accompaniment might provide a background for soft assembly and bias participants towards falsely detecting a together performance when two solo performances are combined.

Finally, though this study focused on novice observers—not privy to the type of contemporary dance vocabulary the dancers performed—they were nevertheless able to successfully perceive them as dancing together or separately. It is likely that expert observers, such as other dancers from the same company, might show even higher detection rates and faster overall reaction times. It is an open question whether dancers trained in other dancing styles, such as ballet, would show similar improvements.

Our findings address a seldom researched aspect of social interactions: how we perceive joint spontaneous action. While the coordination observed is critical to decision-making, access to facial information might provide a cue to the presence of shared intentionality. These findings provide valuable information toward our understanding of the cognitive abilities that make it possible for novice individuals to dance through the challenges of living in a social world.

References

Barbosa, A. V., & Vatikiotis-Bateson, E. (2016). *FlowAnalyzer: Measuring behavioral motion from video*.

Bernieri, F., Davis, J., & Rosenthal, C. R. (1994). Interactional Synchrony and Rapport: Measuring Synchrony in Displays Devoid of Sound and Facial Affect. *Personality and Social Psychology Review*, *20*(3), 303–311.

Bishop, L., Cancino-Chacón, C., & Goebl, W. (2019). Eye gaze as a means of giving and seeking information during musical interaction. *Consciousness and Cognition*, *68*, 73–96. https://doi.org/10.1016/j.concog.2019.01.002

Böckler, A., Knoblich, G., & Sebanz, N. (2011). Observing shared attention modulates gaze following. *Cognition*, *120*(2),292–298. https://doi.org/10.1016/j.cognition.2011.05.002

Brinck, I. (2016). *Joint improvisation in the arts practices: Entrainment, engagement, and expert skill*.

ByteDance. (2023). CapCut [Computer Software].

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: An FMRI study with expert dancers. *Cerebral Cortex (New York, N.Y.: 1991)*, *15*(8), 1243–1249. https://doi.org/10.1093/cercor/bhi007

Cross, E. (2011). The impact of aesthetic evaluation and physical ability on dance perception. *Frontiers in Human Neuroscience*, *5*. https://doi.org/10.3389/fnhum.2011.00102

Cross, E. S., Hamilton, A. F. de C., & Grafton, S. T. (2006). Building a motor simulation de novo: Observation of dance by dancers. *NeuroImage*, *31*(3), 1257–1267. https://doi.org/10.1016/j.neuroimage.2006.01.033

Evola, V., & Skubisz, J. (2019). Coordinated Collaboration and Nonverbal Social Interactions: A Formal and Functional Analysis of Gaze, Gestures, and Other Body Movements in a Contemporary Dance Improvisation Performance. *Journal of Nonverbal Behavior*, *43*(4), 451–479. https://doi.org/10.1007/s10919-019-00313-2

Evola, V., Skubisz, J., & Fernandes, C. (2015). *The Role of Eye Gaze and Body Movements in Turn-Taking during a Contemporary Dance Improvisation*.

Hart, Y., Noy, L., Feniger-Schaal, R., Mayo, A. E., & Alon, U. (2014). Individuality and Togetherness in Joint Improvised Motion. *PLoS ONE*, *9*(2), e87213. https://doi.org/10.1371/journal.pone.0087213

Huang, C.-M., Andrist, S., Sauppé, A., & Mutlu, B. (2015). Using gaze patterns to predict task intent in collaboration. *Frontiers in Psychology*, *6*. https://www.frontiersin.org/articles/10.3389/fpsyg.2015.01049

Knoblich, G. (2011). Psychological research on joint action: Theory and Data. *The Psychology of Learning and Motivation*, *54*, 59–101.

Knoblich, G., & Sebanz, N. (2006). The social nature of perception and action. *Current Directions in Psychological Science*, *15*, 99–104.

Lakens, & Stel. (2011). If they move in sync, they must feel in sync: Movement synchrony leads to attributions of rapport and entitativity. *Social Cognition*, *29*(11), 1–14.

Loula, F., Prasad, S., Harber, K., & Shiffrar, M. (2005). Recognizing people from their movement. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(1), 210–220. https://doi.org/10.1037/0096-1523.31.1.210

McEllin, L., Knoblich, G., & Sebanz, N. (2020). Synchronicities that shape the perception of joint action. *Scientific Reports*, *10*(1), 15554. https://doi.org/10.1038/s41598-020-72729-6

Noy, L., Dekel, E., & Alon, U. (2011). The mirror game as a paradigm for studying the dynamics of two people improvising motion together. *Proceedings of the National Academy of Sciences*, *108*(52), 20947–20952. https://doi.org/10.1073/pnas.1108155108

Ramenzoni, V. C., Riley, M. A., Shockley, K., & Baker, A. A. (2012). Interpersonal and intrapersonal coordinative modes for joint and single task performance. *Human Movement Science*, *31*(5), 1253–1267. https://doi.org/10.1016/j.humov.2011.12.004

Reddish, P., Fischer, R., & Bulbulia, J. (2013). Let's Dance Together: Synchrony, Shared Intentionality and Cooperation. *PLoS ONE*, *8*(8), e71182. https://doi.org/10.1371/journal.pone.0071182

Repp, B. H., & Knoblich, G. (2004). Perceiving action identity: How pianists recognize their own performances. *Psychological Science*, *15*, 604–609. https://doi.org/10.1111/j.0956-7976.2004.00727.x

Riley, M. A., Richardson, M. J., Shockley, K., & Ramenzoni, V. C. (2011). Interpersonal Synergies. *Front. Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00038

Saint-Germier, P., Paternotte, C., & Canonne, C. (2021). Joint Improvisation, Minimalism and Pluralism about Joint action. *Journal of Social Ontology*, *7*(1), 97–118. https://doi.org/10.1515/jso-2020-0068

Sebanz, N., & Knoblich, G. (2021). Progress in Joint-Action Research. *Current Directions in Psychological Science*, *30*(2), 138–143. https://doi.org/10.1177/0963721420984425

Stephenson, L. J., Edwards, S. G., & Bayliss, A. P. (2021). From Gaze Perception to Social Cognition: The Shared-Attention System. *Perspectives on Psychological Science*, *16*(3), 553–576. https://doi.org/10.1177/1745691620953773

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behav Brain Sci*, *28*(5), 675–691. https://doi.org/10.1017/S0140525X05000129

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

Vesper, C., & Sebanz, N. (2016). Acting together: Representations and coordination processes. In *Shared representations: Sensorimotor foundations of social life* (pp. 216–235). Cambridge University Press. https://doi.org/10.1017/CBO9781107279353.012

Vicary, S., Sperling, M., von Zimmermann, J., Richardson, D. C., & Orgs, G. (2017). Joint action aesthetics. *PLOS ONE*, *12*(7), e0180101. https://doi.org/10.1371/journal.pone.0180101

Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psych Bull*, *131*, 460–473.

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