



HAL
open science

Bridging the gap between emotion and joint action

Marta Bienkiewicz, Andrii P Smykovskyi, Temitayo Olugbade, Stefan Janaqi,
Antonio Camurri, Nadia Bianchi-Berthouze, Mårten Björkman, Benoît G.
Bardy

► **To cite this version:**

Marta Bienkiewicz, Andrii P Smykovskyi, Temitayo Olugbade, Stefan Janaqi, Antonio Camurri, et al.. Bridging the gap between emotion and joint action. *Neuroscience and Biobehavioral Reviews*, 2021, 131, pp.806 - 833. 10.1016/j.neubiorev.2021.08.014 . hal-03383388

HAL Id: hal-03383388

<https://hal.umontpellier.fr/hal-03383388v1>

Submitted on 18 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Bridging the gap between emotion and joint action

Marta M.N. Bienkiewicz^{a,*}, Andrii P. Smykovskiy^a, Temitayo Olugbade^b, Stefan Janaqi^a, Antonio Camurri^c, Nadia Bianchi-Berthouze^b, Mårten Björkman^d, Benoît G. Bardy^{a,*}

^a EuroMov Digital Health in Motion, Univ. Montpellier IMT Mines Ales, Montpellier, France

^b UCLIC, University College London, UK

^c Casa Paganini-InfoMus, DIBRIS, Univ. Genoa, Genoa, Italy

^d KTH, KTH Royal Institute of Technology, Sweden

ARTICLE INFO

Keywords:

Socio-Motor interaction
Joint action
Cooperation
Emotion
Affective computing
HCI
HRI
Synchronization
Coupling
Machine learning
Artificial intelligence
Multiple timescales
Multi-modal propagation
Models of human behavior

ABSTRACT

Our daily human life is filled with a myriad of joint action moments, be it children playing, adults working together (i.e., team sports), or strangers navigating through a crowd. Joint action brings individuals (and embodiment of their emotions) together, in space and in time. Yet little is known about how individual emotions propagate through embodied presence in a group, and how joint action changes individual emotion. In fact, the multi-agent component is largely missing from neuroscience-based approaches to emotion, and reversely joint action research has not found a way yet to include emotion as one of the key parameters to model socio-motor interaction. In this review, we first identify the gap and then stockpile evidence showing strong entanglement between emotion and acting together from various branches of sciences. We propose an integrative approach to bridge the gap, highlight five research avenues to do so in behavioral neuroscience and digital sciences, and address some of the key challenges in the area faced by modern societies.

1. General introduction

We thrive on being surrounded by others; we wither when isolated (Baumeister and Leary, 1995). The quantity and the quality of our social interactions are one of the most robust predictors to both how well and how long we live, beating the predictive power of exercise or obesity (Holt-Lunstad et al., 2010). Our brains have been carved by evolution to act together with others towards long-term mutual goals, by emergence of ‘self’-transcendental emotions as opposed to immediate and egoistic benefits (Stellar et al., 2017). These emotions (i.e., compassion, gratitude, or awe) exclusively promote coalitional behavior such as caretaking, cooperation and coordination. For these reasons, moving in unison with others is the firmest of social ties, a superglue, pushing us as a group towards more ambitious goals and performance outcomes as opposed to acting as individual units (Duranton and Gaunet, 2016; Hasson et al., 2012; Marsh et al., 2009; Salmela and Nagatsu, 2017; von Zimmermann and Richardson, 2016). Army drills based on marching together bring a feeling of affiliation and facilitate group performance (McNeill, 1995), just like the Haka dance in a rugby match pumps up the morale before

meeting the enemy and perceptually diminishes the strength of the rival (Clément, 2017; Fessler and Holbrook, 2014). We appraise being a member of a larger group (‘tribe’) by aligning our actions with others (Tsai et al., 2011) to show we like them, we care about them and we are ready to work with them (Mogan et al., 2017; Parkinson, 2020). According to anthropological and behavioral research, a group that chants and dances well together also hunts well (von Zimmermann and Richardson, 2016). Throughout centuries, political and religious power holders have engaged crowds during rallies by using repetitive gestures or vocal expressions (Heinskou and Liebster, 2016; Lukes, 1975). Such rituals (for instance the Nazi salute during May Day rally, cf. Allert, 2009) put normative pressure on individuals with the purpose of bringing up a certain collective thought and feeling, captured by a classical, yet poorly understood in neuroscience, sociological concept of ‘collective effervescence’ (Liebster, 2019; Pickering, 2020). The social morphology of being in a crowd is viewed as essential for motor and emotional synchrony through entrainment of rituals (Borch, 2015; Collins, 2004) with the emergence of contagion hot spots (‘transmitters’) being the essence of collective effervescence (Liebster, 2019; Zheng et al., 2020).

* Corresponding authors at: EuroMov, Université de Montpellier, 700 Avenue du Pic Saint Loup, 34090, Montpellier, France.

E-mail addresses: marta.bienkiewicz@umontpellier.fr (M.M.N. Bienkiewicz), benoit.bardy@umontpellier.fr (B.G. Bardy).

<https://doi.org/10.1016/j.neubiorev.2021.08.014>

Received 16 April 2021; Received in revised form 8 August 2021; Accepted 13 August 2021

Available online 18 August 2021

0149-7634/© 2021 The Authors.

Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Surprisingly, although our social interactions are highly intertwined with emotions we convey or receive, emotions and joint actions are primarily analyzed and modeled by different branches of science, and are usually described separately from each other (Salmela and Nagatsu, 2017). Most models of emotion are individualistic and do not explicitly consider the social context of interacting with others, recently identified as the ‘dark matter’ of modern neuroscience (Schilbach et al., 2013). Exceptions exist in research on empathy (e.g., the model of therapist-patient relation by Koole and Tschacher, 2016), and in research dedicated to movement expression and propagation in arts, such as musical ensembles or dance performance (e.g., Alborno et al., 2015; Basso et al., 2021; Camurri et al., 2011; Chabin et al., 2020; Jola et al., 2013). However, the “acting together” component has not yet been tackled (Butler, 2017; Goldenberg et al., 2016; Mayo and Gordon, 2020).

Reciprocally, human synchronization models, which dominate the joint action research, are not usually inclusive of the emotional state of agents and the manifestations of emotions or other affective components in the way agents move to achieve an optimal outcome (Wallot et al., 2016a; Wolpert et al., 2003). This urgency to address emotional social interaction aspects was recently recognized by Shamay-Tsoory et al. (2019) in their social alignment model, which incorporates the emotional component of group behavior. The scarcity of research is quite surprising, taking into consideration that both self-transcendent emotions and cooperation evolved together as strong features of humanity and contributed greatly to human dominance in the animal kingdom (McNeill, 1995).

In this paper, we zoom into a large body of literature in order to give a synthetic review of the current state of the art on emotion and on joint action, currently two separate strands of research¹. We present evidence showing how these fields are intertwined, and in need of marriage to create a more interdisciplinary and ecologically relevant branch of science. Understanding how we feel impacts the way we act together, and how we act together impacts our emotions, is a societal and scientific challenge of the utmost importance (Barrett, 2017a; Feldman Barrett and Finlay, 2018; Salmela and Nagatsu, 2017; Shamay-Tsoory et al., 2019; Wallot et al., 2016b; Wolpert et al., 2003). As posited by Salmela and Nagatsu (2017), we argue that the emotional component can reveal a lot about the prediction of actions of others, but also about the unfolding of joint acts and their outcomes. We propose a pathway for updates of current models on joint action, such as synchronization, that could be a start to a more integrative approach. We intend to encourage the scientific community to join this conversation, by prompting potential research questions, revise current models of emotion and joint action across neuroscience, computer sciences and social sciences, and to move towards more integrative, social approaches (e.g., Hasson et al., 2012; Hoemann and Feldman Barrett, 2019).

Throughout our review, and in the future directions Section 5, we show evidence that witnesses how the stem of this new research avenue is now shaping the future of human machine interfaces (e.g., robotics, interactive art systems, embodied social media). We believe this new avenue will lend itself to informing occupational health; in promoting efficient and human-friendly working environments and workflows (be it digital or physical) - on a micro-scale, but also crowd management during sport, public and emergency events - on a macro-scale.

2. Joint action: humans act together

2.1. What is joint action?

Joint action can be regarded as any form of socio-motor interaction

¹ We ran a Google search and a PubMed search exploiting core keywords linked to emotion and acting together (for the exact terminology used for the search dated 18/05/2020 and updated 04/04/2021 please see Appendix A) and followed up on cross-references.

whereby two or more individuals coordinate their actions in space and time to bring about a change in the environment. Joint action depends on the following mechanisms: joint attention, representations of others, action prediction and coordination, as well as awareness of oneself and the outcome of the actions of others (Sebanz et al., 2006). Acting together can be emergent or planned and it encompasses the level of intentions, action plans and movements (Knoblich et al., 2011). There is a great variety of terms used by scientific communities in reference to the phenomenon itself: joint action (Sebanz et al., 2006), interpersonal coordination (Mayo and Gordon, 2020; Vicaria and Dickens, 2016), interpersonal adaptation (Burgoon et al., 1995), nonverbal adaptation (Bodie et al., 2016) or even social interactions (Hasson and Frith, 2016). Furthermore, scholars have often offered narrow typologies for joint action subclasses such as physiological or behavioral synchrony and behavioral matching (Mayo and Gordon, 2020), interactional synchrony and behavioral matching (Bernieri and Rosenthal, 1991), mimicry and imitation (Chartrand and Bargh, 1999). Differences in the terminology used by researchers usually stem from stressing and developing models of a particular aspect of socio-motor interaction. For example, Jarrassé et al. (2012) distinguished divisible, interactive and antagonistic tasks, focusing on the objectives and roles each interactant follows (see Table 1). Similarly, Clark (1996) and Toma (2014) both accounted for the temporal relationship of interactions and concentrated on perceptuo-motor aspects of interaction during communication. Burgoon et al. (1995) proposed a more holistic framework of different dimensions (e.g., directedness, timing, measure of behavioral change, intentionality, etc.) to study dyadic adaptation (see Fig. 1).

What remains uncertain in this classification is the required number of characteristics that need to be satisfied for interaction to be regarded as joint action, and its sub-level type, and whether the individual change occurs prior, during or after the interaction (Burgoon et al., 1995). Importantly, joint action does not automatically imply cooperation; it can also indicate competition in terms of individual performance within the group, depending on whether action goal is driven by individual gain (be better as a unit than others) versus collective gain (be better as a group than others) (Tuomela, 2011), leading to multiple possible patterns of coordinated dyadic and group joint action. Table 1 presents an overview of the major sub-types of socio-motor interaction discussed further in this review.

2.2. Models of human group synchronization

As a specific branch of joint action research, the synchronization of a group of agents - such as humans and other animals - or by robotic or digital agents, all underlying the achievement of a common goal, is a robust example of dismissing the emotion component during socio-motor interaction. The state of the art in this domain spans over several scientific disciplines including ethology (Couzin et al., 2005) cognitive and movement neurosciences (e.g., Alderisio et al., 2016), robotics (Iqbal and Riek, 2019) and various branches of mathematics and physics (e.g., Ott and Antonsen, 2017; Strogatz, 2004). Synchronization phenomena have been investigated between individuals, ranging from $N = 2$ (e.g., Noy et al., 2011, to 7–10, e.g., Alderisio et al., 2017) to $N > 10$, such as in human crowds, (e.g., Gallup et al., 2012; Rio et al., 2018). Simply stated, synchronization from a physical principle requires two conditions to be met, (1) a certain behavioral proximity of the systems to be synchronized, such as a common movement amplitude or frequency, and (2) a coupling function between them, through for instance informational exchanges. Typically, metrics for synchronization built on frequency (e.g., the relation between individual and group frequencies), phase (e.g., the order parameter of the synchronization), and their stability characteristics (e.g., Pikovsky et al., 2002), allow us to capture socio-motor coordination characterized by periods of synchronization and desynchronization (Feniger-Schaal et al., 2018; Mayo and Gordon, 2020; Noy et al., 2011).

The generalization of synchronization principles to situations

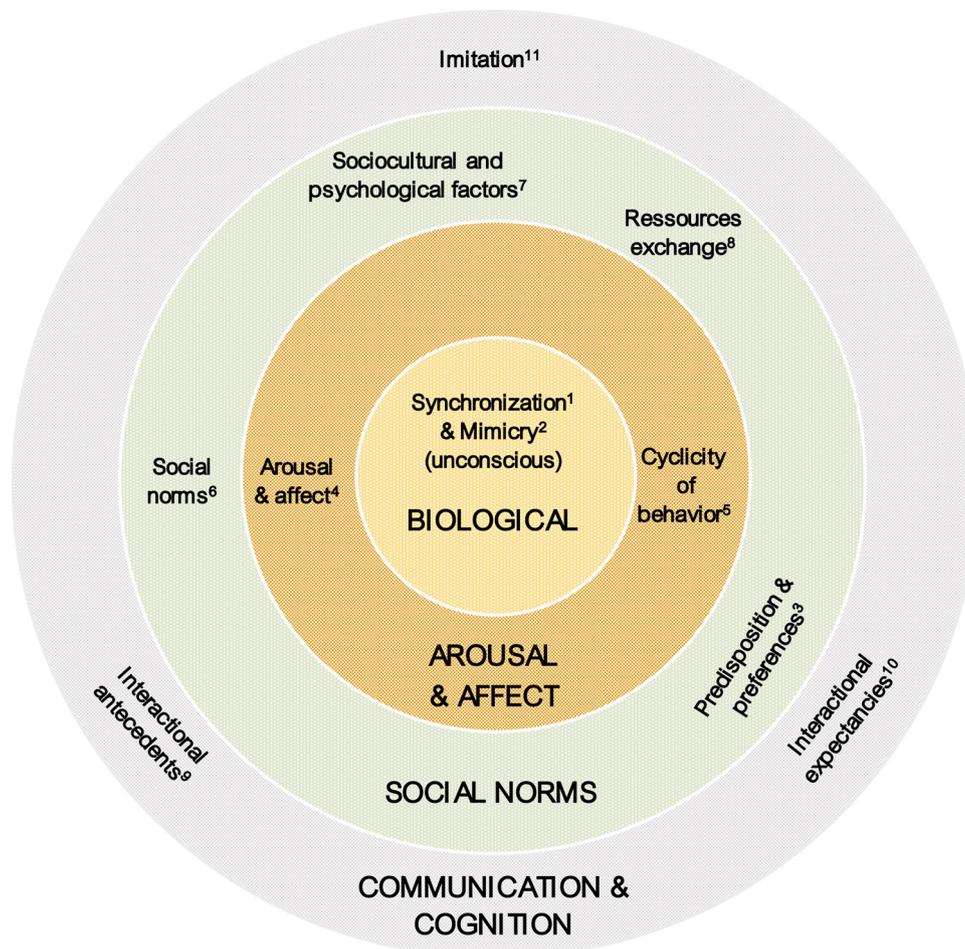


Fig. 1. Levels of joint action (after Burgoon et al., 1995); 1. (Bernieri and Rosenthal, 1991); 2. (Chartrand and Bargh, 1999); 3. (Argyle and Cook, 1976); 4. (Cappella and Greene, 1982); 5. (Altman et al., 1981); 6. (Gouldner, 1960); 7. (Giles et al., 1973); 8. (Roloff et al., 1988); 9. (Patterson, 1983); 10. (Hale and Burgoon, 1984); 11. (Bavelas et al., 1986).

involving more than two agents remains a very recent enterprise². In a nutshell, models of human synchronized behaviors can be categorized in top-down estimation models and in bottom-up self-organized models. In the first category (e.g., Takagi et al., 2019), efficient synchronization between connected participants sharing a common goal is ensured by inference of the shared intention from perceived collective information and consequent adaptation of each individual motion planning. The second category concerns models proposing that synchronized motion observed at the collective level emerges from local interactions between nearby individuals. These models, such as the Kuramoto model (e.g., Strogatz, 2004) or the extended HKB model (for a review see Kelso, 2021) aim to decipher how local informational exchanges and motor adaptations contribute to that emergence. For rhythmic biological movements, coupled oscillator dynamic models have begun to explore perceptuo-motor synchronization phenomena in situations where $N > 2$ (e.g., Alderisio et al., 2017; Bardy et al., 2020; Zhang et al., 2019).

It is striking to observe that a single system of differential equations (Kuramoto or extended HKB model) can capture the complexity of our interactions in perceptuo-motor tasks, including leadership properties (Alborno et al., 2015; Aucoeur and Canonne, 2017; Gnecco et al., 2013; Hilt et al., 2019; Varni et al., 2010), paving the way for the

incorporation of yet missing emotional components (expressive qualities of movement). Anticipating this trend, Varni et al. (2019) recently proposed a Kuramoto-based model of entrainment in music performance of an orchestra, with two components (cf. Phillips-Silver and Keller, 2012): temporal and affective. While the first type relates to rigid temporal alignments in synchronization, the second type, initially developed in childhood during socio-motor interactions with primary carers (Barrett et al., 2007), allows for sharing emotional qualities. Thus, in musical ensembles, entrainment is seen as a temporally flexible organization between musicians, with successive periods of synchronization — stronger at the beginning and at the end of each musical phrase compared to the middle portion (Yoshida, 2002) — creating space for the unfolding of expressive performance cues.

The models of perceptuo-motor social synchronization reviewed above are only a sample of a wide state of the art across several fields of research, and additional branches, for instance social contagion models (Farkas et al., 2002; Mann et al., 2013; Ugander et al., 2012), have been omitted. However, it is striking that *none* of these models have addressed the various emotional qualities that are intrinsic to our efficient and socially relevant cooperative actions. Emotional qualities have occasionally been manipulated (Zhao et al., 2015) or measured (Zhang et al., 2016) during dyadic interaction, although not to our knowledge when $N > 2$, and have not been considered in the various formalisms. In Subsection 5.2, we will propose new directions to incorporate these emotional qualities in joint action models of social synchronization.

² In this article, we are not reviewing experiments and models of collective motion in the animal kingdom such as bird flocks, fish schools, or fireflies' synchronized flashes, although some of them have influenced emergent models of human synchronization (see Clark, 1997; Frijda, 2007).

Table 1
Summary of distinct features of different patterns (subtypes) of socio-motor interaction extracted on the basis of literature review in Section 2.

ARCH TERM	SOCIO-MOTOR INTERACTION (AGENTS SHARING THE SAME SPACE AND TIME)			
SUBTERM	MIMICRY 	JOINT ACTION		
		IMITATION 	SYNCHRONY 	COOPERATION 
SUBCLASS	physiological, behavioural	automatic (spontaneous), intentional		interactive, divisible, antagonistic
METRICS	muscle, physiological, neural activity	body position, neural activity	phase, frequency, time in/out of sync, coupling, re-occurrence	action goals, effort, phase, trajectories, response time
SPACE RELATIONSHIP	copy		copy or not, temporal alignment pivotal	connected, but not a copy
TIME RELATIONSHIP	DELAYED		ALIGNED	
PREDICTION/ ANTICIPATION	NO		YES	
AWARENESS	LOW/ SPONTANEOUS	----->		HIGH/ INTENTIONAL
TIMESCALES	SHORTER PHASIC	----->		LONGER/ CHRONIC
SOCIAL CONSEQUENCES	rapport, empathy, emotion recognition	affiliation, learning, empathy, culture	affiliation, interpersonal attractiveness, cohesion	higher order, complex, longer perspective societal goals
EXAMPLES	facial expression; heart rate	body alignment	hand clapping, haka, brain to brain coupling	orchestra, line manufacturing, see-saw
EMERGENCE DURING LIFESPAN				

Note. Selected features (first column on the left) were identified as informative for understanding the subtle differences between different modes (patterns) of socio-motor interaction. Icons at the bottom depict approximations of developmental stages: babies; toddlers; preschool and school age; teens and young adults.

2.3. Neural origins of acting together

Moving together with others, as discussed in the previous section, results from the coupling between dynamical systems sharing the same space at the same time. Any stable and adaptive behavior emerges as the resultant interaction of internal dynamics and the physical and informational structure of the environment in which the dynamical systems evolve (Gibson and James, 1979; Warren, 2006). Agents learn how to act in the most efficient manner, by searching, for instance, for the most stable attractors (Thelen and Smith, 1994), or by exploiting optimal control (Todorov and Jordan, 2002). This internal self-organization of agent systems in response to external demands is fueled by previous experiences and genetic make-up (e.g., Hainsworth, 1989). Perception naturally plays a key role in delivering information to the brain about the environment and enabling us to efficiently maneuver in it. The investigations into further neural underpinnings of human joint action can be divided into two main areas of research that we herein review in turn: (i) the ability to imitate and understand movement intention in others, and (ii) the synchronization of brain activities during multi-agent

scenarios.

2.3.1. Understanding movement intention in others

The cognitive ability to extract meaning from the perceived actions of others shapes how we cooperate and communicate with them (Rizzolatti and Craighero, 2004). Primates and other social animals, such as birds, are hard-wired to copy the behaviors of group mates (Heyes, 2021). Invasive brain recordings in macaque monkeys identified the neural cornerstone of this ability to visuomotor neurons in the premotor cortex V5. Commonly referred to as mirror neurons, they discharge both when a monkey observes another monkey or another human performing an action, as well as when the monkey performs the action itself (i.e., eating a peanut). Direct single cell recordings are rare in human research, but scarce evidence suggests the existence of similar neurons in the human brain. The supplementary motor area and hippocampus were identified as active during both action observation and action execution, such as reaching (Iacoboni et al., 2005). Interestingly, parts of this network were shown to be active only during observation, indicating an inhibitory role for releasing the imitation response, and other areas

discharged only during execution, suggesting a dissection of this network, referred to as the mirror neuron system (MNS) for action observation (perception) and execution (action). Here, it needs to be noted that the mirror neurons represent an undefined percentage of all neurons in the network for action observation and execution (di Pellegrino et al., 1992), and consequently the validity of using the term MNS itself can be contested; Mirror neurons contribute to, but do not dominate, complex neural networks involved in high-order understanding of action (Heyes and Catmur, 2021). There is an emerging consensus that mirror neurons play a key part only in the processing of low-level features of action for recognition and discrimination, such as body movement topography of another agent, but not of higher-level features such as intention or goal reading (Thompson et al., 2019). Ramachandran (2000) already suggested that the human leap in evolution – the creation of unique aspects of human culture such as language and arts – is endowed to the MNS, as it fosters knowledge transfer between individuals through imitation. Very recently, Heyes (2021) pointed out that imitation is shaped by development and culture, as newborns do not spontaneously imitate but learn to do so via sensori-motor learning, which promotes transformation of neurons to mirror neurons. Emotion and the reward-processing circuitry (the medial orbitofrontal cortex/ventromedial prefrontal cortex) were shown to be linked to the observation of being imitated by another (Kühn et al., 2010), an important hint for the composition of the social glue that acting together promotes.

Some authors (Becchio et al., 2012; Centelles et al., 2011) attribute to MNS, in conjugation with the mentalizing network, the ability to distinguish individual movements from those that are socially connected, and to understand the personal relevance of the movements of others (Kourtis et al., 2010). Exaggeration or modulation of kinematics in order to convey (socially relevant) intention is often referred to as sensori-motor communication (Pezzulo et al., 2013). For example, kinematic differences in the proximal and distal parts of body movements provide pivotal information about whether someone is reaching for a cup to take a sip or to pass it on, and allow us to make adequate predictions and react accordingly before the end of the movement (Ansuini et al., 2014; Cavallo et al., 2016; Soriano et al., 2018). Regardless of the not-yet-fully-understood contribution of mirror neurons themselves in the network, the MNS acts like a bridge between first-and-third-person experiences, allowing for replay of cognitive representation, but safe-guarding the correct placement of the self (Gallese et al., 2004; Meltzoff and Decety, 2003).

2.3.2. Synchronization between brains in shared space

The moment one person interacts with another person in a shared space, we can no longer analyze those entities separately, not only in terms of motor output, but also in terms of brain activity (Balconi and Vanutelli, 2017). Eye-to-eye contact engages language production and reception areas inviting social expression and engagement (Hirsch et al., 2017). Over the last two decades, the state of the art in social neurosciences has indeed shown considerable evidence that our brain is entrained by the structure of physical interaction in the same way it is entrained by the activity of another brain when such interaction is of social nature (Hasson et al., 2012). During social interactions our brain activity is coupled with that of others (Hari and Kujala, 2009), by viewing the same content (Hasson et al., 2004; Nummenmaa et al., 2012), by movements (Basso et al., 2021; Dumas et al., 2010), and even by speech (Hasson et al., 2012; Jasmin et al., 2016). Discussed techniques, referred to as ‘hyperscanning’ (as more than one brain is being recorded) were specifically employed to look at dyads imitating meaningless gestures, and they identified the alpha-mu band as critical for the coordination of interpersonal dynamics, with asymmetrical patterns in brain activity reflecting the imitator-model roles (Dumas et al., 2010). In another context, two particular indexes were identified within this band in the centroparietal brain area - ϕ_1 , ϕ_2 ; and were linked to the coordination of individualistic versus cooperative behaviors in dyads, translating into inhibition and enhancement of MNS activity (Tognoli

et al., 2007). Importantly, the multi-person framework of EEG research has also started to address the niche of social affective interaction (Acquadro et al., 2016). For instance, more pronounced brain-to-brain synchrony (measured with EEG) was found in school classmates ($N > 2$) sharing attention and engaging in face-to-face interactions (Dikker et al., 2017). Also, Babiloni et al. (2012) found that people with high empathic disposition in saxophone quartets ($N > 2$) had higher alpha desynchronization in the BA 44/45 Brodmann’s area during the observation of video recordings of their performance as a musical ensemble. We view this and other recent studies (Chabin et al., 2020; Czeszumski et al., 2020) as promising attempts to investigate the neurophysiology of embodied emotions during joint action.

2.4. Social benefits of acting together

Acting together has been shown to have profound psychosocial consequences, with evidence coming from studies looking primarily at dyads and to lesser extent, groups. The story of the interplay of emotion and synchrony starts in the most primal context of interpersonal relationships – the dyad composed of an infant and a mother. Human infants have no capacity to survive on their own and need a primary carer to regulate their physiological balance (allostasis) (Atzil and Gendron, 2017; McEwen, 2000; Van Der Veer, 1996). Emotionally receptive parents cradle their infants on the left side of their body allowing the flow of visual and auditory information to travel directly to the right hemisphere empathy circuits (Malatesta et al., 2019). Multi-modal channels of bidirectional, physiological concordance with caregivers were identified in infants as early as three months old via heartbeat rate, pupil size mimicry, vocal and affective non-verbal expression (Aktar et al., 2020; Feldman et al., 2011; Palumbo et al., 2017). Further, physiological and movement couplings were found to emerge only when the infant is unsettled, resulting in the parent reducing their own arousal to stabilize the infant (Wass et al., 2019). These first experiences of synchrony were identified to be a cornerstone for healthy emotional development (Feldman, 2012). Parental mirroring allows children to learn their own emotional responses, recognize and label them (Atzil and Gendron, 2017; Pratt et al., 2017). Motor synchrony fosters the building of first trust relationships and prosocial behaviors in human development as early as 14 months, when infants are more likely to pick up a toy dropped by a stranger who bounced with them in synchrony a moment before (Cirelli et al., 2014). Later in development, from toddlerhood to teenage years, other aspects of social cognition and self-regulation are carved by play experiences with caregivers and peers using those very foundations (Nijhof et al., 2018; Viana et al., 2020; Williams et al., 2020). Play promotes learning by imitation in the animal kingdom, through discovering how to act with others, and understanding what others feel (Feldman, 2007). This is a bedrock for empathy and social interactions in humans (Donohue et al., 2020; Viana et al., 2020; Xavier et al., 2016), and perhaps is one of the reasons why ability to synchronize movements with others continues to develop until adulthood (Su et al., 2020b).

Moving in unison thus acts as a social glue and reinforces cooperation (Hoehl et al., 2021; Wiltermuth and Heath, 2009), with a sense of affiliation between strangers (Cacioppo et al., 2014; Kragness and Cirelli, 2021) and boost in self-esteem (Lumsden et al., 2014), arising even from very simple movements such as synchronous finger tapping (Hove and Risen, 2009). The social cohesion instigated by this motor synchronization (Lakens and Stel, 2011) is part of a virtuous cycle and improves actual performance on subsequent joint action, by increasing the perceptual sensitivity of agents towards changes in the environment, those related for instance to the movements of others (Valdesolo et al., 2010).

In that way, motor synchronization seems to be a currency for our social likes and dislikes. Interpersonal attractiveness and likeability of an interaction partner is linked to the magnitude of effort in coordinating with them (in terms for instance of the relative phasing of our

body movements). When engaged in synchronization, we try harder with the person who seems happy and non-threatening (Zhao et al., 2020), or who we find attractive (Zhao et al., 2015). If an interaction partner makes a bad first impression, the chances are we will not put in our best to align our actions with them, as demonstrated by Miles et al. (2009) in a study looking at dyadic synchronous stepping. This relation is bidirectional, as moving in unison fosters interpersonal attractiveness. For instance, Cheng et al. (2020) found that during paired walking tasks a phase synchronization time ratio (how much time people walked in phase with each other) was predicted by how much they liked a stranger based on their initial impression. After a period of silent and ‘chatty’ walking, strangers reported increased affiliation with the person as a consequence of synchronous walking. Atherton et al. (2019) reported decrease in prejudice towards another ethnic group after physical and imagined walking.

Even if people are not ‘on the move’ across space, their bodies show gradual convergence towards postural alignment over the course of almost any dyadic interaction (Chartrand and Lakin, 2013; Fujiwara et al., 2019; Paxton and Dale, 2013). In human psychotherapy, which aims at aiding any shortcomings in emotional regulation in adulthood, therapists usually build a ‘trust’ relation (therapeutic alliance) via the above-mentioned elements of speech synchrony, as well as body movements (Bar-Kalifa et al., 2019; Lutz et al., 2020), to access implicit and explicit emotional regulation of the patient (Koole and Tschacher, 2016). For instance, head position synchrony between therapist and patient has been found to be linked to the overall therapy outcome, whereas body alignments were noted on shorter timescales to be predictive for each session outcome (Ramseyer and Tschacher, 2014) as well as experience of pain with a therapist (Goldstein et al., 2020). Weaker coordination patterns of head motion (angular displacement and velocity of the head’s yaw and pitch) were also reported during conversations involving arguments between romantic partners (Hammal et al., 2014) and were proposed to be predictors of the quality of rapport with a therapist (Goldstein et al., 2020). Interestingly, people who live together, and who are hence in permanent dyadic interaction, such as roommates or couples, become alike, over time, in terms of their emotional reactivity and emotional expressions (Anderson et al., 2003). During conversations adults show convergence in use of certain speech elements such as figures and grammar (Hasson et al., 2012). Joint speech research revealed that merging ourselves with others is visible in the neural activation patterns that are different from speech production alone (Jasmin et al., 2016), putting joint speech in the realm of dynamic interplay of socially shared cognition (Cummins, 2014).

In a study on large groups ($N > 5$), von Zimmermann and Richardson (2016) demonstrated that synchronous vocalization with others enhances memory performance and group effort, providing evidence for hidden wisdom of group rituals such as dancing, or singing, or when marching to fight a rival. Similarly, physiological concordance emerging between newly met group members (i.e. intervals between heart beats) explained one sixth of the variance of the performance in a drumming task in another group study (Mayo and Gordon, 2020). Both physiological and motor synchronization levels were predictive for the subjective sensation of group cohesion. Also, a study by Mønster et al. (2016) looked into the synchronization of bio-signals (heart rate and electromyography of face muscles) during cooperation (line-manufacturing of paper boats by triads). Participants were induced emotionally by the researcher acting either warmly or coldly in the interaction with the group. Exchange of smiles was linked to group cohesion (primed with the warm behavior from the experimenter), whereas synchronous increased skin conduction (interpreted as experiencing tension in a group caused by coldness in the demeanor of the experimenter) correlated negatively with measures of cohesion.

In sum, the motor and physiological coupling between humans is hardwired, and develops through childhood, to deploy group affiliation (we are members of the same tribe) and maximize collaborative efforts. Developments in dyad research have provided ample evidence for the

social benefits of interpersonal alignment, with group research still being under-addressed. Therefore, the future of neuroscience of social interactions needs to include not only a second-person perspective (Schilbach et al., 2013) but also a multi-person or multi-agent perspective, and it constitutes the *raison d’être* of our review.

3. Emotions

The previous section synthesized the current state of the art in the main branches of sciences investigating how we act together, from physics-based models of synchronization to social neuroscience in humans and other animals, embodied cognition and developmental psychology. While we demonstrated that emotion is not a dominant focus to understand how people move in a group, we also mentioned some burgeoning research incorporating emotional qualities and evaluating how they affect embodied social interactions. In the current section, we reciprocate with a first analysis of current models of emotion, how they largely ignore joint action and the above models, with some recent exceptions paving the way for a more integrated approach.

3.1. What is an emotion?

Since James’ attempt to answer the question “What is an *emotion*?” (James, 1884), an unresolved yet debate started on the inherently elusive nature of the phenomenon (Scherer, 2005), highlighted by Fehr and Russell’s (1984) remark that “everyone knows what an emotion is, until they are asked to give a definition” (p. 464). Four decades ago, Kleinginna and Kleinginna (1981) found 92 different definitions of emotion and stressed the need for consensus. Despite substantial advancements, there is not a unified theory of emotion that would exhaustively address all the fundamental questions (Reisenzein, 2015). Interestingly, the etymology of the word “emotion” contains in itself “motion” or *movere*, in Latin “to move”, and the word “affect”, more general, which relates to “producing changes” (2020³). Arguably, a straightforward example of emotion as a driving force of change comes from the so-called fight-or-flight response (Cannon, 1953). When scared, we prepare to run away in order to withdraw ourselves from the perceived source of danger, and when angry, we prepare to stand up and fight against the threat (Cannon, 1953; Jansen et al., 1995). Stemming from this perspective, emotional arousal holds a motivational function (Reisenzein, 2015), previously conceptualized as a mode of ‘action readiness’ (‘Ur-affekte’; Kafka, 1950). Emotions allow us to adapt to a given set of circumstances with the aim of survival and enhancing wellbeing (maintaining allostasis), thus they entrain different action tendencies to satisfy different needs (Frijda, 2007; Frijda et al., 1989; Frijda and Parrott, 2011; Ridderinkhof, 2017), weighed up by cognitive processes against individual cost/gain and previous experience (Ferrari, 2014; Kiverstein and Miller, 2015; Padoa-Schioppa, 2011).

Classically, there were two main conceptual frameworks for studying emotion: (i) discrete models, representative of individuals emotions (e.g., anger, joy, fear, etc.), pioneered by Darwin (1872) and later developed, for instance, by Izard and Carroll (1971); and (ii) dimensional models, with a specific positioning on the coordinate system for which an emotion is an interplay between different dimensions (e.g., valence, or arousal) acting as coordinates (Plutchik, 2001; Russell, 1980) or categories linked by smooth gradients (Cowen and Keltner, 2017). Recent models (e.g., Barrett, 2017a, b) propose that emotions are emergent “constructions” of the world based on interoception (signals within milieu of the body), exteroception (signals from the environment) and previous experiences feeding in dynamic (visceromotor, motor and sensory) predictions. In this context, emotions are not

³ (2020). *Definition of emotion* [online]. Oxford University Press. Available at: <https://www.lexico.com/definition/wake> (Accessed: 14 November 2020).

demarcated events, but derivative of the constant interaction between the complex dynamics of the nervous system, the body, and the surrounding environment. We experience the world factually, but it is the visceral reactions that validate the experience as ‘real’ (Duncan and Barrett, 2007). In a Higher-Order Theory of Emotional Consciousness proposed by LeDoux and Brown (2017), extending the high-order theory of human consciousness to emotions, the sense of self is core to emotional experiences. Although the sub-cortical circuitry, such as fear or survival circuits, are crucial for providing inputs for adaptive behavioral responses, they place emphasis on all conscious states (i.e., emotions) being instantiated in the general cortical network of cognition (cortical circuits).

Recent neuroimaging studies have shown that the emotional network is widespread in the human brain, and embraces cognitive areas (such as anterior frontal areas) beyond typical affective ones, such as amygdala, diencephalon and brainstem (Duncan and Barrett, 2007; Kiverstein and Miller, 2015) and motor readiness circuitry - premotor area (BA6) (Costa and Crini, 2011). Among other studies, Jastorff et al. (2015) found that distinct categories of emotion emerged only when looking at multi-voxel activity patterns in fMRI during discrimination of visual and dynamic emotional stimuli of different saturation. During the resting state in the MNS network, four hubs were identified as connecting points in the right hemisphere: anterior insula, right anterior cingulate, right precentral sulcus and right fusiform gyrus. Four points of connection with other structures were identified in the emotional network: right amygdala, right insula, left putamen, and left middle STS. Interestingly, Costa et al. (2014) demonstrated, using EEG recordings, that emotions from different categories overlap spatially in activation patterns with the emotional brain network, but also show distinct temporal signatures (i.e. time to peak). This aligns with Barrett (2017a, 2017b) argumentation that there are no specific pathways for emotion categories, but different neural activation can lead to the same emotion (“many-to-one”) and reversely the same network can give rise to different emotions (“one-to-many”) (i.e., notion of degeneracy, Edelman and Gally, 2001).

3.2. Understanding others’ emotions

Despite the recent progress in emotion research, little is known about the actual dynamic link between emotion and social interaction (Butler, 2017, 2015), and the models briefly reviewed above remain largely silent about the social nature of emotion. In this section, we focus on the “shareability” of emotions, first through the prism of empathy and mechanism of mimicry, before addressing scarce evidence for the existence of group emotion.

3.2.1. Empathy

Broadly, empathy relates to understanding and ‘sharing’ what other people feel, need or want to do (Bloom, 2017; Ferrari, 2014), but does not infer action itself. If we refer to empathy as a parameter in prosocial behavior, we mostly mean empathetic distress — experiencing discomfort caused by perception of distress in others — which is equivalent to emotional (‘hot’) empathy (Bloom, 2017). However, other possibilities of sharing an affective state include cognitive (‘cold’) empathy (conceptually understanding what another person experiences), emotional contagion (i.e. ‘catching’ anxiety because we share the same physical space with someone who is anxious), and finally, compassion leading to helping behavior and altruism (Bloom, 2017; Preston and de Waal, 2002).

The well-known Perception-Action model (PAM) of empathy (Preston and de Waal, 2002) proposes that the perception of another agent’s affective state activates the same neural representation of the observer (without any particular ‘empathy’ center), leading to activation of the same somatic and autonomic responses, an idea that was first conceived by Darwin (1872), later reinforced by Hommel (1998,1997). Research on MNS has revealed that we do internally simulate the actions we

perceive (measured as shorter reaction times in consecutive execution of action performed by another agent) (Craighero et al., 1998). For example, whether we observe a facial expression of disgust, or imitate it, the same parts of the brain activate as during actual experience of disgust (Carr et al., 2003). In the same vein, observing fearful bodily expressions in others activates the motor readiness circuitry (Borgomaneri et al., 2015). Human toddlers show distress when observing others in distress (Zahn-Waxler et al., 1992), and they ‘try on’ emotional reactions they observe in others to see how they feel, without particular need of their own to be broadcasted (Einon and Potegal, 1994). This behavior in adults, e.g., bursting out crying when someone else does, would be considered as pathological or maladaptive, deprived of emotional containment and dissociation between self and others.

The neural underpinnings of embodied empathy, i.e., the imitation of emotional facial expressions, have so far been pinpointed to the right ventral premotor cortex (Leslie et al., 2004), with the inferior parietal lobule identified as a locus attributing a sense of agency for the self and others (Meltzoff and Decety, 2003). The brain activation of visual information containing key action features flows from the superior temporal cortex to the posterior parietal (to simulate action and code kinematics), involves frontal MNS (to identify action goal), and flows back to the superior temporal cortex to inform it about action prediction and imitation plan if needed (Carr et al., 2003). In this connectivity model the insula relays the action representation to the limbic system from the MNS and motor areas. Although simulation is a key response for understanding what others might feel, people are usually not an echo chamber for other people’s feelings. The neural architecture of empathy is complex and sophisticated, endowing a multitude of social interaction scenarios, and is intertwined with cognition (Bernhardt and Singer, 2012; Ferrari, 2014). A double dissociation neurophysiological mechanism was proposed based on lesion studies for (i) cognitive (cold) empathy embedded in the ventromedial prefrontal areas, and for (ii) emotional (hot) empathy rooted in the inferior frontal gyrus (Shamay-Tsoory et al., 2009).

3.2.2. The prism of mimicry

The motor phenomenon of mimicry, or ‘chameleon effect’ (Chartrand and Bargh, 1999), overlapping in some ways but not to be confounded with empathy (‘I felt what you feel’), relates to involuntary mirroring of expressions of our interaction partners (‘I saw and show what you feel’). Mimicry, unlike imitation and synchronization, is unconscious⁴, and at least partially independent of the performance ability in the latter, despite clearly being nested in the same behavioral spectrum with strong functional interconnections (Genschow et al., 2017; Rauchbauer and Grosbras, 2020). This comes with the caveat that mimicry is usually recorded in a naturalistic observation, such as matching up facial expression (rapid facial reaction) during face-to-face contact with another person (Moody et al., 2007), whereas imitation and synchronization are most often elicited and captured in less ecological scenarios (with instruction). Unintentional mimicry of body and vocal expressions unifies emotional state by means of evoking the same internal responses in agents (Hatfield et al., 2011, 1993), and is perhaps the closest to the concept of emotional contagion. Neuroimaging studies (hyperscanning mock setup for fMRI) show a similar time-locked pattern of brain activity between people subjected to watching emotional excerpts from movies (Nummenmaa et al., 2012) and listening to autobiographical stories retold by interaction partner (Smirnov et al., 2019). Adults also align their heart rate variability (Scarpa et al., 2018), pupil diameter, respiration rate, body temperature and electrodermal activity

⁴ While for some scholars e.g., (Barrett et al., 2019b) mimicry is the process of an unconscious copying of other’s postural, facial and other behaviors, for other authors e.g., (Centelles et al., 2011; Clarke et al., 2005; Nackaerts et al., 2012) to mimic is to intentionally imitate the behavior of the other, we use term ‘mimicry’ in this review in the former definition.

by their mere physical presence in one space (see Palumbo et al., 2017) for a comprehensive review on physiological synchrony). Emotional response to others via mimicry is also mediated by oxytocin hormone (Korb et al., 2016) which, in surplus, can increase mimicry of emotional expressions such as demonstrated in a double-blind study on males looking at adult and infant expressions. Similarly, Festante et al. (2020) looked into EEG after intranasal oxytocin administration and found that the enhanced alpha range of the mu rhythm desynchronization had an impact on the sensorimotor circuits involved in social perception and action understanding. Complimentary to this evidence, The Neuro-cognitive Model of Emotional Contagion (Prochazkova and Kret, 2017) proposed multi-modal connections between motor mimicry (facial and body, inclusive of eye synchrony) and autonomic mimicry (physiological synchrony). In this model, both mechanisms are considered separately, but operate on the social interaction interface (cognition), with MNS being an engine for shared emotional arousal and steering empathic behavior. Also, the mimicry model by Wood et al. (2016) proposes a decomposition of the perception-action loop into sensorimotor layers, encompassing functions such as moderation of emotion, prior beliefs, arousal and adaptive behavioral response. Mimicry, even on a fast timescale under 1000 ms (as demonstrated in an EMG study of facial muscles), is intermediated by the persons' own affective state and the environmental context (Hess and Fischer, 2013), as well as by affiliative goals (Rauchbauer et al., 2015). Interfering with facial mimicry (i.e., mouthguards) blocks out emotional recognition of the body and the facial expressions of fear and happiness in others (Borgomaneri et al., 2020; Rychlowska et al., 2014). This effect was reported to be mitigated by individual levels of empathy; people with higher empathy levels rely less on mimicry to recognize emotional states in others, suggesting at least a partial functional independence of empathy from mimicry (Borgomaneri et al., 2020). The social function of mimicry is also associated with longer timescales. Hogeveen et al. (2015) identified that mimicry might increase social attunement for a longer period than initial interaction via increased mu-suppression activity in MNS.

Based on the evidence above, both empathy and mimicry are key players in emotion propagation, interconnected at the functional level and bridged on the neural level by MNS. However, the dynamics of spreading motor, neural and physiological embodiment of emotions between several agents, on short and transient timescales, as well as its impact on joint action, short and long term, remains largely unexplored today. One early exception is the model proposed by Kelly and Barsade (2001), considering group emotion as resulting from individual states of the agents (shared implicitly and explicitly), captured as an affective tone that molds cohesion in the group and fluctuates. For example, adopting a group identity can inflate saliency of negative emotions, such as anger, through linking group-based appraisal to group emotion and prompting pre-designated behavioral response (Kuppens et al., 2013). One important consequence of this line of research is linked to education, as group emotions in classrooms were reported to mediate attention sharing and learning, cornerstone of long-term academic achievements (Eilam, 2019). To understand further the dynamics of 'sharing' emotion between multiple agents as they occupy or move in the same space at the same time, we give an overview of the affective embodiment research on the multi-modal signals that can carry information about emotion qualities of the agents.

3.3. The embodiment and automatic recognition of emotions

The role played by the various layers of the moving body as both receptacles and vehicles of emotional experiences has been largely addressed. Here we briefly synthesize research in psychology, neuroscience and affective computing revealing how emotional qualities emerge from multi-modal inputs (i.e., Arias et al., 2018), from face to whole body and physiology, before showing the unanswered questions at the heart of this review.

3.3.1. Face

As the most ancient and still most dominant area of emotion research concerns faces (Ekman, 1992), it is not surprising that the majority of research efforts in affective computing have been targeted to facial expressions (de Gelder, 2009). For instance, the Facial Action Coding System (FACS) was developed to provide an objective, standardized, and measurable coding system of emotional facial expressions (Ekman and Friesen, 1978). Combination of facial muscle activations (e.g., raised eyebrows, wrinkled nose, tightened lips) are differentiated as Action Units (AU) of micro expressions — i.e., instantaneous facial movements hardly perceived by the naked eye — and are subsequently identifiable as an experienced emotion (Ekman and Rosenberg, 2005). For instance, lip-corner raising is identified as an AU12 that, among others, is associated with joy. Thanks to progress in computational techniques and the increase in size of facial expression datasets, raw data rather than FACS or hybrid machine learning architectures are now used to let the mathematical models identify the relevant muscle patterns. For a detailed overview on the embodiment of emotions in facial expressions please refer to Barrett et al. (2019a, 2019b). For an overview on automatic recognition of facial expression see (Küntzler et al., 2021).

3.3.2. Whole body

A growing body of literature has shown that body expressions are at least as powerful as facial expressions in conveying affect (Atkinson et al., 2004; Hadjikhani and Thilenius, 2005; Wallbott, 1998). Studies have shown that in certain situations or for certain emotional states, the body is more informative than the face. For example, In the case of incongruence between facial and body expressions, studies (Meeren et al., 2005; Van den Stock et al., 2007) show that body posture has a strong influence on the perceived emotion. These findings were also supported by Aviezer et al. (2012). They show that evaluations made on body expressions rather than on facial expressions lead to more accurate assessment of the affective valence of the situation that triggered such expressions. De Gelder (2009) added that the body does not only convey a person's affective emotional state but also her actions and intention in response to it. Further, it should be considered that at close distance people may possibly rely on the face, but at a distance, where the facial expressions are hardly perceived, the body becomes prevalent to understand and express emotions. For an overview of bodily manifestation of emotions, please refer to Kleinsmith and Bianchi-Berthouze (2013), Melzer et al. (2019) and Witkower and Tracy (2019). Unfortunately, there is no equivalent to a FACS for the body. An initial equivalent model, called Body Action Coding system (BACS), was proposed quite recently by Huis In't Veld et al. (2014a, 2014b). They investigated covert muscle activation across various body parts in the context of anger and fear. Due to the lack of formal models from psychology and neuroscience fields, researchers in affective computing have hence turned to other fields for driving the design of automatic body expression recognition models. The four factors of the Laban Notation System (effort, shape, space, direction) (Laban and Ullmann, 1988) have provided the foundation for most of the pioneering work in this direction. A multi-layered approach inspired by Laban's Effort Theory for the automated recognition of emotion in dance performance was proposed by Camurri et al. (2003), through a computational model capturing how different dancers perform the same choreography with different emotions. Speed and energy showed to be correlated with the arousal dimension of the affective states while an extended body was generally associated with more positive states than a closed body posture. De Gelder and Poyo Solanas (2021, p. 1) defined these features as middle level features, and suggested that "behaviorally relevant information from bodies and body expressions is coded at the levels of mid-level features in the brain". A computational framework to model non-verbal emotions was proposed in the MEGA European project (Camurri et al., 2005) and a more recent approach was proposed in Camurri et al. (2016). By taking advantage of advanced machine learning architectures, there is today the tendency to use a more agnostic approach based on the temporal sequence of row

data (e.g., Wang et al., 2021) from movement (e.g., 3D position of each joint, angles between body segments) or muscle activity sensors (e.g., intensity of muscle activity). Still, given the limited size of the datasets and also the complexity of body expressions when embodied in everyday activity, recognition performance may gain by a combination of low- and mid-level features. By being not so directly connected to the specific high-level semantic emotions, mid-level features provide a more functional but still adaptive description of body expressions (de Gelder and Poyo Solanas, 2021) possibly enabling computational recognition systems to be able to generalize across different contextual situations.

3.3.3. Physiology

Physiological changes in relation to affect have been for long investigated. While pioneering work in the area of affective computing had initially leveraged medical devices, technical advances in the low-cost wearable sensing technology area has opened the possibility to seamlessly set up and explore ubiquitous applications (for a survey see Shu et al., 2018). Differently from facial and body expressions, physiological changes are generally used to build systems that automatically infer affective changes along the valence and arousal dimensions. This is due to the lack of clear physiological patterns associated with discrete emotions (for a review see Siegel et al., 2018). Applications for stress and anxiety levels automatic detection are possibly the most investigated areas in the computing domain (e.g., for a survey see Panicker and Gayathri 2019). General approaches in building physiological-based affect recognition models built on general statistical features (e.g., max, mean, std) extracted from the physiological responses to an emotional event. To improve performances, more specific features are extracted for each type of physiological signals. Heart-related physiological activity is possibly the most explored metrics in affective computing beyond galvanic skin conductance, as it appears related to both valence and arousal. Features related to both sympathetic and parasympathetic activities, in both time and frequency domains, have been explored (e.g., Alberdi et al., 2016), for instance the ratio between high and low frequencies. Beyond heart rate and skin conductance, respiration (Cho et al., 2019), skin temperature (e.g., Goulart et al., 2019; Wang et al., 2014), and brain signals (Alarcão and Fonseca, 2019; Torres et al., 2020) have started to gain increasing attention, demonstrating complementary performances. Similarly, research using electromyography (EMG) has shown evidence of muscle tension that is often linked to anxiety (Pluess et al., 2009). In the work by Olugbade et al. (2019), the use of EMG in concomitance with motion capture leads to clear increase in automatic pain level recognition performances in people with chronic pain. This is again thanks to a technology that is more portable and acceptable for everyday use, enabling the extraction of continuous signals, extending the possibility for measuring a large set of statistical features, and in particular features that characterize the variability of these signals (Cho et al., 2019). In a similar way to the work on non-verbal modalities, there is also the tendency to use advanced machine learning techniques that can work directly on raw data or on low-level statistical features extracted continuously over moving windows of the signals (Wang et al., 2021). However, this approach is still challenged by the limited size of the available datasets.

While each modality carries emotion information, studies have shown that multimodal recognition systems tend to lead to better performances (Al Osman and Falk, 2017; D’Mello and Kory, 2015; Poria et al., 2017). Since modalities work at different temporal scales as responses to emotional triggers, how to fuse such modalities has been, and is still, a crucial question in the affective computing community. A variety of fusion approaches have been considered. Solutions have explored low-, mid- and high-level fusions, as well as more complex architectures to fully capture the relationship between such modalities.

In particular, a typical issue in multimodal modeling is that some sensors may only be available during the training phases of the model. This could be due to sensor malfunctioning or sensor availability (e.g., privacy) during deployment. Some of the explored fusion approaches have tackled such problems by learning the relationship between modalities in order to infer the missing ones when the problem occurs (e.g., Cheng et al., 2016; Rivas et al., 2021; Wagner et al., 2011). Transfer learning approaches have also been used to this purpose together with addressing the problem of limited dataset size (for a review, see Feng and Chaspari, 2020).

However, these are not the only critical questions that challenge the affective computing community. Most of the work so far has focused on mapping face, body and physiological features or their combination into emotion semantic concepts. As we move into real-world applications, such approaches are quite limited as affective experience, and its perception, are subjective processes shaped by various factors such as context (Barrett et al., 2019a) and personality (Komulainen et al., 2014). Transfer learning approaches have been used to support the development of models between for example datasets built in the lab and smaller ecological dataset, or to compensate for the limited size of such datasets (Feng and Chaspari, 2020). Other approaches have more specifically attempted to integrate the context directly in the model. For example, the use of hierarchical architectures leveraging automatic human automatic activity recognition as contextual information to body expression recognition have shown to reach better recognition performances and generalization capabilities across a variety of activities (Wang et al., 2021). Such an approach was further supported by the use of graphical algorithms that intrinsically capture body configuration information critical to both the prediction of the activity performed and the emotion expressed by the body. Similarly, Zhao et al. (2019) have explored how personality can be leveraged to improve recognition performances of personalized emotion recognition models. Using a hypergraph learning framework, they captured the relationship between individual personalities and physiological responses to stimuli, showing a clear improvement in recognition rates, and suggested that the next step would be to co-learn the personality scores of participants.

While the work above is supported by an increasing number of multimodal and also multi-factors datasets, e.g., MAHNOB–HCI (Sol-eymani et al., 2012); DEAP (Koelstra et al., 2012); EMOPAIN (Aung et al., 2016), ASCERTAIN (Subramanian et al., 2018), there is a real need for larger and real-life datasets that are more inclusive, and that can capture the variety of (social and activity) contexts and emotional expressions. While personality surely contributes to the experience, response and perception of emotions, there are many other personal factors (e.g., cognitive and physical impairments) that are critical to these processes. Existing datasets are still largely lacking the investigations of the above questions.

3.3.4. Multi-agent embodiment of emotion

One common characteristic of all the studies reviewed above, and of the-state of the art of embodied emotion in general, is that they all, almost exclusively, investigate the embodied manifestation of emotions in one individual in space and time (see Niedenthal, 2007). However, as said in our general introduction, humans are rarely withdrawn from natural interaction with other conspecifics. Particularly challenging is the issue put forward by leaders in emotion research (e.g., Ekman, 1992) that, arguably, the main job of emotions is to facilitate the engagement in perceived appropriate behaviors, in situational encounters with conspecifics and others. Dyadic and group situations are not only natural vectors of emotion diffusion, they are the instances where this diffusion contributes to a successful communication and enhances prosocial behaviors. For instance, Mou et al. (2016) have shown that body behavior

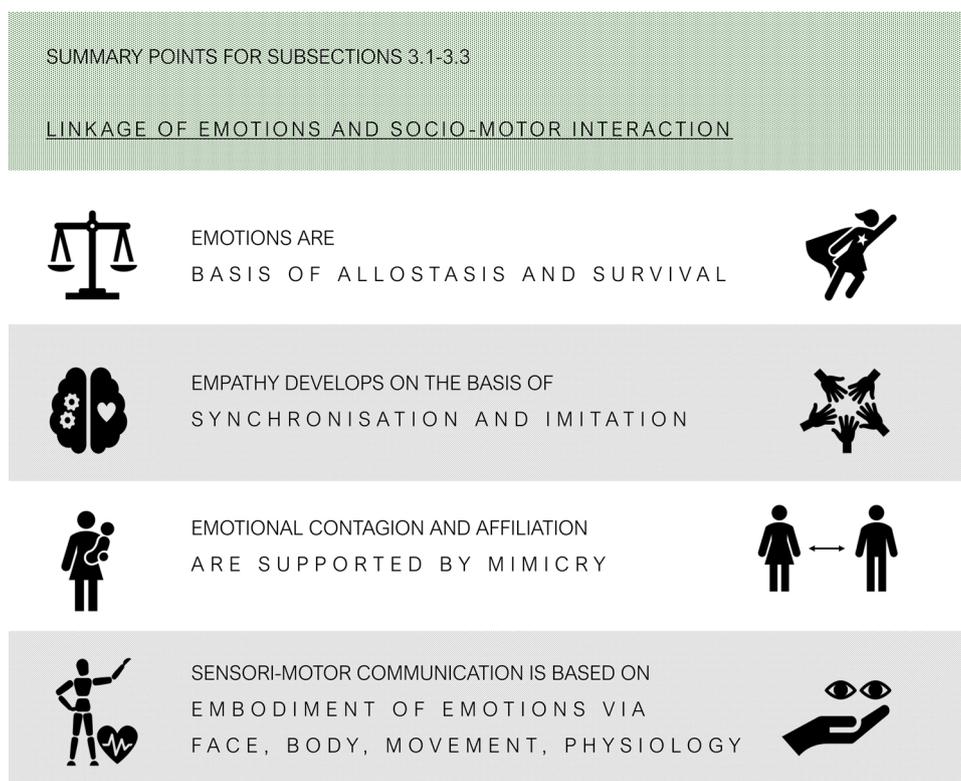


Fig. 2. Graphical summary of scientific findings reviewed in Subsections 3.1–3.3 and referred to in further sections of the manuscript.

is a better predictor of emotion-group membership than facial expression, possibly because of the mirroring that may occur in group interactions. Playing music together is one of the most significant examples of non-verbal human interactive, creative and social activities, and, as music is widely regarded as the medium of emotional expression in full body movement *par excellence*, it is not surprising to witness the first layer of research focusing on emotion transmission in embodied joint action in this domain. For example, [Glowinski et al. \(2013\)](#) compared the expressive movement of violinists when playing solo and in ensemble conditions, and showed that when people perform a task as part of a joint action, their behavior is not the same as it would be if they were performing the same task alone. In the same vein, [Varni et al. \(2010\)](#) showed, in a multi-modal interactive context of a violin duo and a string quartet, that enacting pleasure while playing enhanced movement synchrony, whereas enacting anger reduced it. In another study with a triad of musicians, body sway was structured differently with different levels of emotional expressivity during performance ([Chang et al., 2019](#)). Higher Granger coupling within the triad of musicians (piano, cello and violin) was linked to emotional expressions of happiness when compared to sadness. Finally, the quality of dance performance was found to benefit from synchronized interpersonal movements, a quality that was also enjoyed by the spectators ([Vicary et al., 2017](#)). These examples illustrate the very recent move to understand human emotion in the context of joint action. However, the full picture remains obscure as we still have little to no understanding on how emotion dynamically fluctuates and propagates in multi-agent, naturalistic scenarios (where emotion brews as a consequence of interaction between agents and environment, e.g., [Dotov et al., 2021](#)). Fig. 2 represents summary points from the Sections 3.1–3.3 showing how emotion is intertwined with acting together.

3.4. Linkage between joint action and emotion in socio-motor interaction deficits

Earlier (in Section 2.4), we presented evidence showing how moving in synchrony with others can bring positive emotions such as affiliation and attractiveness. The impact of motor behavior and its shaping role for the emotions experienced by individual agents has also been brought to the spotlight by researchers interested in mental health and wellbeing ([Macpherson et al., 2020](#)). Strong incentive for further exploration in these domains comes from clinical research investigating psychiatric conditions - neurodevelopmental ASD and severe long-term disorder of schizophrenia (SCZ), which we now shortly review.

3.4.1. Clinical evidence from ASD studies

ASD is characterized by impaired development in terms of social interaction in general, communication and motor behaviors ([American Psychiatric Association](#)), with the underlying causes being still poorly understood. The ability to share attention with others, as well as imitate them, is pivotal for human development with the first signs of sharing experiences being recorded as early as in the first year of age in typical developing children ([Kellerman et al., 2020](#); [Tomasello, 2011](#)). Those two adaptive mental functions enable symbolic play later in toddlerhood ([Baron-Cohen and Cross, 1992](#)), which lends itself to learning how to cooperate (e.g., turn-take) and communicate with others ([Nadel, 2015](#)). [Hobson and Hobson \(2007\)](#) proposed that ASD come from the difficulty to differentiate oneself from others ('theory of mind'). [Fulceri et al. \(2018\)](#) reported that ASD children synchronize with others better and imitate them more accurately ([Jiménez et al., 2014](#)) if the spatial goal for their own movement is clearly demarcated. This helps to draw a boundary with others. The evidence for the intact ability in ASD to

imitate (Bird et al., 2007; Heyes and Catmur, 2021) is contradictory with studies reporting spectrum of difficulties with imitation and acting in synchrony (Baillin et al., 2020; Brezis et al., 2017; Fitzpatrick et al., 2017, 2016; Forbes et al., 2016; Koehne et al., 2016b; Marton-Alper et al., 2020; Tunçgenç et al., 2021; Williams et al., 2004). Impact of ASD on joint action during daily activities is also not clear from the scientific literature (Cerullo et al., 2021), with reports of children with ASD showing less interactional synchrony during naturalistic conversation with their partners (Zampella et al., 2020). Another study looking at the action of grasping a bottle, found that participants with ASD did not wait for their action partner and showed prolonged movements (Curioni et al., 2017). Trevisan et al. (2021) found that participants with ASD did not perform as well as their typically developing peers (measured as task performance and ability to sync steps with the interaction partner) in the collaborative task of carrying a table. In both reports (Curioni et al., 2017; Trevisan et al., 2021), difficulties with joint action performance were not linked to measures of other ASD-related motor deficits. In opposition to those reports, Scharoun and Bryden (2016) found no differences in healthy controls and ASD in joint dyadic tasks involving daily action (passing an empty glass of water to the researcher).

Complementary evidences exist in the brain neuroimaging literature. Studies using fMRI for action imitation and observation revealed that activity patterns for ASD in the MNS areas were altered compared to healthy control, along with networks involved in the social cognition and executive function (Chan and Han, 2020). Differences in neural activation patterns with healthy controls were also found with fNIRS during action observation and dyadic joint action in a block building task (Su et al., 2020a). Also, a recent facial electromyography study (Schulte-Rüther et al., 2017) demonstrated that although basic mirror mechanisms in ASD are preserved, they do not link to the high-order social cognition that allows emotion understanding and empathy. Difficulties reading facial expressions of negative emotions were reported for young males with ASD using EEG recordings (Van der Donck et al., 2020). Disrupted brain-to-brain coupling in ASD children was also found in the hyperscanning fNIRS study looking at joint dyadic action of ASD children with their parents during keypress tasks (Wang et al., 2020), but not in the study by Kruppa et al. (2021). Linked to this topic, we identified an fMRI study (Moriguchi et al., 2009) looking at alexithymia in adults (self-awareness of emotions), demonstrating that people with alexithymia have higher activation in the MNS area, therefore showing similar difficulties of differentiation between self and others (i.e., a neural signature in the right superior parietal structure) to those found in the ASD population. Dunsmore et al. (2019) investigated the physiological linkage between interaction partners (heart inter-beat intervals) and found that the patients suffering from ASD do not sync their heart activity with a physical presence of another person in the room as observed in healthy controls (Scarpa et al., 2018).

In sum, a prolific state of the art on ASD reveals differences at both neural and behavioral levels to neurotypical peers, in the ability to share attention during interaction with another person, to perform a joint task together, and to read and recognize their emotions. There is a contradictory evidence concerning the specific role of imitation ability in those deficits.

3.4.2. Clinical evidence from SCZ studies

Schizophrenia (SCZ) is usually diagnosed by the presence of negative symptoms, understood as social withdrawal and emotional flatness, and positive symptoms, understood as change in behavior or thoughts due to hallucinations or delusions. Green et al. (2015) reviewed the literature describing the difficulties with social interaction characteristics for SCZ and summarized them as deficits in empathy, reflective social processing - mentalizing, emotion regulation, face and voice perception. In parallel, the decrease in synchronization performances in people with SCZ compared to healthy controls was found in multiple studies, particularly for intentional synchronization (Manschreck, 1981; Varlet et al., 2012), predictive timing (Wilquin et al., 2018) and lack of sensitivity to social

cues facilitating the synchronous performance with others (Cohen et al., 2017) or imitation (Sansonetti et al., 2020). In a pendulum-based synchronization task, Del-Monte et al. (2013) also showed that non-affected relatives of people with SCZ exhibit a decrease in synchronization ability, due to compromised visual tracking, pointing toward genetic SCZ phenotype for cognition. Interestingly, Raffard et al. (2018) found compromised stability of synchronization in patients with SCZ, but observed it to a lesser degree if the participants were positively primed to improve their sense of 'connectedness' to their task partners before performing the task. In a more ecologically set study by Kupper et al. (2015), participants with a paranoid type of SCZ revealed broken body synchrony patterns during dyadic role play conversation with healthy people. Synchrony negatively correlated to the social competence and severity of the negative symptoms, with participants not imitating the movements of their interlocutors, regardless of SCZ medication. Positive symptoms interacted with a lack of synchronous behavior by the healthy interlocutor, which could display affiliation, perhaps linked to the erratic movements caused by psychotic behavior.

Neurophysiological investigations of social and emotional syndromes in SCZ patients have also revealed interesting findings in the context of this review. For instance, searching for brain activation patterns in a fMRI scanner when observing recordings of finger movement and facial expressions, Horan et al. (2014) reported that people with SCZ did not differ from healthy controls, but showed a disconnection between brain activation and self-reported empathy reported through the Interpersonal Reactivity Index. In the same vein, Marosi et al. (2019) used EEG to investigate face and facial affect recognition in people with SCZ and revealed irregularities in the activity of the magnocellular pathway for face and face emotion processing. In a large study comparing MRI scans, Schilbach et al. (2016) reported differences between people with SCZ and healthy controls with regards to the connectivity in the MNS network and mentalizing network (left dorsomedial prefrontal cortex, left praecuneus, right and left temporo-parietal junction). Irregular connectivity in this area sheds light on the interpersonal difficulties that the patients with SCZ experience and on their ability to act with others. Together, these clinical studies, along with Subsections 2.3 and 3.2, mount evidence that the MNS network might be a double agent for emotion and acting together.

3.4.3. Interventions for ASD and SCZ

There is a great interest in using imitation as leverage for intervention studies for both ASD and SCZ. In an early intervention report, ASD children who received structured intervention focused on imitation and joint attention improved their social interaction skills, such as gaze following and requesting (Warreyn and Roeyers, 2014). Similarly, Landa et al. (2011) ran a randomized control study looking at the intervention targeted at imitation versus other therapy approaches long term improvement of positive affect and joint attention. Koehne et al. (2016a) reported benefits of intervention for adults with ASD using a dance therapy program focused on movement imitation and synchronization which over three months improved emotion interference along with other abilities to imitate and synchronize with others. For individuals with SCZ who participated in the therapy sessions, involving imitation of others and other theory of mind components (with a control group in a therapy focused on the problem solving skills) improved emotion recognition from the social situation and from the understanding of their intention of the movement (Mazza et al., 2010).

In sum, research in ASD and SCZ (both socio-motor interaction deficits) shows complex relationship between multi-faceted difficulties (such as differentiation from others, poor synchrony and imitation) and the ability to act together with others and understand their emotions (see Fig. 3). Neuroimaging studies pinpoint differences in information processing in those populations related to the MNS network along with its linkages to higher social cognition and mentalizing networks.

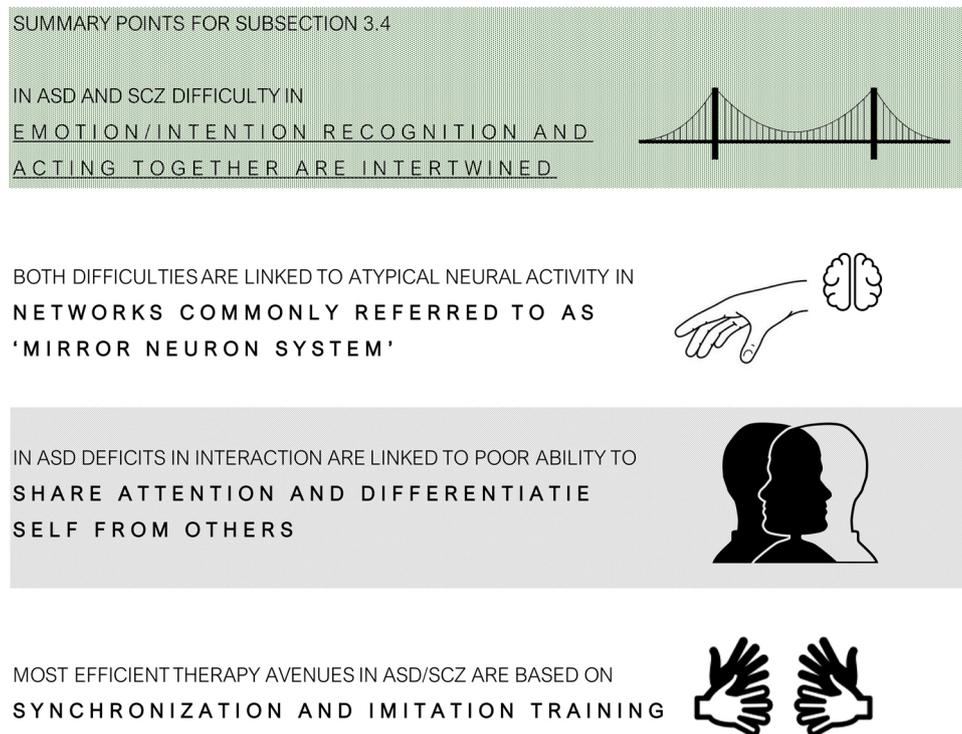


Fig. 3. Graphical summary of Subsection 3.4 highlights. Images: Suspension Bridge by Pechristener; Brain, idea, mind icon from iconfinder.com; Grabbing Hand by Oleksandr Panasovskiy/Psychologist by Dirk-Pieter van Walsum from the Noun Project.

4. Emoting joint action with non-human agents

Consistently over the last years, research in Human-Robot Interaction (HRI) and in Human-Computer Interaction (HCI) focused its efforts on developing social artificial agents that can initiate and maintain efficient and pleasurable interaction with a human. Emotional convergence benefits coordination between partners (Butler, 2015), and design studies in HCI have marshaled evidence that synchrony in movement qualities for facial and body expressions is essential for fluent human communication with virtual agents (Castellano et al., 2010; Marin et al., 2009; Numata et al., 2020).

Current developments in HRI are targeting real-time sensitivity to human expressions and behavior to promote long-term relationships (Castellano et al., 2008; Terada and Takeuchi, 2017). One of the main challenges is the capacity of the agent to operate on fast timescales, under one second, to be able to capture ‘social moments’ (Durantin et al., 2017). For instance, the PEPPER robot can infer possible interactive scenarios with customers via algorithms analyzing facial movement and voice signals in that time frame (Aaltonen et al., 2017). In other studies, robots adjust the interpersonal distance as a function of the estimated level of experienced emotions of the human in front of them (Bajones et al., 2017). The affect control theory offers a guiding principle used to create AI systems which are sensitive to affective states, adjusting their operations as a function of the context and need of their human interacting partners (Hoey et al., 2016). According to this theory, humans engage in situations that evoke emotions and feelings corresponding to one’s culturally built affective span. In general, captured data are used to infer human affective states to which social robots adapt in various semi-autonomous ways.

Scholars in HRI have adapted and simplified the human emotional repertoire to social robots. For instance, the ASIMO, JUSTIN and NAO robots are programmed to express six basic emotions (and their various combinations): anger, disgust, fear, happiness, sadness, and surprise. In general, human participants correctly recognize all basic emotions from the upper-body movements with a success rate of 75 %–100 %, with some exceptions however (van de Perre et al., 2015). Other robots, such

as the iCUB robot (Metta et al., 2008), crawl or semi-autonomously manipulate objects in various dyadic contexts, and learn by doing and imitating (Billard and Dautenhahn, 1998; Boucenna et al., 2014).

Core research activities in the field of affective interaction with artificial agents have been established around two main populations sensitive to affective interaction: people suffering from social disorders (with a particular focus on children with ASD) and elderly people. Here we briefly summarize the state of the art in these two domains, and then present recent trends in multi-agent and collaborative robotics.

4.1. Social HRI and HCI in ASD research

As it is unrealistic to expect children with ASD to continuously and smoothly interact with affectively embodied robots, adjustments have been made to simplify the child-robot interaction and discriminate between positive and negative emotions, particularly to launch social interaction (e.g., Feil-Seifer and Mataric, 2011). The PROBO robot, for instance, imitates animal movements which helps ASD children to recognize basic emotions (Pop et al., 2013). While some robots are only able to detect emotion displayed by humans, others, such as QRIO, also depict facial expressions and include corresponding body manifestations of some emotions, for instance happiness and fear, in a way that is recognizable by humans (Tanaka et al., 2004). Another example of a robot expressing emotions is MONARCH (Sequeira and Ferreira, 2016). This is a companion robot deployed in children’s hospital facilities and successfully integrated into a rich and complex clinical environment. In the related field of HCI, social robots are often replaced by virtual agents designed to create a specific social relationship with their human counterparts. Virtual and augmented realities are commonly used to help ASD children to focus on and recognize facial nonverbal cues (Chen et al., 2016), to learn to recognize and express emotions with their full-body movement (Alborno et al., 2016), to learn the required social skills (Lorenzo et al., 2019) or to promote verbal and nonverbal communication skills via joint actions (Srinivasan et al., 2016, 2015).

4.2. Social robotics for the elderly

The ageing population is another major category targeted by research on HRI, at home (e.g., [Fischinger et al. \(2016\)](#)) or in nursing institutions ([Moyle et al., 2013](#)). Older people develop a wide variety of age-related conditions that can cause their vulnerability to minor stressor events and lead to loss of autonomy: this phenomenon is commonly known as “frailty”. A number of interventions have been developed to target negative symptoms such as loneliness, anxiety and depression, which can also accompany dementia ([Cifuentes et al., 2020](#); [Kachouie et al., 2014](#); [Valentí Soler et al., 2015](#)). One example is a NAO robot-based rehabilitation program for people with dementia based in a geriatric ward, which reported higher outcome scores than conventional therapy on the immediate well-being and satisfaction ([Rouaix et al., 2017](#)). Similarly, *PARO* is a robotic seal that elderly residents in nursing homes benefit from by verbally interacting with it ([Moyle et al., 2013](#)). *HOBBIT* is another emotional assistive caregiving robot used at home to prevent falling ([Fischinger et al., 2016](#)). Further, a Social Assistive Robot exercise system was reported as more engaging the elderly in aerobic physical activity than virtual coach ([Fasola and Mataric, 2012](#)). The robot in [Zhang et al. \(2019\)](#) computes continuously the person’s movement trajectory, while assisting with their dressing, but is not emotion aware. These are only five examples that have been extracted from a plethora of research and proof-of-concept studies and have demonstrated how useful HRI and HCI approaches can be in the clinical context, as well as at home, to accompany healthy aging.

4.3. Environments of multiple human and artificial agents

When it comes to the environment being social, i.e., acting together in a group, a modest number of studies overcome the limitation of a robot(s) of working with more than one human, and only very few

robots are adapted to such interactions. For example, the interactive robot *KEEPON* can engage in both dyadic and triadic interaction due to emotional expressivity which aids to build joint attention with the interaction partner, e.g., looking in the same direction as the human ([Kozima et al., 2005](#)). Besides *NAO*, which is known to be able to work as a guide in a museum for a group of visitors ([Gehle et al., 2014](#)) or with school-aged children ([Hood et al., 2015](#); [Ros et al., 2014](#)), and *TIRO* which serves as a teaching assistant in musical classes ([Han et al., 2009](#)), the literature on human-robot group interaction remains scarce and almost exclusively in the form of “one-to-many” (a *star* graph as in the guide situation) in contrast to a more generic form of “many-to-many” (a *complete* graph, see [Bardy et al., 2020](#)). One of the unique endeavors employed triads of *BEAM* robots in a semi-autonomous control mode (Wizard of Oz) during game playing scenarios with human triads ([Fraune et al., 2019](#)). Human participants reported changes in subjective fear and motivation moderated by the perceived cohesion of the robot group, in comparison to other typologies with one human versus three robots and vice versa, and one-to-one interaction between a human and a robot. This indicates a breadth of emotional component to be explored in intergroup dynamics between human and artificial multiple agents, despite robots not being embodied with sensori-motor communication abilities (embodiment of emotion). However, despite being ‘emotionally neutral’, in a study by [Kochigami et al. \(2018\)](#), robots *NAO* and *PEPPER* played social roles by creating social ties between human group members (children and adults), and successfully facilitated interaction between them. Examples of similar studies are limited in number (see [Fig. 4](#)). [Sebo et al. \(2020\)](#) pinpointed key messages emerging from the current state of the art: (i) behavior in one person to one robot does not interpolate on the group behavior; (ii) verbal and non-verbal robot behavior shapes the response within the group and can support cohesion; (iii) people are more likely to engage with a robot when they are in groups; (iv) similarity (anthropomorphism) to humans plays a role in the

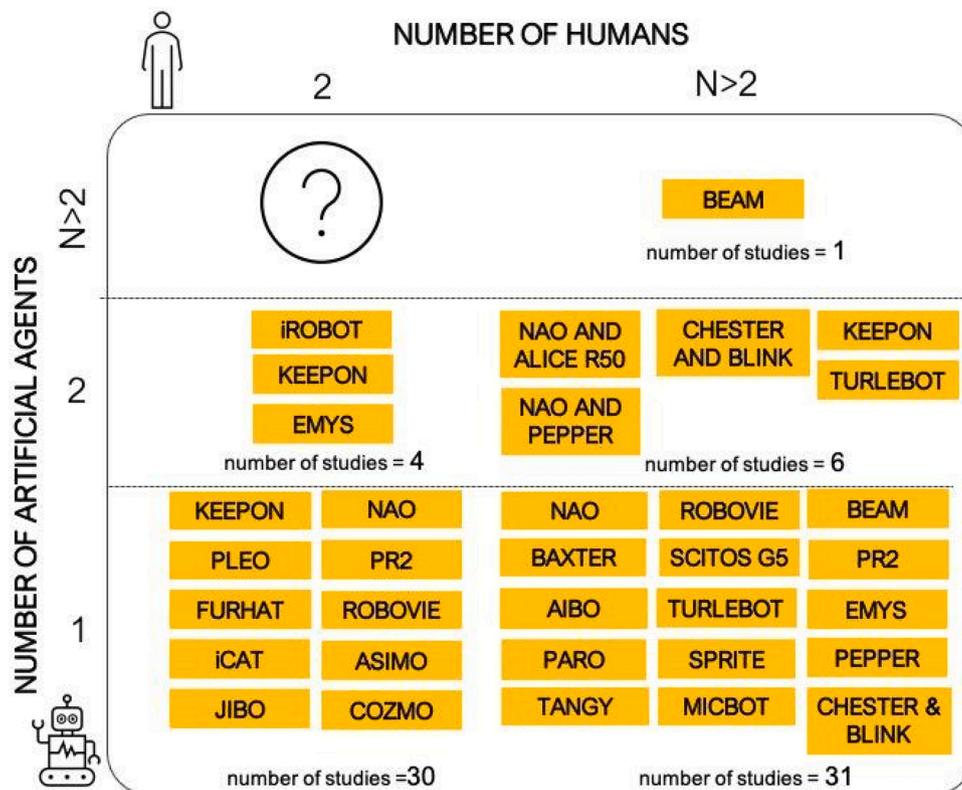


Fig. 4. The 2D matrix of the number of agents (persons ≥ 2 and robots) in HRI interaction extracted from the literature review of [Sebo et al. \(2020\)](#). Examples of the names of the robots used are placed on orange tiles with the count of overall studies identified by the researchers (refer to [Table 1](#) in [Sebo et al. \(2020\)](#) for further details).

integration of robots in a group. The dynamics of how emotion can be shared or propagated through the heterogeneous networks of several humans interacting with several artificial agents is unknown and has profound implications for the future of collaborative robotics.

4.4. Collaborative robotics in the industry

Emotion sharing as a means to facilitate social interactions in HRI has so far mostly been applied in therapeutic settings (see Subsections 4.1–4.2). In industrial settings, however, despite the heavy reliance on industrial robots for manufacturing, such examples are still rare. It is believed though that an essential component of the next industrial revolution, often referred to as Industry 5.0, will be that of the collaborative robot, a robot that can complement human co-workers, performing tasks that are either tedious or dangerous (Demir et al., 2019). The reason is because industrial sectors still lacking in terms of automation are those that cannot be fully automated, as they require human participation (Elprama et al., 2016). It has been recognized, however, that introducing collaborative robots to workplaces might have an adverse effect on social interactions in these workplaces. Untrained personnel, in particular, tend to expect the same social signals from robots as they would from human colleagues and expect the robots to adhere to existing social practices (Fischer, 2019). If collaborative robots fail to understand social signals and respond accordingly, they will be seen as impolite, cold and uncooperative. It also represents a missed opportunity to convey the robot's capabilities, while making communication more dependent on disruptive explicit signals, when more fluent implicit signals would have been preferable for seamless collaboration (Breazeal et al., 2005). Fischer (2019) further argues that collaborative robots do not only need to understand and produce social signals but that these signals need to include emotional expression. The reason, as seen in the above sections, is because sensori-motor communication of emotion and intention is an integral part of conventional social practice. A robot is simply expected to be sad when delivering bad news or happy when successfully completing a challenging task. Emotional expression may also be used to communicate real needs, such as when the system is running out of power and needs to be recharged. Recognition of human emotional expression under natural industrial conditions is difficult, as the technology needs to be both non-intrusive and robust over time. Speech (Khalil et al., 2019), gaze (Admoni and Scassellati, 2017) and facial expressions (Li and Deng, 2020) become more convenient cues than gestures or full body movement (Liu and Wang, 2018). However, with the introduction of cheaper wearable sensors, emotion recognition from EEG has recently become a viable alternative (Toichoa Eyam, 2019; Zheng et al., 2019). For expression of emotions, collaborative robots are limited by their embodiment and interfaces have more often been used for conveying information than for social signaling. Most collaborative robots only rely on projections of faces on flat screens to express emotions (Kalegina et al., 2018), if such expressions are used at all. There are recent examples, however, where the embodiment has in fact been exploited for social signaling, even highlighting the importance of a breathing motion (Maric et al., 2020; Terzioğlu et al., 2020), opening new emotion-based perspectives in collaborative robotics.

5. Avenues for future research

Sections 1–4 presented ample evidence for the interplay between emotions and joint action. Humans are unequivocally attuned to each other, with body movement being a powerful carrier of idiosyncratic information (Cutting and Kozlowski, 1977; Loula et al., 2005; Troje, 2002) and socially meaningful qualities (e.g., Centelles et al., 2011; Clarke et al., 2005; de Gelder and Poyo Solanas, 2021; Nackaerts et al., 2012). Information about the arousal of a person encoded in movement and their intention can be transferred to another person, for example, as

a forewarning of a threat. Being able to detect those non-verbal signals from others, along with the ability of humans to couple their body and brain activity, is the cornerstone of successful communication and cooperation between people. We strongly believe that interaction is key to understanding the human brain, as the human brain, through interaction with the environment, is of a physical, but also fundamentally of a social nature (Section 4 recaps why this also applies to hybrid interactions between humans and artificial agents). We acknowledge, as do other researchers, that social interaction should be at the forefront of neuroscience research (Schilbach et al., 2013). Some recent attempts, such as the social alignment theory, using herding modeling by Sharmay-Tsoory et al. (2019), are providing other important milestones towards this venture.

In this last Section (5) of our review, we put forward the idea that emotional arousal should be considered as an integral part of the so-called 'motor system', shaping and fine-tuning the real-time socio-motor interaction with others. Emotions arise as responses to the stimuli in the environment (with a function to maintain/restore allostasis) and bear subsequent impact on one's perception, affective state, ongoing and future movements (e.g., Wood et al., 2016). Thus, we propose to incorporate emotion in a joint action context as one entity, a 'third eye' that steers other mental and physiological processes to navigate the rich, multi-modal layers of the multi-agent social space. As emphasized before, the scientific evidence on the emotional embodiment during socio-motor interaction is limited (see Section 3.3 for overview), especially in terms of studies exploring real, not enacted, emotional arousal with naturalistic scenarios as a backdrop. To disentangle and decode the emotion propagation and socio-motor interaction not only via movement, but also via physiological processes, we propose to follow new research avenues (Subsections 5.1–5.5) to decipher the unknowns about interplay of emotion and joint action (see below the research questions we have identified as a part of literature review process, continued further on Fig. 6).

Research questions

We now know that some features of body posture and movement carry emotional qualities (de Gelder and Poyo Solanas, 2021), but can we find a motor signature of emotional arousal and valence in body movements in the context of socio-motor interaction (regardless of particular body parts) specific to i.e., particular levels of valence or arousal? To what extent context and culture shapes this signature during joint action? Is there a group emotional signature emerging from individuals sharing space at the same time (i.e., euphoria of football fans in the tribunes) (Subsection 5.1)? How do emotions evolve and propagate through the network of agents (humans or hybrid groups of human and artificial agents) and if so, how do they influence the outcome of joint performance (Subsection 5.2)? Further, we dive into the need for adoption of multiple timescales approaches, which emerged throughout this review (Subsections 2.4, 3.2–3.4, 4), showing that environmental function of emotion unfolds over multiple time windows, throughout physiological and movement qualities, as well as that some factors in socio-motor interaction are only meaningful when looked through an appropriate temporal lens (i.e., expertise, culture, previous experience with an agent). What are then the crucial timescales we need to integrate into the research agenda to advance this enquiry (Subsection 5.3)? How can AI techniques assist in this process, and unravel the patterns of information about emotion/intentional qualities and its propagation in agents during joint action (Subsection 5.4)? Finally, we dive into the world of digital, currently disembodied interaction, and highlight the notions that recent pandemic experience has left in terms of interplay of physical presence and emotion embodiment during social interactions (Subsection 5.5).

5.1. The notion of emotional motor signature in joint action

A large body of the work reviewed in Section 3.3 cements the foundations of the embodied nature of emotions, according to which expressions of emotional dimensions, rooted into common neurophysiological structures between cognition and action, are diffusing into movement physiology and are visible in facial, distal, as well as proximal parts of the body (Barrett et al., 2019b; de Gelder and Poyo Solanas, 2021; Kleinsmith and Bianchi-Berthouze, 2013; Melzer et al., 2019; Witkower and Tracy, 2019). Interestingly, this extensive literature does not yet intersect with another parallel body of research developing concepts and methods to assess Individual Motor Signatures (IMS), the idiosyncratic way each individual moves (e.g., Słowiński et al., 2016). The pioneering work of Johansson (1973) showing that - transformational invariants (Malcolm, 1953) persistence in some dimensions (e.g., length, ratios) across the motion of others (e.g., global transformation of the local optic flow) could help observers to quickly extract person-related properties, is very relevant to the socio-motor interaction context. IMS often relies on movement velocity as a key feature, as it is both stable across time and repetitions for each individual (movement similarity) and differential between individuals (inter-individual movement difference). Differences in the way people move during the performance of a motor task can be captured by using 95 % confidence interval ellipses in the similarity space (Słowiński et al., 2016). This is an abstract two-dimensional geometrical space minimizing distances between repetitions and individuals by using ad-hoc dimensional reduction techniques. Ellipses can be large or small depending on intra-individual variability and can be close or distant from each other depending on between-individual variability. The approach has proven useful in identifying IMS in various populations, ranging from healthy individuals to people suffering from schizophrenia (e.g., Słowiński et al., 2017). It has also proven useful in various tasks and contexts such as in the mirror game (Słowiński et al., 2016) or during improvisation movement (e.g., Coste et al., 2019), and at different distal or proximal (and more postural) parts of the body (e.g., Coste et al., 2021). Whether IMS, {as well as Group Motion Signatures (GMS, inter-group movement differences), i.e., the way IMS are assembled together in an ensemble of individuals engaged in reaching a common goal during joint action}, are emotionally neutral, and whether they are of different shapes and locations in the similarity space when produced in various emotional contexts (intra-subject variability) remains open to investigation (see Lozano-Goupil et al., 2021, for a first evaluation of Emotional IMS). Taking this road would not only answer the above questions, but would also offer a way to reconcile existing theories of emotion and those of embodied social interaction, inclusive of intra- and inter-individual/group variability and concepts such as motor accents (e.g., Ting et al., 2015), in a real-life context of a joint action.

5.2. Emotional group synchronization models

As emphasized in Section 2.2, models of perceptuo-motor social synchronization when $N > 2$ have not yet incorporated emotional qualities in their constituents, i.e., they remain emotionally neutral, despite the evidence gathered in this review that emotions are contagious, propagate through the social network, and constitute the essence of joint action. One urgent avenue of research requires a complementary approach incorporating emotional qualities in experimental and modeling scenarios. On the experimental side, the manipulation of positive, negative, and mixed emotional qualities, be they enacted or (ideally) induced, and the observation of how these qualities propagate from one node to the next across the collective sensori-motor network, converge or conflict, is requested. On the modeling side, coloring coupled oscillator models of

synchronization with those emotional qualities would help to better understand, and generalize, the underlying propagation mechanisms. For instance, the network of coupled Kuramoto oscillators presented in Section 2.2, capturing group synchronization regimes when perception is present (see Bardy et al., 2020, for details), needs to be adapted to incorporate emotional qualities at individual levels, such as:

$$\dot{\theta}_i(t) = \omega_i + c \sum_{j=1}^N a_{ij} \sin(\theta_{jEM}(t) - \theta_{iEM}(t))$$

Where N is the number of agents, θ_{iEM} the phase of the movement of the i -th agent under emotion EM , ω_i their natural frequency, and c the strength of the coupling with the other agents when perceptual coupling is established. Coefficients a_{ij} are set to a value between 0 and 1, depending on the dyadic perceptual coupling between agents i and j , i.e., the spatial configuration of the group. Coloring such oscillatory models with emotion-aware individual signatures (see Fig. 5, Example C), nourished by experimental data, would therefore be an operational way to close the gap between Sections 2–3 of the present review.

5.3. Embodied emotion across multiple timescales

In this review, we have hinted at the concept of multiple timescales on a handful of occasions. In Greek pre-Socratic philosophy time was represented by two notions; Chronos, sequential and linear time as we currently understand and apply it in a metric system (chronological time), and Kairos; which resembled the ‘right’ time, especially in the context of an action affordance (i.e., time for harvest). A myriad of research studies has provided data-driven rationale in favor of the use of multiple timescales to capture animal behavior and physiological processes (be it signal duration, temporal resolution, units applied or temporal dynamics). The evidence for multi-timescale behavioral organization has been recently investigated in *C. elegans*; showing how neural dynamics in this much simpler organism (slow - low frequency; fast - high frequency) orchestrates different movements and allows for flexibility of behavior (Kaplan et al., 2019). In humans, communication is regarded as a robust example of multi-modal behavior stretching across multiple levels of temporal structures due to the variety of interconnections between internal systems such as respiration and movement (Bardy et al., 2015; Pouw et al., 2021).

5.3.1. What do we know about temporal aspects of emotions?

Since the works of Solomon and Corbit (1974) it has been widely accepted that emotions unfold their dynamic over time, rather than spike events, with complex temporal structure (De Gelder et al., 2004; Frijda, 2007). Regardless if the stimulus is aversive or hedonic, the response curve for high arousal physiological reactions (i.e., heartbeat) unfurls as (i) rise to peak, (ii) adaptation period, (iii) recovery with reversed peak to re-establish baseline within 30–60 s. In a LORETA EEG paradigm, Costa et al. (2014) found a precise pattern of neural signatures of fear, disgust, happiness and sadness, with differences emerging mostly in the temporal characteristics of neural activation, but not the spatial spread. The differences that emerged are as follows: (i) Early onset (around 200 ms post exposure) and shorter duration characterized emotions - fear and disgust, which are associated with a need for quick body reaction; - (ii) Early onset (around 260 ms post stimuli) with a second processing peak at around 400 ms in different areas - happiness; and (iii) Late onset (around 400 ms post exposure) and longest duration (90 ms) - sadness. Personal diary study (Verduyn et al., 2015) reported similar temporal patterns linking to the adaptive behavior evoked by emotion, meaning that fear and disgust operate on fast timescales as they require quick fight-or-flight reaction; whereas emotions like anger

or joy, on average, take longer to disperse through action (Costa and Crini, 2011; De Gelder et al., 2004; Feldman Barrett and Finlay, 2018; Frijda, 2007). Notably, sadness was earmarked as the longest in duration, perhaps because its presumed function is to pave a pathway to rumination, motivation for change in personal circumstances or acceptance. Personal dispositions, from reactivity and resilience on a physiological level to higher cognitive functions such as emotion regulation processes (reappraisal), were highlighted as subjected to inter-subject differences (Solomon and Corbit, 1974)

5.3.2. How emotions influence motor timing?

Distal movements in particular (i.e., object manipulation) are a clear *reservoir* of emotions, a cornerstone assumption of forensic criminology. Gao et al. (2012) analyzed movement in a touch-based game and found (i) the length of the stroke to be indicative of the dimensional quality of valence, (ii) speed and direction to be indicative of arousal, while (iii) pressure specifically discriminated anger from other states (where increase in energy transmitted to movement has functional significance). Similarly, more frequent manipulation of the computer mouse was found to be associated with higher stress (Hernandez et al., 2014). During paced synchronization, adults and children tapped faster if they were primed with negatively valenced pictorial stimuli before the trial (Monier and Droit-Volet, 2018). The speeding up of motor response was interpreted as activation of the fear circuitry evoked by negative emotional induction (LeDoux, 2014), leading to the speeding up of the internal clock system (Cheng et al., 2016) and shifting movement towards faster timescales. Further, emotional arousal leads to subjective perception of time in some tasks (Gil and Droit-Volet, 2012), making a point that time perception is tied to the quality of stimuli (Grisey, 1987). This becomes particularly relevant in the context of untangling the dynamics of joint action and emotion (i.e., during synchronization).

5.3.3. What do we know about group temporal dynamics?

Vesper et al. (2011) has demonstrated that better coordination is achieved in dyadic action when participants make themselves more predictable (less temporally variable), in comparison to performing identical (pointing) movement alone or next to another person without intention to act together. Grammer et al. (1998) demonstrated that opposite sex pairs show a complex temporal structure of interaction patterns of body movement during conversation, which is repeated if both sides show interest, and is unique for each dyad. In a previously mentioned model of a psychotherapeutic alliance (Koole and Tschacher, 2016) three timescales were proposed for interpersonal synchrony, namely; (i) a *phasic* time-scale, which runs from a few hundred milliseconds to 10 s, characteristic of motor synchrony, (ii) a *tonic* time-scale, which runs from about 10 s to an hour, and involves more complex forms of social cognition, such as language and reasoning, and (iii) a *chronic* time-scale, stretching from weeks to years, and which involves the development of complex emotion-regulatory abilities. Bardy et al. (2020) reported that social memory (expertise in dance practice, related to (iii)) can affect the ability to synchronize with others under different perceptual strains. Similarly, experts in capoeira and tango had higher ability (kinesthetic ability) to imitate and synchronize with others, in comparison to athletes who also practice group sports, but without the synchronization component (Koehne et al., 2016c). More broadly, Burgoon et al. (1995) suggested that behavioral norms can pass from one generation to another as culture (i.e., think of the jovial behavior expected from a salesman versus the stoicism of a medical professional). An unexplored territory is investigation of previous personal experience with agents partaking in socio-motor interaction, which can trigger

certain emotions prior or during the execution, due to predictions of the internal model (Barrett, 2017b).

Taken together, these findings hint at a hidden hierarchy of socio-motor interaction, from low-level, fast timescales, which are more appropriate for immediate behavioral synchronization and responses, to high-level, slower timescales, which involve complex cognition and emotion regulation, linked to perception of emotional qualities and social memory. In this context, future research should apply a multi- (or inter-) modal multiple timescale approach for studies set on the interface of emotion and human joint action.

5.4. Leveraging artificial intelligence (AI) methods to capture emotions in socio-motor interactions

AI offers powerful analytical tools to deal with complex data, and so it is a valuable method for investigating individual and group motion signatures of emotions within the context of joint action. As discussed in Section 5.3, a multiple timescale approach is required to build a solid foundation of how various emotional qualities propagate in joint action could provide a window into relative aspects of time and how it is linked to the differentiation of emotional qualities in movement and the opportunity to act collectively. However, AI methods that address multiple timescales are largely limited to encoding each temporal variable (e.g., each different modality or modality dimension) separately. Such methods can only really create models tuned to a single timescale per temporal variable. Perhaps the more promising direction is methods that capture multiple timescales within each variable itself.

The few studies in this direction (e.g., Gurcan and Nguyen, 2019; Ma et al., 2019; Yamashita and Tani, 2008; Yang et al., 2020) have so far been constrained to individual action modeling. Two of the few exceptions are the multiple timescale recurrent neural network (MTRNN) (Yamashita and Tani, 2008) and the Approach Group Graph Convolutional Neural Network (AG-GCN) (Yang et al., 2020). Whereas the AG-GCN was designed to model group behavior, the MTRNN was originally developed for individual action scenario but has been extended to dyadic (Hinoshita et al., 2009a) and group interaction (Jaderberg et al., 2019), although only for robots in simple robot-robot interactions and bots in a multiplayer computer game. There is not much analysis of the behavior of the MTRNN in the multi-agent settings, but a functional hierarchical structuring of events in the movement sequences sampled was shown to emerge through modeling at multiple timescales in its original use (Yamashita and Tani, 2008). Similar findings reported for multiple timescale AI architectures explored in the context of natural language processing (Chung et al., 2017) underline the value in pursuing multiple timescale modeling for further understanding emotion in joint action. Moving forward, there is the need for new architectures that simultaneously: 1) have pathways for both action and emotion recognition, drawing from neuroscientific findings of multitask coding of observed action (behavior identification and semantic interpretation) in humans (Gallese, 2007; Iacoboni et al., 2005); 2) with multiple timescales processing for the two, rather than for one or the other (such as in Hinoshita et al., 2009b); and 3) capture individual as well as group signatures in group settings. Such architectures have the potential to maximize the value of AI for different aspects of emotion expression, perception, and propagation in the context of joint action.

In the current age of deep learning where low-level layers are entrusted with the extraction of signatures (i.e., features) from continuous streams of sensor data, there is a shift away from focus on hand crafting of computational features, with increasing significance

instead placed on the design of architectures that learn relevant features organically and directly from sensor data. A more valuable approach may be a blend of both methods. On one hand, automatic feature extraction sub-architectures could perhaps lead to (deeper) insight into what behaviors or other responses at individual and group levels characterize emotions in joint action. Contemporary understanding of these individual and group signatures could, on the other hand, be further explored by employing them in the form of hand-crafted features. The large number of existing studies on affect recognition (see Section 3.3 for a discussion of some of these) would be valuable in guiding the choice of individual level features to examine in the context of joint action. The minimal set of studies (Mou et al., 2015; Ukita et al., 2016; Yang et al., 2020; Yücel et al., 2013), that have used group relations features for affect recognition and related AI areas highlight differences in distance, speed, and direction (as well as displacement and/or velocity) between the individuals in a group as additional features to consider. As anticipated by some studies (e.g., Ukita et al., 2016; Yücel et al., 2013) the problem of determining (the extent of) a group of interest is a challenge that may need to first be addressed, especially for autonomous AI to be integrated into real world settings, before understanding how emotion experience in joint action can be possible (for an overview of the current state of the art on emotion recognition on groups please see Veltmeijer et al., 2021). Such an AI system would need to be able to determine if there is any joint action group present in a given location or sequence of events, the number of such groups, and the membership of each of them.

5.5. Emoted disembodied joint action in the digital world

The COVID-19 pandemic has speeded up our move into digital encounters across all aspects of our social life, be it work, education, leisure or even health. Virtual interactions have enabled millions of people to continue working together remotely (video calls e.g., Zoom or virtual spaces with people represented by avatars e.g., Sococo or Virbela). However, these virtual interactions have been impoverished in emotional context due to the lack of information, coming from gestures, body postures and facial expressions, about emotional arousal and agent's intentions. These non-verbal cues are critical to communication, understanding and bonding, recently captured by de Gelder and Poyo Solanas (2021) as mid-level qualities. Musicians, teachers and athletes across the globe have experienced how different it feels to perform without an audience feeding back their reception. After all, part of our identity comes from socio-motor interaction with others, by which we can express our personal qualities, such as being funny or having a preference to lead or to follow. Others reflect our qualities via sensorimotor communication, which is the foundation for validation and updating our sense of self. The lack of such easily accessible non-verbal cues in virtual spaces and the amplification of facial and gaze cues over body cues in video calls (e.g., Zoom) have been suggested to contribute to "Zoom fatigue" (Bailenson, 2021). In sum, lack of socio-motor interaction with others deprives our brain from the habitual process of predicting the unfolding of their actions in order to efficiently affiliate and cooperate with them in real time. Given that a digital environment will most probably last to a certain extent post COVID-19, it opens the opportunity, but also calls for embedding and facilitating joint emotional interaction to become effective. How to enhance the communication of emotional expressions in virtual spaces has been previously investigated. However, these studies have been limited to simply manually expressing such states (e.g., Pita and Pedro, 2011) showing that in such situations people spend more time in carefully crafting verbal affective

expressions than they do in gestural ones, possibly because of the lack of embodiment of the latter. Sensing technology, affective computing and sensory interactions or substitution research can have a crucial role in creating and sharing a sense of agency, a felt embodied affective state and at the same time advancing our understanding of how emotions become joint experiences. Leithinger et al. (2014) have shown how our own hand gestures can be transferred to another physical space as 3D objects for the others to experience in action. Remote tactile interactions, through the use of wearable devices that stimulate the other person skin in response to a remote tactile gesture (e.g., tactile exchanges in Huisman et al., 2013), such as skin stretching or being pinched (Hamdan et al., 2019; Muthukumarana et al., 2020), could help maintain the affective power of our non-verbal behavior during remote communication. Unfortunately, none of these studies have yet explored how such approaches are suitable to transfer the emotion qualities of an action and definitely not how such emotion qualities transfer across a group. Instead, transfer and group dynamics have been explored through disembodied representations of emotion-related signals (e.g., galvanic skin response or HRV) or inferred emotions through computational algorithms (Ardizzi et al., 2020; Gashi et al., 2019). Also, very little attention has been given to the spatial and temporal aspects which characterize joint emotional experiences, which are becoming even more critical than before. Studies have also shown that the perception of self-location can be altered through the right manipulation of sensory feedback, as in Lenggenhager et al. (2007). As Nadler (2020) highlights, space takes new meaning and creates new affordances in these virtual spaces that alter the meaning of joint interaction. From a computational perspective, modeling group emotion may require us to integrate in the computational models the dynamic characteristics of such virtual spaces that are affected by their properties and typology of information flow (Bardy et al., 2020).

6. Summary

Emotions move us across multiple levels of qualities and timescales, for our own survival, and for higher collective purposes. The sheer physical presence of others in shared space and time fulfills the most primal of human needs, which is to belong to a group. The recent pandemic experience (COVID-19) has demonstrated a devastating effect of disruption of routine social interactions. Joint actions have been obstructed by social distancing measures or by being moved entirely to the digital world. Confinement has had profound and not yet fully understood effects on mental wellbeing across all age and gender groups, (Ammar et al., 2020), and has had impact on the development of social skills in children deprived of contact with their peers (Giménez-Dasí et al., 2020). In this review we have highlighted a need to close the gap in the research between emotion and socio-motor interaction across different disciplines, and have prompted specific questions to the scientific community to do so. Although various branches of science have separately focused on joint action and on emotion, there is a growing necessity to understand how emotions flow across our embodied social interactions, and how they affect us as individuals, as a group and as a society.

Fig. 5 depicts dimensions that were identified in this review as meaningful to obtain a full picture of emoted socio-motor interaction, inspired by the research "landscape" of second-person neuroscience, proposed by (Schilbach et al., 2013). Sections 2–3 are represented in the lateral panel as emotion depicted a shade gradient of agents engaged in interaction (where color denotes interaction type), and organized the possible consequences of socio-motor interaction, into three 'working' categories that emerged during our literature search: performance,

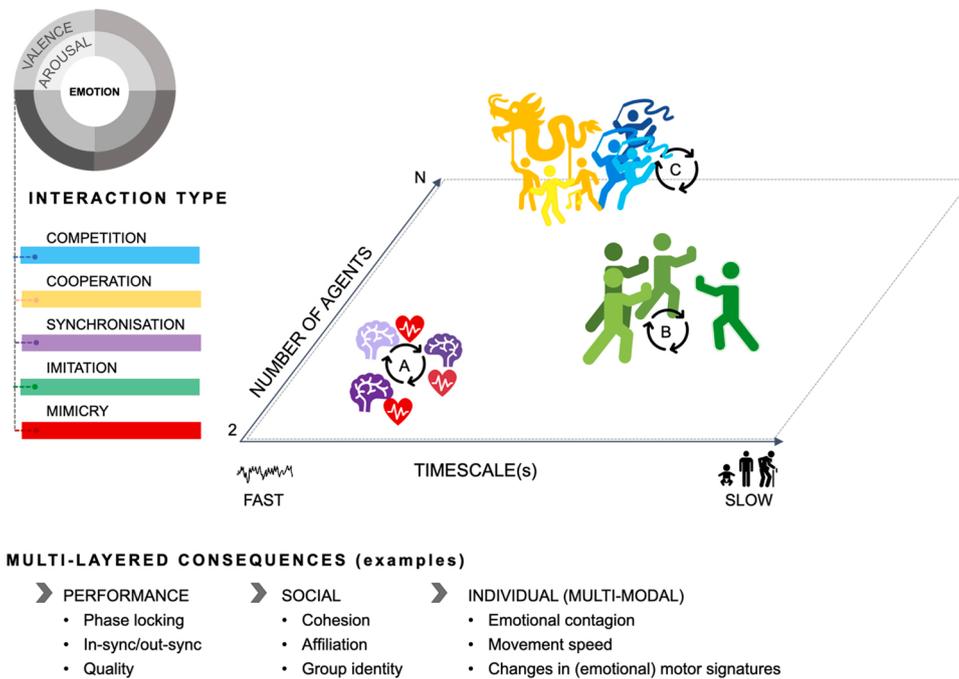


Fig. 5. Presentation of the future research landscape for emotion and joint action. The **bottom horizontal axis** represents **multiple timescales** that can be extracted from data (ranging from *ms*, i.e. brain and neural activity, to *hours and years*, i.e. expertise). The **vertical axis** denotes the number of agents engaged in socio-motor interaction, being perceptually and motorically active in the same physical or virtual space. The **colorful legend on the left side** represents possible **types of socio-motor interaction** that emerged from the literature review (delineating different spatio-temporal relationships between agents - see [Table 1](#) for summary). The **circle in the left corner** represents models of emotions (gradients relate to dimensions of valence and arousal) that need to be adapted to multiagent scenarios, and here are injected into a color scheme of interaction types. **Bottom panel** lists **multi-layered consequences of the socio-motor interaction** across three main categories: performance (quantitative and qualitative), social (i.e., affiliation and cohesion) and individual (impact on personal motor characteristics and emotional contagion from interaction with other agents), identified from this review. Example A – recordings of brain activity (hyper scanning) and heart activity during a naturalistic conversation between three people; Example B – A group is trying to follow and repeat the pattern from the

leader (Tai Chi class), captured with motion and respiration recordings; Example C – Group of agents dancing in and out of synchrony with each other (music street festival/parade), intertwined with periods of coordinated competition with each other, captured with motion recordings, following different typologies ([Bardy et al., 2020](#)). Images: The Tai Chi and Old Man icons come from <http://www.flaticon.com> (author: Freepik).

social and individual.

Recent models of emotion (i.e., [Barrett, 2017b](#)) have paved a theoretical path to integrate aspects of the presence of others and acting together in order to bridge a new, more informed, interdisciplinary avenue of research that is inclusive of dynamic relationships between emotion and joint action performance, when more than one agent is present, and of action context. The scientific evidence gathered and synthesized in Sections 1–3 of this review provides a weighty incentive to embrace more holistic and interdisciplinary approaches that are built on the assumption that our brain is primarily predictive, over reactive, and that our emotions are based of interoception and exteroception, and play an important allostatic function. Human abilities to understand emotions and to act together develop

simultaneously throughout the lifespan and show neural overlap in brain activity suggesting that both have been shaped by evolution to be interdependent. Therefore, the unraveling of the intrinsic relationship between emotion and socio-motor interaction needs to be built on modeling multi/inter-modal emotion propagation models, conceptualizing what group emotion is, and whether some emotions are exclusive to interaction (see [Fig. 6](#)).

This will nourish new modes of social interaction with non-human agents, which can provide personalized care, entertainment and life-long education for fragile populations. We argue that the deployment of machine learning algorithms and models supporting multiple timescales might provide the apposite caliber of research machinery to advance our comprehension of the dynamics of this ‘dark matter’

OUTSTANDING QUESTIONS

- How do individual emotions shape the way we move during joint action?
- Does joint action require agents to tune into the emotional state of other group members via empathy or to acknowledge it via mimicry?
- How can we track multi/(inter)-modal propagation of emotional embodiment (i.e. qualities/features)?
- Are certain emotions amplified by being with others while other emotions are exclusive to being with others (i.e., timidity, shame or desire)?

Fig. 6. Questions we have not found answer to during our literature review and need to be addressed by inter-disciplinary research.

(Schilbach et al., 2013), and dissect the multi-layered physiological, socio-behavioral consequences of acting together, inclusive of the emotion component and cohesion between agents. Progress in this cross-disciplinary field will feed future research and development in many technological areas such as collaborative robotics for industry and healthcare (i.e., incorporating the design of artificial agent principles, sensors and effectors for social-signaling and sensori-motor

communication), and will provide the tools for embodied digital interactions (for virtual workspaces and education).

Funding statement

This work was financed by the ‘EnTimeMent’ project funded by EU Horizon 2020 FET PROACTIVE (No 824160).

Appendix A

Keywords (acting together)	Keywords (emo)
joint action	emotion *
acting together	
acting in unison	
moving in unison	
moving together	Keywords (clinical)
group movement	Autism
group synchronisation	ASD
action coordination	Autistic
social motor coordination	Alexithymia
action cooperation	
entrainment	
action coupling	
motor coupling	
group imitation	
motor mimicry	
action mimicry	
propagation	
synchronisation	
chameleon effect	
synchrony	

PUBMED SYNTAX/ MAIN SEARCH	REVELANT HITS (repetitions were ignored ignored from the count)
(group movement [Title/Abstract]) AND emotion [Title/Abstract]	0/5
(group synchronisation [Title/Abstract]) AND emotion [Title/Abstract]	1/5
(synchronisation [Title/Abstract]) AND emotion [Title/Abstract]	3/24
(moving in unison [Title/Abstract]) AND emotion [Title/Abstract]	0
(moving together [Title/Abstract]) AND emotion [Title/Abstract]	4/78
(acting in unison [Title/Abstract]) AND emotion [Title/Abstract]	0/64
(acting together [Title/Abstract]) AND emotion [Title/Abstract]	2/2
(action coordination [Title/Abstract]) AND emotion [Title/Abstract]	1/1
(social motor coordination [Title/Abstract]) AND emotion [Title/Abstract]	1/39
(action cooperation [Title/Abstract]) AND emotion [Title/Abstract]	1/1
(motor coupling [Title/Abstract]) AND emotion [Title/Abstract]	0/2
(action coupling [Title/Abstract]) AND emotion [Title/Abstract]	2/5
(motor mimicry [Title/Abstract]) AND emotion [Title/Abstract]	2/11
(action mimicry [Title/Abstract]) AND emotion [Title/Abstract]	8/27
(group imitation [Title/Abstract]) AND emotion [Title/Abstract]	5/46
(entrainment [Title/Abstract]) AND emotion [Title/Abstract]	5/97
(joint action [Title/Abstract]) AND emotion [Title/Abstract]	0/5
(propagation [Title/Abstract]) AND emotion [Title/Abstract]	0/33
(Interpersonal coordination [Title/Abstract]) AND emotion [Title/Abstract]	5/25
(chameleon effect [Title/Abstract]) AND emotion [Title/Abstract]	2/2
(group synchronization) AND emotion [Title/Abstract]	3/96
(motor synchronization) AND emotion [Title/Abstract]	2/57
((emotion[Title/Abstract]) AND synchrony[Title/Abstract]) AND (motor[Title/Abstract] OR movement[Title/Abstract] OR motion[Title/Abstract])	11/17

PUBMED SYNTAX/ CLINICAL SEARCH	
AUTISTIC SPECTRUM DISORDER	
((autism[Title/Abstract] OR autistic[Title/Abstract])) AND (joint action[Title/Abstract] OR synchronisation[Title/Abstract] OR chameleon effect[Title/Abstract] OR synchronization[Title/Abstract] OR propagation[Title/Abstract] OR action mimicry[Title/Abstract] OR motor mimicry[Title/Abstract] OR group imitation[Title/Abstract] OR motor coupling[Title/Abstract] OR action coupling[Title/Abstract] OR entrainment[Title/Abstract] OR action cooperation[Title/Abstract] OR social motor coordination[Title/Abstract] OR action coordination[Title/Abstract] OR group synchronisation[Title/Abstract] OR group movement[Title/Abstract] OR moving together[Title/Abstract] OR moving in unison[Title/Abstract] OR acting in unison[Title/Abstract] OR acting together[Title/Abstract] OR group synchronization[Title/Abstract])	40/ 441
((autism[Title/Abstract]) AND synchrony[Title/Abstract]) AND (motor[Title/Abstract] OR movement[Title/Abstract] OR moving[Title/Abstract])	8/42
ALEXITHYmia	
((alexithymia[Title/Abstract] OR alexithymic[Title/Abstract])) AND (joint action[Title/Abstract] OR synchronisation[Title/Abstract] OR chameleon effect[Title/Abstract] OR synchronization[Title/Abstract] OR propagation[Title/Abstract] OR action mimicry[Title/Abstract] OR motor mimicry[Title/Abstract] OR group imitation[Title/Abstract] OR motor coupling[Title/Abstract] OR action coupling[Title/Abstract] OR entrainment[Title/Abstract] OR action cooperation[Title/Abstract] OR social motor coordination[Title/Abstract] OR action coordination[Title/Abstract] OR group synchronisation[Title/Abstract] OR group movement[Title/Abstract] OR moving together[Title/Abstract] OR moving in unison[Title/Abstract] OR acting in unison[Title/Abstract] OR acting together[Title/Abstract] OR group synchronization[Title/Abstract])	1/9
SCHIZOPHRENIA	
schizophrenia[Title/Abstract] AND (joint action[Title/Abstract] OR synchronisation[Title/Abstract] OR chameleon effect[Title/Abstract] OR synchronization[Title/Abstract] OR propagation[Title/Abstract] OR action mimicry[Title/Abstract] OR motor mimicry[Title/Abstract] OR group imitation[Title/Abstract] OR motor coupling[Title/Abstract] OR action coupling[Title/Abstract] OR entrainment[Title/Abstract] OR action cooperation[Title/Abstract] OR social motor coordination[Title/Abstract] OR action coordination[Title/Abstract] OR group synchronisation[Title/Abstract] OR group movement[Title/Abstract] OR moving together[Title/Abstract] OR moving in unison[Title/Abstract] OR acting in unison[Title/Abstract] OR acting together[Title/Abstract] OR group synchronization[Title/Abstract])	15/539
((schizophrenia[Title/Abstract]) AND synchrony[Title/Abstract]) AND (motor[Title/Abstract] OR movement[Title/Abstract] OR	0/28

References

Aaltonen, I., Arvola, A., Heikkilä, P., Lammi, H., 2017. Hello pepper, May I tickle you? In: Mutlu, B., Tscheligi, M., Weiss, A., Young, J.E. (Eds.), Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction. ACM, New York, NY, USA, pp. 53–54. <https://doi.org/10.1145/3029798.3038362>.

Acquadro, M.A.S., Congedo, M., De Ridder, D., 2016. Music performance as an experimental approach to hyperscanning studies. *Front. Hum. Neurosci.* 10, 242. <https://doi.org/10.3389/fnhum.2016.00242>.

Admoni, H., Scassellati, B., 2017. Social eye gaze in human-robot interaction: a review. *J. Human-Robot Interact.* 6, 25. <https://doi.org/10.5898/jhri.6.1.admoni>.

Aktar, E., Raijmakers, M.E.J., Kret, M.E., 2020. Pupil mimicry in infants and parents. *Cogn. Emot.* 34, 1160–1170. <https://doi.org/10.1080/02699931.2020.1732875>.

Al Osman, H., Falk, T.H., 2017. Multimodal Affect Recognition: Current Approaches and Challenges, in: Emotion and Attention Recognition Based on Biological Signals and Images. InTech. <https://doi.org/10.5772/65683>.

Alarcão, S.M., Fonseca, M.J., 2019. Emotions recognition using EEG signals: A survey. *IEEE Trans. Affect. Comput.* <https://doi.org/10.1109/TAFCC.2017.2714671>.

Alberdi, A., Aztiria, A., Basarab, A., 2016. Towards an automatic early stress recognition system for office environments based on multimodal measurements: a review. *J. Biomed. Inform.* <https://doi.org/10.1016/j.jbi.2015.11.007>.

Alborno, P., Volpe, G., Camurri, A., Clayton, M., Keller, P., 2015. Automated video analysis of interpersonal entrainment in Indian music performance. 2015 7th

- International Conference on Intelligent Technologies for Interactive Entertainment (INTETAIN) 57–63.
- Alborno, P., Piana, S., Mancini, M., Niewiadomski, R., Volpe, G., Camurri, A., 2016. Analysis of intrapersonal synchronization in full-body movements displaying different expressive qualities. In: Costabile, M.F., Buono, P., Matera, M., Lanzilotti, R. (Eds.), Proceedings of the International Working Conference on Advanced Visual Interfaces - AVI '16. ACM Press, New York, New York, USA, pp. 136–143. <https://doi.org/10.1145/2909132.2909262>.
- Alderisio, F., Bardy, B.G., Di Bernardo, M., 2016. Entrainment and synchronization in networks of Rayleigh-van der Pol oscillators with diffusive and Haken-Kelso-Bunz couplings. *Biol. Cybern.* 110, 151–169. <https://doi.org/10.1007/s00422-016-0685-7>.
- Alderisio, F., Fiore, G., Salesse, R.N., Bardy, B.G., Bernardo, M., Di, 2017. Interaction patterns and individual dynamics shape the way we move in synchrony. *Sci. Rep.* 7 <https://doi.org/10.1038/s41598-017-06559-4>.
- Allert, T., 2009. The Hitler Salute: on the meaning of a gesture. *Picador*.
- Altman, I., Vinsel, A., Brown, B.B., 1981. Dialectic conceptions in social psychology: an application to social penetration and privacy regulation. *Adv. Exp. Soc. Psychol.* 14, 107–160. [https://doi.org/10.1016/S0065-2601\(08\)60371-8](https://doi.org/10.1016/S0065-2601(08)60371-8).
- Ammar, A., Chtourou, H., Boukhris, O., Trabelsi, K., Masmoudi, L., Brach, M., Bouaziz, B., Bentlage, E., How, D., Ahmed, M., Mueller, P., Mueller, N., Hsouna, H., Aloui, A., Hammouda, O., Paineiras-Domingos, L.L., Braakman-Jansen, A., Wrede, C., Bastoni, S., Pernambuco, C.S., Mataruna, L., Taheri, M., Irandoust, K., Khacharem, A., Bragazzi, N.L., Strahler, J., Adrian, J., Andreeva, A., Khoshnami, S. C., Samara, E., Zisi, V., Sankar, P., Ahmed, W.N., Romdhani, M., Delhey, J., Bailey, S. J., Bott, N.T., Gargouri, F., Chaari, L., Batatia, H., Ali, G.M., Abdelkarim, O., Jarraya, M., Abed, K., El Souissi, N., Van Gemert-Pijnen, L., Riemann, B.L., Riemann, L., Moalla, W., Gómez-Raja, J., Epstein, M., Sanderman, R., Schulz, S., Jerg, A., Al-Horani, R., Mansi, T., Jmail, M., Barbosa, F., Santos, F., Šimunić, B., Pišot, R., Pišot, S., Gaggioli, A., Zmijewski, P., Apfelbacher, C., Steinacker, J., Saad, H.B., Glenn, J.M., Chamari, K., Driss, T., Hoekelmann, A., 2020. Covid-19 home confinement negatively impacts social participation and life satisfaction: a worldwide multicenter study. *Int. J. Environ. Res. Public Health* 17, 1–17. <https://doi.org/10.3390/ijerph17176237>.
- Anderson, C., Keltner, D., John, O.P., 2003. Emotional convergence between people over time. *J. Pers. Soc. Psychol.* 84, 1054–1068. <https://doi.org/10.1037/0022-3514.84.5.1054>.
- Ansuini, C., Cavallo, A., Bertone, C., Becchio, C., 2014. The visible face of intention: why kinematics matters. *Front. Psychol.* 5, 815. <https://doi.org/10.3389/fpsyg.2014.00815>.
- Ardizzi, M., Calbi, M., Tavaglione, S., Umiltà, M.A., Gallese, V., 2020. Audience spontaneous entrainment during the collective enjoyment of live performances: physiological and behavioral measurements. *Sci. Rep.* 10, 1–12. <https://doi.org/10.1038/s41598-020-60832-7>.
- Argyle, M., Cook, M., 1976. *Gaze and Mutual Gaze. Gaze and Mutual Gaze*. Cambridge U Press, Oxford, England.
- Arias, P., Belin, P., Aucouturier, J.J., 2018. Auditory smiles trigger unconscious facial imitation. *Curr. Biol.* <https://doi.org/10.1016/j.cub.2018.05.084>.
- Atherton, G., Sebanz, N., Cross, L., 2019. Imagine all the Synchrony: the effects of actual and imagined synchronous walking on attitudes towards marginalised groups. *PLoS One* 14, e0216585. <https://doi.org/10.1371/journal.pone.0216585>.
- Atkinson, A.P., Dittrich, W.H., Gemmell, A.J., Young, A.W., 2004. Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception* 33, 717–746. <https://doi.org/10.1068/p5096>.
- Atzil, S., Gendron, M., 2017. Bio-behavioral synchrony promotes the development of conceptualized emotions. *Curr. Opin. Psychol.* <https://doi.org/10.1016/j.copsyc.2017.07.009>.
- Aucouturier, J.J., Canonne, C., 2017. Musical friends and foes: the social cognition of affiliation and control in improvised interactions. *Cognition* 161, 94–108. <https://doi.org/10.1016/j.cognition.2017.01.019>.
- Aung, M.S.H., Kaltwang, S., Romera-Paredes, B., Martinez, B., Singh, A., Cella, M., Valstar, M., Meng, H., Kemp, A., Shafizadeh, M., Elkins, A.C., Kanakam, N., De Rothschild, A., Tyler, N., Watson, P.J., Williams, A.C.D.C., Pantic, M., Bianchi-Berthouze, N., 2016. The automatic detection of chronic pain-related expression: requirements, challenges and the multimodal EmoPain dataset. *IEEE Trans. Affect. Comput. 7*, 435–451. <https://doi.org/10.1109/TAFFC.2015.2462830>.
- Aviezer, H., Trope, Y., Todorov, A., 2012. Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science* 338, 1225–1229. <https://doi.org/10.1126/science.1224313>.
- Babiloni, C., Buffo, P., Vecchio, F., Marzano, N., Del Percio, C., Spada, D., Rossi, S., Bruni, I., Rossini, P.M., Perani, D., 2012. Brains “in concert”: frontal oscillatory alpha rhythms and empathy in professional musicians. *Neuroimage* 60, 105–116. <https://doi.org/10.1016/j.neuroimage.2011.12.008>.
- Bailenson, J.N., 2021. Nonverbal overload: a theoretical argument for the causes of Zoom fatigue. *Technol. Mind, Behav.* 2 <https://doi.org/10.1037/tmb0000030>.
- Baillin, F., Lefebvre, A., Pedoux, A., Beauxis, Y., Engemann, D.A., Maruani, A., Amsellem, F., Kelso, J.A.S., Bourgeron, T., Delorme, R., Dumas, G., 2020. Interactive psychometrics for autism with the human dynamic clamp: interpersonal synchrony from sensorimotor to Sociocognitive Domains. *Front. Psychiatry* 11. <https://doi.org/10.3389/fpsyg.2020.510366>.
- Bajones, M., Zillich, M., Vincze, M., 2017. Robots Should Respect Your Feelings As Well Adapting Distance Between a Robot and a User Based on Expressed Emotions.
- Balconi, M., Vanutelli, M.E., 2017. Cooperation and competition with hyperscanning methods: review and future application to emotion domain. *Front. Comput. Neurosci.* <https://doi.org/10.3389/fncom.2017.00086>.
- Bardy, B.G., Hoffmann, C.P., Moens, B., Leman, M., Dalla Bella, S., 2015. Sound-induced stabilization of breathing and moving. *Ann. N. Y. Acad. Sci.* 1337, 94–100. <https://doi.org/10.1111/nyas.12650>.
- Bardy, B.G., Calabrese, C., de Lellis, P., Bourgeaud, S., Colomer, C., Pla, S., Di Bernardo, M., 2020. Moving in unison after perceptual interruption. *Sci. Rep.* 10 <https://doi.org/10.1038/s41598-020-74914-z>.
- Bar-Kalifa, E., Prinz, J.N., Atzil-Slonim, D., Rubel, J.A., Lutz, W., Rafaeli, E., 2019. Physiological synchrony and therapeutic alliance in an imagery-based treatment. *J. Couns. Psychol.* 66, 508–517. <https://doi.org/10.1037/cou0000358>.
- Baron-Cohen, S., Cross, P., 1992. *Reading the Eyes: Evidence for the Role of Perception in the Development of a Theory of Mind*. *Mind Lang.* 7.
- Barrett, L.F., 2017a. Categories and their role in the science of emotion. *Psychol. Inq.* <https://doi.org/10.1080/1047840X.2017.1261581>.
- Barrett, L.F., 2017b. The theory of constructed emotion: an active inference account of interoception and categorization. *Soc. Cogn. Affect. Neurosci.* 12, 1–23. <https://doi.org/10.1093/scan/nsw154>.
- Barrett, L.F., Niedenthal, P.M., Winkielman, P. (Eds.), 2007. *Emotion and Consciousness*. Guilford, New York and London.
- Barrett, L.F., Adolphs, R., Marsella, S., Martinez, A.M., Pollak, S.D., 2019a. Emotional expressions reconsidered: challenges to inferring emotion from human facial movements. *Psychol. Sci. Public Interest* 20, 1–68. <https://doi.org/10.1177/1529100619832930>.
- Barrett, L.F., Adolphs, R., Marsella, S., Martinez, A.M., Pollak, S.D., 2019b. Emotional expressions reconsidered: challenges to inferring emotion from human facial movements. *Psychol. Sci. Public Interest* 20, 1–68. <https://doi.org/10.1177/1529100619832930>.
- Basso, J.C., Satyal, M.K., Rugh, R., 2021. Dance on the brain: enhancing intra- and inter-brain synchrony. *Front. Hum. Neurosci.* 14 <https://doi.org/10.3389/fnhum.2020.584312>.
- Baumeister, R.F., Leary, M.R., 1995. The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychol. Bull.* <https://doi.org/10.1037/0033-2909.117.3.497>.
- Bavelas, J.B., Black, A., Lemery, C.R., Mullett, J., 1986. I show how you feel”. Motor mimicry as a communicative act. *J. Pers. Soc. Psychol.* 50, 322–329. <https://doi.org/10.1037/0022-3514.50.2.322>.
- Becchio, C., Manera, V., Sartori, L., Cavallo, A., Castiello, U., 2012. Grasping intentions: from thought experiments to empirical evidence. *Front. Hum. Neurosci.* 6 <https://doi.org/10.3389/fnhum.2012.00117>.
- Bernhardt, B.C., Singer, T., 2012. The neural basis of empathy. *Annu. Rev. Neurosci.* 35, 1–23. <https://doi.org/10.1146/annurev-neuro-062111-150536>.
- Bernieri, F.J., Rosenthal, R., 1991. *Interpersonal coordination: behavior matching and interactional synchrony. Fundamentals of Nonverbal Behavior, Studies in Emotion & Social Interaction*. Editions De La Maison Des Sciences De l’Homme, Paris, France, pp. 401–432.
- Billard, A., Dautenhahn, K., 1998. Grounding communication in autonomous robots: An experimental study. *Rob. Auton. Syst.* 24, 71–79. [https://doi.org/10.1016/S0921-8890\(98\)00023-2](https://doi.org/10.1016/S0921-8890(98)00023-2).
- Bird, G., Leighton, J., Press, C., Heyes, C., 2007. Intact automatic imitation of human and robot actions in autism spectrum disorders. *Proc. R. Soc. B Biol. Sci.* 274, 3027–3031. <https://doi.org/10.1098/rspb.2007.1019>.
- Bloom, P., 2017. Empathy and its discontents. *Trends Cogn. Sci. (Regul. Ed.)*. <https://doi.org/10.1016/j.tics.2016.11.004>.
- Bodie, G.D., Cannava, K.E., Vickery, A.J., Jones, S.M., 2016. Patterns of nonverbal adaptation in supportive interactions. *Commun. Stud.* 67, 3–19. <https://doi.org/10.1080/10510974.2015.1036168>.
- Borch, C., 2015. The politics of the senses: crowd formation through sensory manipulation. *From Multitude to Crowds: Collective Action and Media*. Peter Lang, Frankfurt Am Main, pp. 53–69. <https://doi.org/10.3726/978-3-653-05926-7>.
- Borgomaneri, S., Vitale, F., Gazzola, V., Avenanti, A., 2015. Seeing fearful body language rapidly freezes the observer’s motor cortex. *Cortex* 65, 232–245. <https://doi.org/10.1016/j.cortex.2015.01.014>.
- Borgomaneri, S., Bolloni, C., Sessa, P., Avenanti, A., 2020. Blocking facial mimicry affects recognition of facial and body expressions. *PLoS One* 15. <https://doi.org/10.1371/journal.pone.0229364>.
- Boucenna, S., Narzisi, A., Tilmont, E., Muratori, F., Pioggia, G., Cohen, D., Chetouani, M., 2014. Interactive technologies for autistic children: a review. *Cognit. Comput.* 6, 722–740. <https://doi.org/10.1007/s12559-014-9276-x>.
- Breazeal, C., Kidd, C.D., Thomaz, A.L., Hoffman, G., Berlin, M., 2005. Effects of nonverbal communication on efficiency and robustness in human-robot teamwork. 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS. IEEE Computer Society 708–713. <https://doi.org/10.1109/IROS.2005.1545011>.
- Brezis, R.S., Noy, L., Alony, T., Godlieb, R., Cohen, R., Golland, Y., Levit-Binnun, N., 2017. Patterns of joint improvisation in adults with autism spectrum disorder. *Front. Psychol.* 8 <https://doi.org/10.3389/fpsyg.2017.01790>.
- Burgoon, J.K., Stern, L.A., Dillman, L., 1995. *Interpersonal Adaptation: Dyadic Interaction Patterns*. Interpersonal Adaptation: Dyadic Interaction Patterns. Cambridge University Press, New York, NY, US. <https://doi.org/10.1017/CBO9780511720314>.
- Butler, E.A., 2015. Interpersonal affect dynamics: it takes two (and time) to tango. *Emot. Rev.* 7, 336–341. <https://doi.org/10.1177/1754073915590622>.
- Butler, E.A., 2017. Emotions are temporal interpersonal systems. *Curr. Opin. Psychol.* <https://doi.org/10.1016/j.copsyc.2017.07.005>.
- Cacioppo, S., Zhou, H., Monteleone, G., Majka, E.A., Quinn, K.A., Ball, A.B., Norman, G. J., Semin, G.R., Cacioppo, J.T., 2014. You are in sync with me: neural correlates of interpersonal synchrony with a partner. *Neuroscience* 277, 842–858. <https://doi.org/10.1016/j.neuroscience.2014.07.051>.

- Camurri, A., Lagerlöf, I., Volpe, G., 2003. Recognizing emotion from dance movement: comparison of spectator recognition and automated techniques. *Int. J. Hum. Stud.* 59, 213–225. [https://doi.org/10.1016/S1071-5819\(03\)00050-8](https://doi.org/10.1016/S1071-5819(03)00050-8).
- Camurri, A., Volpe, G., De Poli, G., Leman, M., 2005. Communicating expressiveness and affect in multimodal interactive systems. *IEEE Multimed.* 12, 43–53. <https://doi.org/10.1109/MMUL.2005.2>.
- Camurri, A., Canepa, C., Ferrari, N., Mancini, M., Volpe, G., 2011. Modelling and analysing creative communication within groups of people: the artistic event at FET11. *Procedia Computer Science*. Elsevier B.V., pp. 144–145. <https://doi.org/10.1016/j.procs.2011.09.085>
- Camurri, A., Volpe, G., Piana, S., Mancini, M., Niewiadomski, R., Ferrari, N., Canepa, C., 2016. The dancer in the eye. In: Mitsaris, S. (Ed.), *Proceedings of the 3rd International Symposium on Movement and Computing - MOCO' 16*. ACM Press, New York, New York, USA, pp. 1–7. <https://doi.org/10.1145/2948910.2948927>.
- Cannon, W., 1953. Bodily changes in pain, hunger, fear and rage: *ann. Intern. Med.* 39, 383. <https://doi.org/10.7326/0003-4819-39-2-383.1>.
- Cappella, J.N., Greene, J.O., 1982. A discrepancy-arousal explanation of mutual influence in expressive behavior for adult and infant-adult interaction. *Commun. Monogr.* 49, 89–114. <https://doi.org/10.1080/03637758209376074>.
- Carr, L., Iacoboni, M., Dubeaut, M.C., Mazziotta, J.C., Lenzi, G.L., 2003. Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. *Proc. Natl. Acad. Sci. U. S. A.* 100, 5497–5502. <https://doi.org/10.1073/pnas.0935845100>.
- Castellano, G., Aylett, R., Dautenhahn, K., Paiva, A., McOwan, P.W., Ho, S., 2008. Long-term affect sensitive and socially interactive companions. *Proc. 4th Int. Work. Human-Computer Conversat.* 5.
- Castellano, G., Leite, I., Pereira, A., Martinho, C., Paiva, A., McOwan, P.W., 2010. Affect recognition for interactive companions: challenges and design in real world scenarios. *J. Multimodal User Interfaces* 3, 89–98. <https://doi.org/10.1007/s12193-009-0033-5>.
- Cavallo, A., Koul, A., Ansuini, C., Capozzi, F., Becchio, C., 2016. Decoding intentions from movement kinematics. *Sci. Rep.* 6, 1–8. <https://doi.org/10.1038/srep37036>.
- Centelles, L., Assaiante, C., Nazarian, B., Anton, J.L., Schmitz, C., 2011. Recruitment of both the mirror and the mentalizing networks when observing social interactions depicted by point-lights: a neuroimaging study. *PLoS One* 6, e15749. <https://doi.org/10.1371/journal.pone.0015749>.
- Cerullo, S., Fulceri, F., Muratori, F., Contaldo, A., 2021. Acting with shared intentions: A systematic review on joint action coordination in Autism Spectrum Disorder. *Brain Cogn.* 149. <https://doi.org/10.1016/j.bandc.2021.105693>.
- Chabin, T., Tio, G., Comte, A., Joucla, C., Gabriel, D., Pazart, L., 2020. The relevance of a conductor competition for the study of emotional synchronization within and between groups in a natural musical setting. *Front. Psychol.* 10. <https://doi.org/10.3389/fpsyg.2019.02954>.
- Chan, M.M.Y., Han, Y.M.Y., 2020. Differential mirror neuron system (MNS) activation during action observation with and without social-emotional components in autism: a meta-analysis of neuroimaging studies. *Mol. Autism* 11. <https://doi.org/10.1186/s13229-020-00374-x>.
- Chang, A., Kragness, H.E., Livingstone, S.R., Bosnyak, D.J., Trainor, L.J., 2019. Body sway reflects joint emotional expression in music ensemble performance. *Sci. Rep.* 9, 1–11. <https://doi.org/10.1038/s41598-018-36358-4>.
- Chartrand, T.L., Bargh, J.A., 1999. The chameleon effect: the perception-behavior link and social interaction. *J. Pers. Soc. Psychol.* 76, 893–910. <https://doi.org/10.1037/0022-3514.76.6.893>.
- Chartrand, T.L., Lakin, J.L., 2013. The antecedents and consequences of human behavioral mimicry. *Annu. Rev. Psychol.* 64, 285–308. <https://doi.org/10.1146/annurev-psych-113011-143754>.
- Chen, C.-H., Lee, I.-J., Lin, L.-Y., 2016. Augmented reality-based video-modeling storybook of nonverbal facial cues for children with autism spectrum disorder to improve their perceptions and judgments of facial expressions and emotions. *Comput. Human Behav.* 55, 477–485. <https://doi.org/10.1016/j.chb.2015.09.033>.
- Cheng, R.K., Tipples, J., Narayanan, N.S., Meck, W.H., 2016. Clock speed as a window into dopaminergic control of emotion and time perception. *Timing Time Percept.* 4, 99–122. <https://doi.org/10.1163/202134468-00002064>.
- Cheng, M., Kato, M., Saunders, J.A., Tseng, C., 2020. Paired walkers with better first impression synchronize better. *PLoS One* 15, e0227880. <https://doi.org/10.1371/journal.pone.0227880>.
- Cho, Y., Bianchi-Berthouze, N., Oliveira, M., Holloway, C., Julier, S., 2019. Nose heat: exploring stress-induced nasal thermal variability through Mobile thermal imaging. 2019 8th International Conference on Affective Computing and Intelligent Interaction, ACII 2019. Institute of Electrical and Electronics Engineers Inc. 566–572. <https://doi.org/10.1109/ACII.2019.8925453>.
- Chung, J., Ahn, S., Bengio, Y., 2017. Hierarchical multiscale recurrent neural networks. 5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings.
- Cifuentes, C.A., Pinto, M.J., Céspedes, N., Múnera, M., 2020. Social robots in therapy and care. *Curr. Robot. Reports* 1, 59–74. <https://doi.org/10.1007/s43154-020-00009-2>.
- Cirelli, L.K., Einarson, K.M., Trainor, L.J., 2014. Interpersonal synchrony increases prosocial behavior in infants. *Dev. Sci.* 17, 1003–1011. <https://doi.org/10.1111/desc.12193>.
- Clark, H.H., 1996. *Using Language*. Cambridge University Press, Cambridge.
- Clark, A., 1997. The dynamical challenge. *Cogn. Sci.* 21, 461–481. <https://doi.org/10.1207/s15516709cog2104.3>.
- Clarke, T.J., Bradshaw, M.F., Field, D.T., Hampson, S.E., Rose, D., 2005. The perception of emotion from body movement in point-light displays of interpersonal dialogue. *Perception* 34, 1171–1180. <https://doi.org/10.1068/p5203>.
- Clément, V., 2017. Dancing bodies and Indigenous ontology: what does the haka reveal about the Māori relationship with the Earth? *Trans. Inst. Br. Geogr.* 42, 317–328. <https://doi.org/10.1111/tran.12157>.
- Cohen, L., Khoramshahi, M., Salese, R.N., Bortolon, C., Slowiński, P., Zhai, C., Tsaneva-Atanasova, K., Di Bernardo, M., Capdevielle, D., Marin, L., Schmidt, R.C., Bardy, B. G., Billard, A., Raffard, S., 2017. Influence of facial feedback during a cooperative human-robot task in schizophrenia. *Sci. Rep.* 7. <https://doi.org/10.1038/s41598-017-14773-3>.
- Collins, R., 2004. *Interaction Ritual Chains*. Princeton University Press.
- Costa, T., Crini, M., 2011. Basic emotions: differences in time sequence and functional imaging with low resolution brain electrical tomography (LORETA). *Nat. Preced.* <https://doi.org/10.1038/npre.2011.5566.1.1-1>.
- Costa, T., Cauda, F., Crini, M., Tatu, M.-K., Celeghini, A., de Gelder, B., Tamietto, M., 2014. Temporal and spatial neural dynamics in the perception of basic emotions from complex scenes. *Soc. Cogn. Affect. Neurosci.* 9, 1690–1703. <https://doi.org/10.1093/scan/nst164>.
- Coste, A., Bardy, B.G., Marin, L., 2019. Towards an embodied signature of improvisation skills. *Front. Psychol.* 10, 2441. <https://doi.org/10.3389/fpsyg.2019.02441>.
- Coste, A., Bardy, B.G., Janagi, S., Slowiński, P., Tsaneva-Atanasova, K., Goupil, J.L., Marin, L., 2021. Decoding identity from motion: how motor similarities colour our perception of self and others. *Psychol. Res.* 85, 509–519. <https://doi.org/10.1007/s00426-020-01290-8>.
- Couzin, I.D., Krause, J., Franks, N.R., Levin, S.A., 2005. Effective leadership and decision-making in animal groups on the move. *Nature* 433, 513–516. <https://doi.org/10.1038/nature03236>.
- Cowen, A., Keltner, D., 2017. Self-report captures 27 distinct categories of emotion bridged by continuous gradients. *Proc. National Acad. Sci. United States Am.* 114 (38), E7900–E7909. <https://doi.org/10.1073/pnas.1702247114>.
- Craigero, L., Fadiga, L., Rizzolatti, G., Umiltà, C., 1998. Visuomotor priming. *Vis. cogn.* 5, 109–125. <https://doi.org/10.1080/1713756780>.
- Cummins, F., 2014. Voice, (inter-)subjectivity, and real time recurrent interaction. *Front. Psychol.* 5, 760. <https://doi.org/10.3389/fpsyg.2014.00760>.
- Curioni, A., Minio-Paluello, I., Sachelì, L.M., Candidi, M., Aglioti, S.M., 2017. Autistic traits affect interpersonal motor coordination by modulating strategic use of role-based behavior. *Mol. Autism* 8. <https://doi.org/10.1186/s13229-017-0141-0>.
- Cutting, J.E., Kozlowski, L.T., 1977. Recognizing friends by their walk: gait perception without familiarity cues. *Bull. Psychon.* 9, 353–356. <https://doi.org/10.3758/BF03337021>.
- Czeszumski, A., Eustergerling, S., Lang, A., Menrath, D., Gerstenberger, M., Schuberth, S., Schreiber, F., Rendon, Z.Z., König, P., 2020. Hyperscanning: a valid method to study neural inter-brain underpinnings of social interaction. *Front. Hum. Neurosci.* <https://doi.org/10.3389/fnhum.2020.00039>.
- D'Mello, S.K., Kory, J., 2015. A review and meta-analysis of multimodal affect detection systems. *ACM Comput. Surv.* <https://doi.org/10.1145/2682899>.
- Darwin, C., 1872. *The Expression of Emotions in Man and Animals*. Murray, London, England.
- de Gelder, B., 2009. Why bodies? Twelve reasons for including bodily expressions in affective neuroscience. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 364, 3475–3484. <https://doi.org/10.1098/rstb.2009.0190>.
- de Gelder, B., Poyo Solanas, M., 2021. A computational neuroethology perspective on body and expression perception. *Trends Cogn. Sci.* <https://doi.org/10.1016/j.tics.2021.05.010>.
- De Gelder, B., Snyder, J., Greve, D., Gerard, G., Hadjikhani, N., 2004. Fear fosters flight: a mechanism for fear contagion when perceiving emotion expressed by a whole body. *Proc. Natl. Acad. Sci. U. S. A.* 101, 16701–16706. <https://doi.org/10.1073/pnas.0407042101>.
- Del-Monte, J., Capdevielle, D., Varlet, M., Marin, L., Schmidt, R.C., Salese, R.N., Bardy, B.G., Boulenger, J.P., Gély-Nargeot, M.C., Attal, J., Raffard, S., 2013. Social motor coordination in unaffected relatives of schizophrenia patients: a potential intermediate phenotype. *Front. Behav. Neurosci.* 7. <https://doi.org/10.3389/fnbeh.2013.00137>.
- Demir, K.A., Döven, G., Sezen, B., 2019. Industry 5.0 and Human-Robot Co-working, in: *Procedia Computer Science*. Elsevier B.V., pp. 688–695. <https://doi.org/10.1016/j.procs.2019.09.104>
- di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., Rizzolatti, G., 1992. Understanding motor events: a neurophysiological study. *Exp. Brain Res.* 91, 176–180. <https://doi.org/10.1007/BF00230027>.
- Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., Rowland, J., Michalareas, G., Van Bavel, J.J., Ding, M., Poeppel, D., 2017. Brain-to-Brain synchrony tracks real-world dynamic group interactions in the classroom. *Curr. Biol.* 27, 1375–1380. <https://doi.org/10.1016/j.cub.2017.04.002>.
- Donohue, M.R., Williamson, R.A., Tully, E.C., 2020. Toddlers imitate prosocial demonstrations in bystander but not transgressor contexts. *J. Exp. Child Psychol.* 192. <https://doi.org/10.1016/j.jecp.2019.104776>.
- Dotov, D., Bosnyak, D., Trainor, L.J., 2021. Collective music listening: movement energy is enhanced by groove and visual social cues. *Q. J. Exp. Psychol.* <https://doi.org/10.1177/1747021821991793>, 1747021821991799.
- Dumas, G., Nadel, J., Soussignan, R., Martinerie, J., Garnero, L., 2010. Inter-brain synchronization during social interaction. *PLoS One* 5, e12166. <https://doi.org/10.1371/journal.pone.0012166>.
- Duncan, S., Barrett, L.F., 2007. Affect is a form of cognition: a neurobiological analysis. *Cogn. Emot.* 21, 1184–1211. <https://doi.org/10.1080/02699930701437931>.
- Dunsmore, J.C., Ashley, R.A., Zhou, Y., Swain, D.M., Factor, R.S., Broomell, A.P., Waldron, J.C., Bell, M.A., Scarpa, A., 2019. Marching to the beat of your own drum?: a proof-of-concept study assessing physiological linkage in Autism Spectrum disorder. *Biol. Psychol.* <https://doi.org/10.1016/j.biopsycho.2019.03.001>.

- Durantini, G., Heath, S., Wiles, J., 2017. Social moments: a perspective on interaction for social robotics. *Front. Robot. AI* 4, 24. <https://doi.org/10.3389/frobt.2017.00024>.
- Duranton, C., Gaunet, F., 2016. Behavioural synchronization from an ethological perspective: overview of its adaptive value. *Adapt. Behav.* 24, 181–191. <https://doi.org/10.1177/1059712316644966>.
- Edelman, G.M., Gally, J.A., 2001. Degeneracy and complexity in biological systems. *Proc. Natl. Acad. Sci. U. S. A.* 98, 13763–13768. <https://doi.org/10.1073/pnas.231499798>.
- Eilam, E., 2019. Synchronization: a framework for examining emotional climate in classes. *Palgrave Commun.* 5, 1–11. <https://doi.org/10.1057/s41599-019-0356-0>.
- Einon, D., Potegal, M., 1994. Temper tantrums in young children. *The Dynamics of Aggression: Biological and Social Processes in Dyads and Groups*. Lawrence Erlbaum Associates, Inc, Hillsdale, NJ, US, pp. 157–194.
- Ekman, P., 1992. An argument for basic emotions. *Cogn. Emot.* 6, 169–200. <https://doi.org/10.1080/02699939208411068>.
- Ekman, P., Friesen, W., 1978. Facial action coding system: a technique for the measurement of facial movement. Palo Alto. Differences among Unpleasant Feelings. *Motivation and Emotion*, pp. 271–302.
- Ekman, P., Rosenberg, E.L., 2005. What the face reveals: basic and applied studies of spontaneous expression using the facial action coding system (FACS). *Series in Affective Science*, 2nd ed. Oxford University Press, Oxford and New York.
- Elprama, B., El Makrini, I., Jacobs, A., 2016. Acceptance of collaborative robots by factory workers: a pilot study on the importance of social cues of anthropomorphic robots. *International Symposium on Robot and Human Interactive Communication*.
- Farkas, I., Helbing, D., Vicsek, T., 2002. Mexican waves in an excitable medium. *Nature* 419, 131–132. <https://doi.org/10.1038/419131a>.
- Fasola, J., Mataric, M.J., 2012. Using Socially Assistive Human–Robot Interaction to Motivate Physical Exercise for Older Adults. *Proc. IEEE* 100, 2512–2526. <https://doi.org/10.1109/JPROC.2012.2200539>.
- Fehr, R., Russell, J.A., 1984. Concept of emotion viewed from a prototype perspective. *J. Exp. Psychol. Gen.* 113, 464–486. <https://doi.org/10.1037/0096-3445.113.3.464>.
- Feil-Seifer, D., Mataric, M., 2011. Automated detection and classification of positive vs. Negative robot interactions with children with autism using distance-based features. In: Billard, A., Kahn, P., Adams, J.A., Trafton, G. (Eds.), *Proceedings of the 6th International Conference on Human-Robot Interaction - HRI' 11*. ACM Press, New York, New York, USA, p. 323. <https://doi.org/10.1145/1957656.1957785>.
- Feldman, R., 2007. On the origins of background emotions: from affect synchrony to symbolic expression. *Emotion*. <https://doi.org/10.1037/1528-3542.7.3.601>.
- Feldman, R., 2012. Parent-infant synchrony: a biobehavioral model of mutual influences in the formation of affiliative bonds. *Monogr. Soc. Res. Child Dev.* 77, 42–51. <https://doi.org/10.1111/j.1540-5834.2011.00660.x>.
- Feldman, R., Magori-Cohen, R., Galili, G., Singer, M., Louzoun, Y., 2011. Mother and infant coordinate heart rhythms through episodes of interaction synchrony. *Infant Behav. Dev.* 34, 569–577. <https://doi.org/10.1016/j.infbeh.2011.06.008>.
- Feldman Barrett, L., Finlay, B.L., 2018. Concepts, goals and the control of survival-related behaviors. *Curr. Opin. Behav. Sci.* <https://doi.org/10.1016/j.cobeha.2018.10.001>.
- Feng, K., Chaspari, T., 2020. A review of generalizable transfer learning in automatic emotion recognition. *Front. Comput. Sci. China* 2, 9. <https://doi.org/10.3389/fcomp.2020.00009>.
- Feniger-Schaal, R., Hart, Y., Lotan, N., Koren-Karie, N., Noy, L., 2018. The body speaks: using the mirror game to link attachment and non-verbal behavior. *Front. Psychol.* 9, 1560. <https://doi.org/10.3389/fpsyg.2018.01560>.
- Ferrari, P.F., 2014. The neuroscience of social relations. A comparative-based approach to empathy and to the capacity of evaluating others' action value. *Behaviour* 151, 297–313. <https://doi.org/10.1163/1568539X-00003152>.
- Fessler, D.M.T., Holbrook, C., 2014. Marching into battle: synchronized walking diminishes the conceptualized formidability of an antagonist in men. *Biol. Lett.* 10, 20140592. <https://doi.org/10.1098/rsbl.2014.0592>.
- Festante, F., Ferrari, P.F., Thorpe, S.G., Buchanan, R.W., Fox, N.A., 2020. Intranasal oxytocin enhances EEG mu rhythm desynchronization during execution and observation of social action: an exploratory study. *Psychoneuroendocrinology* 111. <https://doi.org/10.1016/j.psyneuen.2019.104467>.
- Fischer, K., 2019. Why collaborative robots must be social (and even emotional) actors. *Techné Res. Philos. Technol.* 23, 270–289. <https://doi.org/10.5840/techné20191120104>.
- Fischinger, D., Einramhof, P., Papoutsakis, K., Wohlkinger, W., Mayer, P., Panek, P., Hofmann, S., Koertner, T., Weiss, A., Argyros, A., Vincze, M., 2016. Hobbit, a care robot supporting independent living at home: first prototype and lessons learned. *Rob. Auton. Syst.* 75, 60–78. <https://doi.org/10.1016/j.robot.2014.09.029>.
- Fitzpatrick, P., Frazier, J.A., Cochran, D.M., Mitchell, T., Coleman, C., Schmidt, R.C., 2016. Impairments of social motor synchrony evident in autism spectrum disorder. *Front. Psychol.* 7. <https://doi.org/10.3389/fpsyg.2016.01323>.
- Fitzpatrick, P., Romero, V., Amaral, J.L., Duncan, A., Barnard, H., Richardson, M.J., Schmidt, R.C., 2017. Social motor synchronization: insights for understanding social behavior in autism. *J. Autism Dev. Disord.* 47, 2092–2107. <https://doi.org/10.1007/s10803-017-3124-2>.
- Forbes, P.A.G., Pan, X., Antonia, A.F., 2016. Reduced mimicry to virtual reality avatars in autism spectrum disorder. *J. Autism Dev. Disord.* 46, 3788–3797. <https://doi.org/10.1007/s10803-016-2930-2>.
- Fraune, M.R., Sherrin, S., Sabanovic, S., Smith, E.R., 2019. Is human-robot interaction more competitive between groups than between individuals? *ACM/IEEE International Conference on Human-Robot Interaction*. IEEE Computer Society 104–113. <https://doi.org/10.1109/HRI.2019.8673241>.
- Frijda, N.H., 2007. *The Laws of emotion, the Laws of Emotion*. Lawrence Erlbaum Associates Publishers, Mahwah, NJ, US.
- Frijda, N.H., Parrott, W.G., 2011. Basic emotions or ur-emotions? *Emot. Rev.* 3, 406–415. <https://doi.org/10.1177/1754073911410742>.
- Frijda, N.H., Kuipers, P., ter Schure, E., 1989. Relations among emotion, appraisal, and emotional action readiness. *J. Pers. Soc. Psychol.* 57, 212–228. <https://doi.org/10.1037/0022-3514.57.2.212>.
- Fujiwara, K., Kimura, M., Daibo, I., 2019. Gender differences in synchrony: females in sync during unstructured dyadic conversation. *Eur. J. Soc. Psychol.* 69, 518. <https://doi.org/10.1002/ejsp.2587>.
- Fulceri, F., Tonacci, A., Lucaferro, A., Apicella, F., Narzisi, A., Vincenti, G., Muratori, F., Contaldo, A., 2018. Interpersonal motor coordination during joint actions in children with and without autism spectrum disorder: The role of motor information. *Res. Dev. Disabil.* 80, 13–23. <https://doi.org/10.1016/j.ridd.2018.05.018>.
- Gallese, V., 2007. Before and below 'theory of mind': embodied simulation and the neural correlates of social cognition. *Philos. Trans. R. Soc. B Biol. Sci.* 362, 659–669. <https://doi.org/10.1098/rstb.2006.2002>.
- Gallese, V., Keysers, C., Rizzolatti, G., 2004. A unifying view of the basis of social cognition. *Trends Cogn. Sci.* 8, 396–403. <https://doi.org/10.1016/j.tics.2004.07.002>.
- Gallup, A.C., Hale, J.J., Sumpter, D.J.T., Garnier, S., Kacelnik, A., Krebs, J.R., Couzin, I. D., 2012. Visual attention and the acquisition of information in human crowds. *Proc. Natl. Acad. Sci. U. S. A.* 109, 7245–7250. <https://doi.org/10.1073/pnas.1116141109>.
- Gao, Y., Bianchi-Berthouze, N., Meng, H., 2012. What does touch tell us about emotions in touchscreen-based gameplay? *ACM Trans. Comput. Interact.* 19, 1–30. <https://doi.org/10.1145/2395131.2395138>.
- Gashi, S., Di Lascio, E., Santini, S., 2019. Using unobtrusive wearable sensors to measure the physiological synchrony between presenters and audience members. *ACM Interactive, Mobile, Wearable Ubiquitous Technol.* 3, 1–19. <https://doi.org/10.1145/3314400>.
- Gehle, R., Pitsch, K., Wrede, S., 2014. Signaling trouble in robot-to-group interaction. Emerging visitor dynamics with a museum guide robot. In: Kuzuoka, H., Ono, T., Imai, M., Young, J.E. (Eds.), *Proceedings of the Second International Conference on Human-Agent Interaction - HAI' 14*. ACM Press, New York, New York, USA, pp. 361–368. <https://doi.org/10.1145/2658861.2658887>.
- Genschow, O., Van Den Bossche, S., Cracco, E., Bardi, L., Rigoni, D., Brass, M., 2017. Mimicry and automatic imitation are not correlated. *PLoS One* 12. <https://doi.org/10.1371/journal.pone.0183784>.
- Gibson, J.J., James, J., 1979. *The Ecological Approach to Visual Perception*. Houghton Mifflin, Boston.
- Gil, S., Droit-Volet, S., 2012. Emotional time distortions: the fundamental role of arousal. *Cogn. Emot.* 26, 847–862. <https://doi.org/10.1080/02699931.2011.625401>.
- Giles, H., Taylor, D.M., Bourhis, R., 1973. Towards a theory of interpersonal accommodation through language: some Canadian data. *Lang. Soc.* 2, 177–192. <https://doi.org/10.1017/S0047404500000701>.
- Giménez-Dasí, M., Quintanilla, L., Lucas-Molina, B., Sarmento-Henrique, R., 2020. Six weeks of confinement: psychological effects on a sample of children in early childhood and primary education. *Front. Psychol.* 11. <https://doi.org/10.3389/fpsyg.2020.590463>.
- Glowinski, D., Mancini, M., Cowie, R., Camurri, A., Chiorri, C., Doherty, C., 2013. The movements made by performers in a skilled quartet: a distinctive pattern, and the function that it serves. *Front. Psychol.* 4, 841. <https://doi.org/10.3389/fpsyg.2013.00841>.
- Gnecco, G., Badino, L., Camurri, A., D'Ausilio, A., Fadiga, L., Glowinski, D., Sanguineti, M., Varni, G., Volpe, G., 2013. Towards automated analysis of joint music performance in the orchestra. *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICTST*. Springer, Berlin, Heidelberg, pp. 120–127. https://doi.org/10.1007/978-3-642-37982-6_16.
- Goldenberg, A., Halperin, E., van Zomeren, M., Gross, J.J., 2016. The process model of group-based emotion: integrating intergroup emotion and emotion regulation perspectives. *Pers. Soc. Psychol. Rev.* 20, 118–141. <https://doi.org/10.1177/1088868315581263>.
- Goldstein, P., Losin, E.A.R., Anderson, S.R., Schelkun, V.R., Wager, T.D., 2020. Clinician-patient movement synchrony mediates social group effects on interpersonal trust and perceived pain. *J. Pain* 21, 1160–1174. <https://doi.org/10.1016/j.jpain.2020.03.001>.
- Goulart, C., Valadao, C., Delisle-Rodriguez, D., Funayama, D., Favarato, A., Baldo, G., Binotte, V., Caldeira, E., Bastos-Filho, T., 2019. Visual and thermal image processing for facial specific landmark detection to infer emotions in a child-robot interaction. *Sensors* 19, 2844. <https://doi.org/10.3390/s19132844>.
- Gouldner, A.W., 1960. The norm of reciprocity: a preliminary statement. *Am. Sociol. Rev.* 25, 161. <https://doi.org/10.2307/2092623>.
- Grammer, K., Kruck, K.B., Magnusson, M.S., 1998. The courtship dance: patterns of nonverbal synchronization in opposite-sex encounters. *J. Nonverbal Behav.* 22, 1–27. <https://doi.org/10.1023/A:1022986608835>.
- Green, M.F., Horan, W.P., Lee, J., 2015. Social cognition in schizophrenia. *Nat. Rev. Neurosci.* <https://doi.org/10.1038/nrn4005>.
- Grisey, G., 1987. Tempus ex Machina: a composer's reflections on musical time. *Contemp. Music Rev.* 2, 239–275. <https://doi.org/10.1080/07494468708567060>.
- Gurcan, I., Nguyen, H.V., 2019. Surgical activities recognition using multi-scale recurrent networks. In: *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*. Institute of Electrical and Electronics Engineers Inc, pp. 2887–2891. <https://doi.org/10.1109/ICASSP.2019.8683849>.
- Hadjikhani, A., Thilenius, P., 2005. The impact of horizontal and vertical connections on relationships' commitment and trust. *J. Bus. Ind. Mark.* 20, 136–147. <https://doi.org/10.1108/08858620510592759>.

- Hainsworth, F.R., 1989. Wing movements and positioning for aerodynamic benefit by Canada geese flying in formation. *Can. J. Zool.* 67, 585–589. <https://doi.org/10.1139/z89-084>.
- Hale, J., Burgoon, J.K., 1984. Models of reactions to changes in nonverbal immediacy. Dyadic Power Theory development View project Game-Based Learning View project. Researchgate.Net 8, 287–314. <https://doi.org/10.1007/BF00985984>.
- Hamdan, A.N., Wagner, A., Voelker, S., Steimle, J., Borchers, J., 2019. Springlets: expressive, flexible and silent on-skin tactile interfaces. *Conference on Human Factors in Computing Systems - Proceedings. Association for Computing Machinery*, pp. 1–14.
- Hammal, Z., Cohn, J.F., George, D.T., 2014. Interpersonal coordination of headmotion in distressed couples. *IEEE Trans. Affect. Comput.* 5, 155–167. <https://doi.org/10.1109/TAFFC.2014.2326408>.
- Han, J.H., Kim, D.H., Kim, J.W., 2009. Physical learning activities with a teaching assistant robot in elementary school music class. NCM 2009 - 5th International Joint Conference on INC, IMS, and IDC 1406–1410. <https://doi.org/10.1109/NCM.2009.407>.
- Hari, R., Kujala, M.V., 2009. Brain basis of human social interaction: from concepts to brain imaging. *Physiol. Rev.* <https://doi.org/10.1152/physrev.00041.2007>.
- Hasson, U., Frith, C.D., 2016. Mirroring and beyond: coupled dynamics as a generalized framework for modelling social interactions. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 371 <https://doi.org/10.1098/rstb.2015.0366>.
- Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., Malach, R., 2004. Intersubject synchronization of cortical activity during natural vision. *Science (80-)* 303, 1634–1640. <https://doi.org/10.1126/science.1089506>.
- Hasson, U., Ghazanfar, A.A., Galantucci, B., Garrod, S., Keysers, C., 2012. Brain-to-brain coupling: a mechanism for creating and sharing a social world. *Trends Cogn. Sci.* <https://doi.org/10.1016/j.tics.2011.12.007>.
- Hatfield, E., Cacioppo, J.T., Rapson, R.L., 1993. Emotional contagion. *Curr. Dir. Psychol. Sci.* 2, 96–100. <https://doi.org/10.1111/1467-8721.ep10770953>.
- Hatfield, E., Rapson, R.L., Le, Y.-C.L., 2011. Emotional contagion and empathy. In: Decety, J. (Ed.), *The Social Neuroscience of Empathy, Social Neuroscience Series*. MIT Press, Cambridge, Mass, pp. 19–30. <https://doi.org/10.7551/mitpress/9780262012973.003.0003>.
- Heinskou, M.B., Liebst, L.S., 2016. On the elementary neural forms of micro-interactional rituals: integrating autonomic nervous system functioning into interaction ritual theory. *Sociol. Forum Randolph N (Randolph N J)* 31, 354–376. <https://doi.org/10.1111/sofcl.12248>.
- Hernandez, J., Paredes, P., Roseway, A., Czerwinski, M., 2014. Under pressure. In: Jones, M., Palanque, P., Schmidt, A., Grossman, T. (Eds.), *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems - CHI '14*. ACM Press, New York, New York, USA, pp. 51–60. <https://doi.org/10.1145/2556288.2557165>.
- Hess, U., Fischer, A., 2013. Emotional mimicry as social regulation. *Pers. Soc. Psychol. Rev.* 17, 142–157. <https://doi.org/10.1177/1088868312472607>.
- Heyes, C., 2021. Imitation. *Curr. Biol.* 31, R228–R232. <https://doi.org/10.1016/j.cub.2020.11.071>.
- Heyes, C., Catmur, C., 2021. What Happened to Mirror Neurons? <https://doi.org/10.1177/1745691621990638> <https://doi.org/10.1177/1745691621990638>.
- Hilt, P., Badino, L., D'Ausilio, A., Gualtieri, V., Serà, T., Luciano, F., Antonio, C., 2019. Multi-layer adaptation of group coordination in musical ensembles. *Sci. Rep.* 9, 1–10. <https://doi.org/10.1038/s41598-019-42395-4>.
- Hinoshita, W., Ogata, T., Kozima, H., Kanda, H., Takahashi, T., Okuno, H.G., 2009a. Emergence of evolutionary interaction with voice and motion between two robots using RNN. In: 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, Piscataway, NJ, pp. 4186–4192. <https://doi.org/10.1109/IROS.2009.5353887>.
- Hinoshita, W., Ogata, T., Kozima, H., Kanda, H., Takahashi, T., Okuno, H.G., 2009b. Emergence of evolutionary interaction with voice and motion between two robots using RNN. 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2009 4186–4192. <https://doi.org/10.1109/IROS.2009.5353887>.
- Hirsch, J., Zhang, X., Noah, J.A., Ono, Y., 2017. Frontal temporal and parietal systems synchronize within and across brains during live eye-to-eye contact. *Neuroimage* 157, 314–330. <https://doi.org/10.1016/j.neuroimage.2017.06.018>.
- Hobson, J.A., Hobson, P.R., 2007. Identification: the missing link between joint attention and imitation? *Dev. Psychopathol.* 19, 411–431. <https://doi.org/10.1017/S0954579407070204>.
- Hoehl, S., Fairhurst, M., Schirmer, A., 2021. Interactional synchrony: signals, mechanisms and benefits. *Soc. Cogn. Affect. Neurosci.* 16, 5–18. <https://doi.org/10.1093/scan/nsaa024>.
- Hoemann, K., Feldman Barrett, L., 2019. Concepts dissolve artificial boundaries in the study of emotion and cognition, uniting body, brain, and mind. *Cogn. Emot.* 33, 67–76. <https://doi.org/10.1080/02699931.2018.1535428>.
- Hoey, J., Schröder, T., Althof, A., 2016. Affect control processes: intelligent affective interaction using a partially observable Markov decision process. *Artif. Intell.* 230, 134–172. <https://doi.org/10.1016/j.artint.2015.09.004>.
- Hogeveen, J., Chartrand, T.L., Obhi, S.S., 2015. Social mimicry enhances mu-suppression during action observation. *Cereb. Cortex* 25, 2076–2082. <https://doi.org/10.1093/cercor/bhu016>.
- Holt-Lunstad, J., Smith, T.B., Layton, J.B., 2010. Social relationships and mortality risk: a meta-analytic review. *PLoS Med.* 7, e1000316. <https://doi.org/10.1371/journal.pmed.1000316>.
- Hommel, B., 1997. Toward an action-concept model of stimulus-response compatibility. *Adv. Psychol.* 118, 281–320. [https://doi.org/10.1016/S0166-4115\(97\)80041-6](https://doi.org/10.1016/S0166-4115(97)80041-6).
- Hommel, B., 1998. Automatic stimulus-response translation in Dual-Task performance. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 1368–1384. <https://doi.org/10.1037/0096-1523.24.5.1368>.
- Hood, D., Lemaignan, S., Dillenbourg, P., 2015. When children teach a robot to write: an Autonomous teachable humanoid which uses simulated handwriting. In: *ACM/IEEE International Conference on Human-Robot Interaction*. IEEE Computer Society, New York, NY, USA, pp. 83–90. <https://doi.org/10.1145/2696454.2696479>.
- Horan, W.P., Iacuboni, M., Cross, K.A., Korb, A., Lee, J., Nori, P., Quintana, J., Wynn, J. K., Green, M.F., 2014. Self-reported empathy and neural activity during action imitation and observation in schizophrenia. *Neuroimage Clin.* 5, 100–108. <https://doi.org/10.1016/j.nicl.2014.06.006>.
- Hove, M.J., Risen, J.L., 2009. It's all in the timing: interpersonal synchrony increases affiliation. *Soc. Cogn.* 27, 949–960. <https://doi.org/10.1521/soco.2009.27.6.949>.
- Huisin' t Veld, E.M.J., Van Boxtel, G.J.M., de Gelder, B., 2014b. The Body Action Coding System I: muscle activations during the perception and expression of emotion. *Soc. Neurosci.* 9, 249–264. <https://doi.org/10.1080/17470919.2014.890668>.
- Huisin' t Veld, E.M.J., van Boxtel, G.J.M., de Gelder, B., 2014a. The body action coding system II: Muscle activations during the perception and expression of emotion. *Front. Behav. Neurosci.* 8, 1–13. <https://doi.org/10.3389/fnbeh.2014.00330>.
- Huisman, G., Darriba Frederiks, A., Van Dijk, B., Hevlen, D., Kroese, B., 2013. The TaSSt: tactile sleeve for social touch. 2013 World Haptics Conference, WHC 2013 211–216. <https://doi.org/10.1109/WHC.2013.6548410>.
- Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J.C., 2005. Grasping the intentions of others with one's own mirror neuron system. In: *Ashe, J. (Ed.), PLoS Biology. Public Library of Science*, pp. 0529–0535. <https://doi.org/10.1371/journal.pbio.0030079>.
- Iqbal, T., Riek, L.D., 2019. Human-robot teaming: approaches from joint action and dynamical systems. *Humanoid Robotics: A Reference*. Springer, Netherlands, pp. 2293–2312. https://doi.org/10.1007/978-94-007-6046-2_137.
- Izard, C.E., Carroll, E., 1971. *The Face of Emotion*. Appleton-Century-Crofts.
- Jaderberg, M., Czarnecki, W.M., Dunning, I., Marris, L., Lever, G., Castañeda, A.G., Beattie, C., Rabinowitz, N.C., Morcos, A.S., Ruderman, A., Sonnerat, N., Green, T., Deason, L., Leibo, J.Z., Silver, D., Hassabis, D., Kavukcuoglu, K., Graepel, T., 2019. Human-level performance in 3D multiplayer games with population-based reinforcement learning. *Science (80-)* 364, 859–865. <https://doi.org/10.1126/science.aau6249>.
- James, W., 1884. What is an emotion? *Mind* 9, 188–205. <https://doi.org/10.1093/mind/os-IX.34.188>.
- Jansen, A.S.P., Van Nguyen, X., Karpitskiy, V., Mettenleiter, T.C., Loewy, A.D., 1995. Central command neurons of the sympathetic nervous system: Basis of the fight-or-flight response. *Science (80-)* 270, 644–646. <https://doi.org/10.1126/science.270.5236.644>.
- Jarrassé, N., Charalambous, T., Burdet, E., 2012. A framework to describe, analyze and generate interactive motor behaviors. *PLoS One* 7. <https://doi.org/10.1371/journal.pone.0049945>.
- Jasmin, K.M., McGettigan, C., Agnew, Z.K., Lavan, N., Josephs, O., Cummins, F., Scott, S. K., 2016. Cohesion and joint speech: right hemisphere contributions to synchronized vocal production. *J. Neurosci.* 36, 4669–4680. <https://doi.org/10.1523/JNEUROSCI.4075-15.2016>.
- Jastorff, J., Huang, Y.-A., Giese, M.A., Vandenbulcke, M., 2015. Common neural correlates of emotion perception in humans. *Hum. Brain Mapp.* 36, 4184–4201. <https://doi.org/10.1002/hbm.22910>.
- Jiménez, L., Lorda, M.J., Méndez, C., 2014. Emulation and mimicry in school students with typical development and with high functioning autism. *J. Autism Dev. Disord.* 44, 1597–1608. <https://doi.org/10.1007/s10803-013-2027-0>.
- Johansson, G., 1973. Visual perception of biological motion and a model for its analysis. *Percept. Psychophys.* 14, 201–211. <https://doi.org/10.3758/BF03212378>.
- Jola, C., McAleer, P.M.A.P., Grosbras, M.H., Love, S.A., Morison, G., Pollick, F.E., 2013. Uni- and multisensory brain areas are synchronised across spectators when watching unedited dance recordings. *Iperception* 4, 265–284. <https://doi.org/10.1068/i0536>.
- Kachouie, R., Sedighadel, S., Khosla, R., Chu, M.T., 2014. Socially assistive robots in elderly care: a mixed-method systematic literature review. *Int. J. Hum. Interact.* 30, 369–393. <https://doi.org/10.1080/10447318.2013.873278>.
- Kafka, G., 1950. Über urafekte. *Acta Psychol. (Amst.)* 7, 256–278. [https://doi.org/10.1016/0001-6918\(50\)90018-7](https://doi.org/10.1016/0001-6918(50)90018-7).
- Kalegina, A., Schroeder, G., Allchin, A., Berlin, K., Cakmak, M., 2018. Characterizing the design space of rendered robot faces. In: *ACM/IEEE International Conference on Human-Robot Interaction*. IEEE Computer Society, New York, NY, USA, pp. 96–104. <https://doi.org/10.1145/3171221.3171286>.
- Kaplan, H.S., Thula, O.S., Khoss, N., Zimmer, M., 2019. Nested neuronal dynamics orchestrate a behavioral hierarchy across timescales (review). *Neuron* 1–15. <https://doi.org/10.1016/j.neuron.2019.10.037>.
- Kellerman, A.M., Schwichtenberg, A.J., Abu-Zhaya, R., Miller, M., Young, G.S., Ozonoff, S., 2020. Dyadic synchrony and responsiveness in the first year: associations with autism risk. *Autism Res.* 13, 2190–2201. <https://doi.org/10.1002/aur.2373>.
- Kelly, J.R., Barsade, S.G., 2001. Mood and emotions in small groups and work teams. *Organ. Behav. Hum. Decis. Process.* 86, 99–130. <https://doi.org/10.1006/obhd.2001.2974>.
- Kelso, J.A.S., 2021. Unifying large- and small-scale theories of coordination. *Entropy* 23, 537. <https://doi.org/10.3390/e23050537>.
- Khalil, R.A., Jones, E., Babar, M.I., Jan, T., Zafar, M.H., Alhussain, T., 2019. Speech emotion recognition using deep learning techniques: a review. *IEEE Access* 7, 117327–117345. <https://doi.org/10.1109/ACCESS.2019.2936124>.

- Kiverstein, J., Miller, M., 2015. The embodied brain: towards a radical embodied cognitive neuroscience. *Front. Hum. Neurosci.* 9, 237. <https://doi.org/10.3389/fnhum.2015.00237>.
- Kleinginna, P.R., Kleinginna, A.M., 1981. A categorized list of motivation definitions, with a suggestion for a consensual definition. *Motiv. Emot.* 5, 263–291. <https://doi.org/10.1007/BF00993889>.
- Kleinsmith, A., Bianchi-Berthouze, N., 2013. Affective body expression perception and recognition: a survey. *IEEE Trans. Affect. Comput.* 4, 15–33. <https://doi.org/10.1109/T-AFFC.2012.16>.
- Knoblich, G., Butterfill, S., Sebanz, N., 2011. Psychological Research on Joint Action. Theory and Data, Psychology of Learning and Motivation - Advances in Research and Theory. Academic Press. <https://doi.org/10.1016/B978-0-12-385527-5.00003-6>.
- Kochigami, K., Okada, K., Inaba, M., 2018. Does an introduction of a person in a group by a robot have a positive effect on people's communication? RO-MAN 2018 - 27th IEEE International Symposium on Robot and Human Interactive Communication. Institute of Electrical and Electronics Engineers Inc. 692–698. <https://doi.org/10.1109/ROMAN.2018.8525686>.
- Koehne, S., Behrends, A., Fairhurst, M.T., Dziobek, I., 2016a. Fostering social cognition through an imitation- and synchronization-based Dance/Movement intervention in adults with autism Spectrum disorder: a controlled proof-of-Concept study. *Psychother. Psychosom.* 85, 27–35. <https://doi.org/10.1159/000441111>.
- Koehne, S., Hatri, A., Cacioppo, J.T., Dziobek, I., 2016b. Perceived interpersonal synchrony increases empathy: insights from autism spectrum disorder. *Cognition* 146, 8–15. <https://doi.org/10.1016/j.cognition.2015.09.007>.
- Koehne, S., Schmidt, M.J., Dziobek, I., 2016c. The role of interpersonal movement synchronisation in empathic functions: insights from Tango Argentino and Capoeira. *Int. J. Psychol.* 51, 318–322. <https://doi.org/10.1002/ijop.12213>.
- Koelstra, S., Mühl, C., Soleymani, M., Lee, J.S., Yazdani, A., Ebrahimi, T., Pun, T., Nijholt, A., Patras, I., 2012. DEAP: a database for emotion analysis; using physiological signals. *IEEE Trans. Affect. Comput.* 3, 18–31. <https://doi.org/10.1109/T-AFFC.2011.15>.
- Komulainen, E., Meskanen, K., Lipsanen, J., Lahti, J.M., Jylhä, P., Melartin, T., Wichers, M., Isometsä, E., Ekelund, J., 2014. The effect of personality on daily life emotional processes. *PLoS One* 9, e110907. <https://doi.org/10.1371/journal.pone.0110907>.
- Koole, S.L., Tschacher, W., 2016. Synchrony in psychotherapy: a review and an integrative framework for the therapeutic alliance. *Front. Psychol.* 7, 862. <https://doi.org/10.3389/fpsyg.2016.00862>.
- Korb, S., Malsert, J., Strathearn, L., Vuilleumier, P., Niedenthal, P., 2016. Sniff and mimic - Intranasal oxytocin increases facial mimicry in a sample of men. *Horm. Behav.* 84, 64–74. <https://doi.org/10.1016/j.yhbeh.2016.06.003>.
- Kourtis, D., Sebanz, N., Knoblich, G., 2010. Favouritism in the motor system: social interaction modulates action simulation. *Biol. Lett.* 6, 758–761. <https://doi.org/10.1098/rsbl.2010.0478>.
- Kozima, H., Nakagawa, C., Yasuda, Y., 2005. Interactive Robots for Communication-care: a Case-study in Autism Therapy, pp. 341–346. <https://doi.org/10.1109/ROMAN.2005.1513802>.
- Kragness, H.E., Cirelli, L.K., 2021. A syncing feeling: reductions in physiological arousal in response to observed social synchrony. *Soc. Cogn. Affect. Neurosci.* 16, 177–184. <https://doi.org/10.1093/scan/nsaa116>.
- Kruppa, J.A., Reindl, V., Gerloff, C., Oberwiesland Weiss, E., Prinz, J., Herpertz-Dahlmann, B., Konrad, K., Schulte-Rüther, M., 2021. Brain and motor synchrony in children and adolescents with ASD - A fNIRS hyperscanning study. *Soc. Cogn. Affect. Neurosci.* 16, 103–116. <https://doi.org/10.1093/scan/nsaa092>.
- Kühn, S., Müller, B.C.N., van Baaren, R.B., Wietzker, A., Dijksterhuis, A., Brass, M., 2010. Why do I like you when you behave like me? Neural mechanisms mediating positive consequences of observing someone being imitated. *Soc. Neurosci.* 5, 384–392. <https://doi.org/10.1080/17470911003633750>.
- Küntzler, T., Höfling, T.T.A., Alpers, G.W., 2021. Automatic facial expression recognition in standardized and non-standardized emotional expressions. *Front. Psychol.* 12, 627561. <https://doi.org/10.3389/fpsyg.2021.627561>.
- Kuppens, T., Zyerby, V.Y., Dandache, S., Fischer, A.H., van der Schalk, J., 2013. Social identity salience shapes group-based emotions through group-based appraisals. *Cogn. Emot.* 27, 1359–1377. <https://doi.org/10.1080/02699931.2013.785387>.
- Kupper, Z., Ramseyer, F., Hoffmann, H., Tschacher, W., 2015. Nonverbal synchrony in social interactions of patients with schizophrenia indicates socio-communicative deficits. *PLoS One* 10. <https://doi.org/10.1371/journal.pone.0145882>.
- Laban, R., Ullmann, L., 1988. *The Mastery of Movement*. Northcote House, Plymouth, England.
- Lakens, D., Stel, M., 2011. If they move in sync, they must feel in sync: Movement synchrony leads to attributions of rapport and entitativity. *Soc. Cogn.* 29, 1–14. <https://doi.org/10.1521/soco.2011.29.1.1>.
- Landa, R.J., Holman, K.C., O'Neill, A.H., Stuart, E.A., 2011. Intervention targeting development of socially synchronous engagement in toddlers with autism spectrum disorder: a randomized controlled trial. *J. Child Psychol. Psychiatry Allied Discip.* 52, 13–21. <https://doi.org/10.1111/j.1469-7610.2010.02288.x>.
- Ledoux, J.E., Brown, R., 2017. A higher-order theory of emotional consciousness. *Proc. Natl. Acad. Sci. U. S. A.* 114, E2016–E2025. <https://doi.org/10.1073/pnas.1619316114>.
- LeDoux, J.E., 2014. Coming to terms with fear. *Proc. Natl. Acad. Sci. U. S. A.* <https://doi.org/10.1073/pnas.1400335111>.
- Leithinger, D., Follmer, S., Olwal, A., Ishii, H., 2014. Physical telepresence: shape capture and display for embodied, computer-mediated remote collaboration. *UIST 2014 - Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*. Association for Computing Machinery, Inc, New York, NY, USA 461–470. <https://doi.org/10.1145/2642918.2647377>.
- Lenggenhager, B., Tadi, T., Metzinger, T., Blanke, O., 2007. Video ergo sum: manipulating bodily self-consciousness. *Science (80-)* 317, 1096–1099. <https://doi.org/10.1126/science.1143439>.
- Leslie, K.R., Johnson-Frey, S.H., Grafton, S.T., 2004. Functional imaging of face and hand imitation: towards a motor theory of empathy. *Neuroimage* 21, 601–607. <https://doi.org/10.1016/j.neuroimage.2003.09.038>.
- Li, S., Deng, W., 2020. Deep facial expression recognition: a survey. *IEEE Trans. Affect. Comput.* <https://doi.org/10.1109/T-AFFC.2020.2981446>.
- Liebst, L.S., 2019. Exploring the sources of collective effervescence: a multilevel study. *Sociol. Sci.* 6, 27–42. <https://doi.org/10.15195/V6.A2>.
- Liu, H., Wang, L., 2018. Gesture recognition for human-robot collaboration: a review. *Int. J. Ind. Ergon.* 68, 355–367. <https://doi.org/10.1016/j.ergon.2017.02.004>.
- Lorenzo, G., Lledó, A., Arráez-Verá, G., Lorenzo-Lledó, A., 2019. The application of immersive virtual reality for students with ASD: a review between 1990–2017. *Educ. Inf. Technol.* 24, 127–151. <https://doi.org/10.1007/s10639-018-9766-7>.
- Loula, F., Prasad, S., Harber, K., Shiffar, M., 2005. Recognizing people from their movement. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 210–220. <https://doi.org/10.1037/0096-1523.31.1.210>.
- Lozano-Goupil, J., Bardy, B., Marin, L., 2021. Toward an individual emotional motor signature. *Front. Psychol.* 12, 1559. <https://doi.org/10.3389/fpsyg.2021.647704>.
- Lukes, S., 1975. Political ritual and social integration. *Sociology* 9, 289–308. <https://doi.org/10.1177/003803857500900205>.
- Lumsden, J., Miles, L.K., Neil Macrae, C., 2014. Sync or sink? Interpersonal synchrony impacts self-esteem. *Front. Psychol.* 5. <https://doi.org/10.3389/fpsyg.2014.01064>.
- Lutz, W., Prinz, J.N., Schwartz, B., Paulick, J., Schoenherr, D., Deisenhofer, A.K., Terhürne, P., Boyle, K., Altmann, U., Strauß, B., Rafaeli, E., Atzil-Slonim, D., Bar-Kalifa, E., Rubel, J., 2020. Patterns of early change in interpersonal problems and their relationship to nonverbal synchrony and multidimensional outcome. *J. Couns. Psychol.* 67, 449–461. <https://doi.org/10.1037/cou0000376>.
- Ma, Q., Chen, E., Lin, Z., Yan, J., Yu, Z., Ng, W.W.Y., 2019. Convolutional multiresolution echo state network. *IEEE Trans. Cybern.* 1–13. <https://doi.org/10.1109/tycb.2019.2919648>.
- Macpherson, M.C., Marie, D., Schön, S., Miles, L.K., 2020. Evaluating the interplay between subclinical levels of mental health symptoms and coordination dynamics. *Br. J. Psychol.* 111, 782–804. <https://doi.org/10.1111/bjop.12426>.
- Malatesta, G., Marzoli, D., Rapino, M., Tommasi, L., 2019. The left-cradling bias and its relationship with empathy and depression. *Sci. Rep.* 9, 1–9. <https://doi.org/10.1038/s41598-019-42539-6>.
- Malcolm, N., 1953. Direct perception. *Philos. Q.* 3, 301–316. <https://doi.org/10.2307/2217098>.
- Mann, R.P., Faria, J., Sumpter, D.J.T., Krause, J., 2013. The dynamics of audience applause. *J. R. Soc. Interface* 10, 20130466. <https://doi.org/10.1098/rsif.2013.0466>.
- Manschreck, T.C., 1981. Deficient motor synchrony in schizophrenia. *J. Abnorm. Psychol.* 90, 321–328. <https://doi.org/10.1037/0021-843X.90.4.321>.
- Maric, V., Ramanathan, D., Mishra, J., 2020. Respiratory regulation & interactions with neuro-cognitive circuitry. *Neurosci. Biobehav. Rev.* <https://doi.org/10.1016/j.neubiorev.2020.02.001>.
- Marin, L., Issartel, J., Chaminade, T., 2009. Robots in the Wild: exploring human-robot interaction in naturalistic environments. *Interact. Stud. Soc. Behav. Commun. Biol. Artif. Syst.* 10, 479–504. <https://doi.org/10.1075/is.10.3.09mar>.
- Marosi, C., Fodor, Z., Csukly, G., 2019. From basic perception deficits to facial affect recognition impairments in schizophrenia. *Sci. Rep.* 9. <https://doi.org/10.1038/s41598-019-45231-x>.
- Marsh, K.L., Richardson, M.J., Schmidt, R.C., 2009. Social connection through joint action and interpersonal coordination. *Top. Cogn. Sci.* 1, 320–339. <https://doi.org/10.1111/j.1756-8765.2009.01022.x>.
- Marton-Alper, I.Z., Gvirts-Provolovski, H.Z., Nevat, M., Karklinsky, M., Shamay-Tsoory, S.G., 2020. Herding in human groups is related to high autistic traits. *Sci. Rep.* 10. <https://doi.org/10.1038/s41598-020-74951-8>.
- Mayo, O., Gordon, I., 2020. In and out of synchrony—Behavioral and physiological dynamics of dyadic interpersonal coordination. *Psychophysiology* 57, e13574. <https://doi.org/10.1111/psyp.13574>.
- Mazza, M., Lucci, G., Pacitti, F., Pino, M.C., Mariano, M., Casacchia, M., Roncone, R., 2010. Could schizophrenic subjects improve their social cognition abilities only with observation and imitation of social situations? *Neuropsychol. Rehabil.* 20, 675–703. <https://doi.org/10.1080/09602011.2010.486284>.
- McEwen, B.S., 2000. Allostasis and allostatic load: implications for neuropsychopharmacology. *Neuropsychopharmacology* 22, 108–124. [https://doi.org/10.1016/S0893-133X\(99\)00129-3](https://doi.org/10.1016/S0893-133X(99)00129-3).
- McNeill, W.H., 1995. *Keeping Together in Time: Dance and Drill in Human History*. Harvard University Press.
- Meeren, H.K.M., Van Heijnsbergen, C.C.R.J., De Gelder, B., 2005. Rapid perceptual integration of facial expression and emotional body language. *Proc. Natl. Acad. Sci. U. S. A.* 102, 16518–16523. <https://doi.org/10.1073/pnas.0507650102>.
- Meltzoff, A.N., Decety, J., 2003. What imitation tells us about social cognition: a rapprochement between developmental psychology and cognitive neuroscience. *Philos. Trans. R. Soc. B Biol. Sci.* 358, 491–500. <https://doi.org/10.1098/rstb.2002.1261>.
- Melzer, A., Shafir, T., Tsachor, R.P., 2019. How do we recognize emotion from movement? Specific motor components contribute to the recognition of each emotion. *Front. Psychol.* 10, 1389. <https://doi.org/10.3389/fpsyg.2019.01389>.
- Metta, G., Sandini, G., Vernon, D., Natale, L., Nori, F., 2008. The iCub humanoid robot. In: Madhavan, R., Messina, E. (Eds.), *Proceedings of the 8th Workshop on Performance Metrics for Intelligent Systems - PerMIS' 08*. ACM Press, New York, New York, USA, p. 50. <https://doi.org/10.1145/1774674.1774683>.

- Miles, L.K., Griffiths, J.L., Richardson, M.J., Macrae, C.N., 2009. Too late to coordinate: Contextual influences on behavioral synchrony. *Eur. J. Soc. Psychol.* 40 <https://doi.org/10.1002/ejsp.721> n/a/n/a.
- Mønster, D., Håkansson, D.D., Eskildsen, J.K., Wallot, S., 2016. Physiological evidence of interpersonal dynamics in a cooperative production task. *Physiol. Behav.* 156, 24–34. <https://doi.org/10.1016/j.physbeh.2016.01.004>.
- Mogan, R., Fischer, R., Bulbulia, J.A., 2017. To be in synchrony or not? A meta-analysis of synchrony's effects on behavior, perception, cognition and affect. *J. Exp. Soc. Psychol.* 72, 13–20. <https://doi.org/10.1016/j.jesp.2017.03.009>.
- Monier, F., Droit-Volet, S., 2018. Synchrony and emotion in children and adults. *Int. J. Psychol.* 53, 184–193. <https://doi.org/10.1002/ijop.12363>.
- Moody, E.J., McIntosh, D.N., Mann, L.J., Weisser, K.R., 2007. More than mere mimicry? The influence of emotion on rapid facial reactions to faces. *Emotion* 7, 447–457. <https://doi.org/10.1037/1528-3542.7.2.447>.
- Moriguchi, Y., Ohnishi, T., Decety, J., Hirakata, M., Maeda, M., Matsuda, H., Komaki, G., 2009. The human mirror neuron system in a population with deficient self-awareness: an fMRI study in alexithymia. *Hum. Brain Mapp.* 30, 2063–2076. <https://doi.org/10.1002/hbm.20653>.
- Mou, W., Celiktutan, O., Gunes, H., 2015. Group-level arousal and valence recognition in static images: face, body and context. 2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, FG 2015. Institute of Electrical and Electronics Engineers Inc. 1–6. <https://doi.org/10.1109/FG.2015.7284862>.
- Mou, W., Gunes, H., Patras, I., 2016. Automatic recognition of emotions and membership in group videos. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. IEEE Computer Society, pp. 1478–1486. <https://doi.org/10.1109/CVPRW.2016.185>.
- Moyle, W., Cooke, M., Beattie, E., Jones, C., Klein, B., Cook, G., Gray, C., 2013. Exploring the effect of companion robots on emotional expression in older adults with dementia: a pilot randomized controlled trial. *J. Gerontol. Nurs.* 39, 46–53. <https://doi.org/10.3928/00989134-20130313-03>.
- Muthukumarana, S., Elvitigala, D.S., Forero Cortes, J.P., Matthies, D.J.C., Nanayakkara, S., 2020. Touch me gently: recreating the perception of touch using a shape-memory alloy matrix. In: Conference on Human Factors in Computing Systems - Proceedings. Association for Computing Machinery, New York, NY, USA, pp. 1–12. <https://doi.org/10.1145/3313831.3376491>.
- Nackaerts, E., Wagemans, J., Helsen, V., Swinnen, S.P., Wenderoth, N., Alaerts, K., 2012. Recognizing biological motion and emotions from point-light displays in autism Spectrum disorders. *PLoS One* 7, 44473. <https://doi.org/10.1371/journal.pone.0044473>.
- Nadel, J., 2015. Perception-action coupling and imitation in autism spectrum disorder. *Dev. Med. Child Neurol.* 57, 55–58. <https://doi.org/10.1111/dmcn.12689>.
- Nadler, R., 2020. Understanding “zoom fatigue”: theorizing spatial dynamics as third skins in computer-mediated communication. *Comput. Compos.* 58, 102613 <https://doi.org/10.1016/j.compcom.2020.102613>.
- Niedenthal, P.M., 2007. Embodying emotion. *Science* 316, 1002–1005. <https://doi.org/10.1126/science.1136930>.
- Nijhof, S.L., Vinkers, C.H., van Geelen, S.M., Duijff, S.N., Achterberg, E.J.M., van der Net, J., Veltkamp, R.C., Grootenhuis, M.A., van de Putte, E.M., Hillegers, M.H.J., van der Brug, A.W., Wierenga, C.J., Benders, M.J.N.L., Engels, R.C.M.E., van der Ent, C. K., Vanderschuren, L.J.M.J., Lesscher, H.M.B., 2018. Healthy play, better coping: the importance of play in the development of children in health and disease. *Neurosci. Biobehav. Rev.* <https://doi.org/10.1016/j.neubiorev.2018.09.024>.
- Noy, L., Dekel, E., Alon, U., 2011. The mirror game as a paradigm for studying the dynamics of two people improvising motion together. *Proc. Natl. Acad. Sci. U. S. A.* 108, 20947–20952. <https://doi.org/10.1073/pnas.1108155108>.
- Numata, T., Sato, H., Asa, Y., Koike, T., Miyata, K., Nakagawa, E., Sumiya, M., Sadato, N., 2020. Achieving affective human–virtual agent communication by enabling virtual agents to imitate positive expressions. *Sci. Rep.* 10 <https://doi.org/10.1038/s41598-020-62870-7>.
- Nummenmaa, L., Gleason, E., Viinikainen, M., Jääskeläinen, I.P., Hari, R., Sams, M., 2012. Emotions promote social interaction by synchronizing brain activity across individuals. *Proc. Natl. Acad. Sci. U. S. A.* 109, 9599–9604. <https://doi.org/10.1073/pnas.1206095109>.
- Olugbade, T.A., Singh, A., Bianchi-Berthouze, N., Marquardt, N., Aung, M.S.H., De Williams, A.C.C., 2019. How can affect be detected and represented in technological support for physical rehabilitation? *ACM Trans. Comput. Interact.* 26, 1–29. <https://doi.org/10.1145/3299095>.
- Ott, E., Antonsen, T.M., 2017. Frequency and phase synchronization in large groups: low dimensional description of synchronized clapping, firefly flashing, and cricket chirping. *Chaos* 27, 051101. <https://doi.org/10.1063/1.4983470>.
- Padoa-Schioppa, C., 2011. Neurobiology of economic choice: a good-based model. *Annu. Rev. Neurosci.* 34, 333–359. <https://doi.org/10.1146/annurev-neuro-061010-113648>.
- Palumbo, R.V., Marraccini, M.E., Weyandt, L.L., Wilder-Smith, O., McGee, H.A., Liu, S., Goodwin, M.S., 2017. Interpersonal autonomic physiology: a systematic review of the literature. *Pers. Soc. Psychol. Rev.* 21, 99–141. <https://doi.org/10.1177/1088868316628405>.
- Panicker, S.S., Gayathri, P., 2019. A survey of machine learning techniques in physiology based mental stress detection systems. *Biocybern. Biomed. Eng.* <https://doi.org/10.1016/j.bbe.2019.01.004>.
- Parkinson, B., 2020. Intragroup emotion convergence: beyond contagion and social appraisal. *Pers. Soc. Psychol. Rev.* 24, 121–140. <https://doi.org/10.1177/1088868319882596>.
- Patterson, M.L., 1983. *Nonverbal Behavior*. Springer, New York. <https://doi.org/10.1007/978-1-4612-5564-2>.
- Paxton, A., Dale, R., 2013. Argument disrupts interpersonal synchrony. *Q. J. Exp. Psychol. (Hove)* 66, 2092–2102. <https://doi.org/10.1080/17470218.2013.853089>.
- Pezzulo, G., Donnarumma, F., Dindo, H., 2013. Human sensorimotor communication: a theory of signaling in online social interactions. *PLoS One* 8, e79876. <https://doi.org/10.1371/journal.pone.0079876>.
- Phillips-Silver, J., Keller, P.E., 2012. Searching for roots of entrainment and joint action in early musical interactions. *Front. Hum. Neurosci.* 6, 26. <https://doi.org/10.3389/fnhum.2012.00026>.
- Pickering, W.S.F., 2020. The Elementary Forms of the Religious Life. Durkheim on Religion, pp. 102–166. <https://doi.org/10.2307/j.ctt1cgf6tc.14>.
- Pikovsky, A., Rosenblum, M., Kurths, J., Hilborn, R.C., 2002. Synchronization: a universal concept in nonlinear science. *Am. J. Phys.* 70 <https://doi.org/10.1119/1.1475332>, 655–655.
- Pita, S., Pedro, L., 2011. Verbal and Non-Verbal Communication in Second Life, pp. 100–116. <https://doi.org/10.4018/978-1-60960-854-5.ch007>.
- Pluess, M., Conrad, E., Wilhelm, F.H., 2009. Muscle tension in generalized anxiety disorder: a critical review of the literature. *J. Anxiety Disord.* 23, 1–11. <https://doi.org/10.1016/j.janxdis.2008.03.016>.
- Plutchik, R., 2001. The Nature of Emotions. *Am. Sci.* 89 (4), 344. <https://doi.org/10.1511/2001.4.344>.
- Pop, C.A., Simut, R., Pinte, S., Saldien, J., Rusu, A., David, D., Vanderfaellie, J., Lefebvre, D., Vanderborght, B., 2013. Can the social robot Probo help children with autism to identify situation-based emotions? A series of single case experiments. *Int. J. Humanoid Robot.* 10, 1350025 <https://doi.org/10.1142/S0219843613500254>.
- Poria, S., Cambria, E., Bajpai, R., Hussain, A., 2017. A review of affective computing: from unimodal analysis to multimodal fusion. *Inf. Fusion* 37, 98–125. <https://doi.org/10.1016/j.inffus.2017.02.003>.
- Pouw, W., Prokisch, S., Drijvers, L., Gamba, M., Holler, J., Kello, C., Schaefer, R., Wiggins, G., 2021. Multilevel Rhythms in Multimodal Communication (in Press.. <https://doi.org/10.31219/OSF.IO/PSMHN>.
- Pratt, M., Goldstein, A., Levy, J., Feldman, R., 2017. Maternal depression across the first years of life impacts the neural basis of empathy in preadolescence. *J. Am. Acad. Child Adolesc. Psychiatry* 56, 20–29. <https://doi.org/10.1016/j.jaac.2016.10.012> e3.
- Preston, S.D., de Waal, F.B.M., 2002. Empathy: its ultimate and proximate bases. *Behav. Brain Sci.* <https://doi.org/10.1017/S0140525X02000018>.
- Prochazkova, E., Kret, M.E., 2017. Connecting minds and sharing emotions through mimicry: a neurocognitive model of emotional contagion. *Neurosci. Biobehav. Rev.* <https://doi.org/10.1016/j.neubiorev.2017.05.013>.
- Raffard, S., Salesses, R.N., Bortolon, C., Bardy, B.G., Henriques, J., Marin, L., Stricker, D., Capdevielle, D., 2018. Using mimicry of body movements by a virtual agent to increase synchronization behavior and rapport in individuals with schizophrenia. *Sci. Rep.* 8 <https://doi.org/10.1038/s41598-018-35813-6>.
- Ramachandran, V.S., 2000. Mirror Neurons and Imitation Learning As the Driving Force Behind “the Great Leap Forward” in Human Evolution. *Edge.org*, pp. 1–7.
- Ramseyer, F., Tschacher, W., 2014. Nonverbal synchrony of head- and body-movement in psychotherapy: different signals have different associations with outcome. *Front. Psychol.* 5, 979. <https://doi.org/10.3389/fpsyg.2014.00979>.
- Rauchbauer, B., Grosbras, M.H., 2020. Developmental trajectory of interpersonal motor alignment: positive social effects and link to social cognition. *Neurosci. Biobehav. Rev.* <https://doi.org/10.1016/j.neubiorev.2020.07.032>.
- Rauchbauer, B., Majdandžić, J., Hummer, A., Windischberger, C., Lamm, C., 2015. Distinct neural processes are engaged in the modulation of mimicry by social group-membership and emotional expressions. *Cortex* 70, 49–67. <https://doi.org/10.1016/j.cortex.2015.03.007>.
- Reisenzein, R., 2015. A short history of psychological perspectives on emotion. In: Calvo, R.A., D’Mello, S., Gratch, J., Kappas, A. (Eds.), *The Oxford Handbook of Affective Computing*. Oxford University Press, Oxford and New York. <https://doi.org/10.1093/oxfordhb/9780199942237.013.014>.
- Ridderinkhof, K.R., 2017. Emotion in action: a predictive processing perspective and theoretical synthesis. *Emot. Rev.* 9, 319–325. <https://doi.org/10.1177/1754073916661765>.
- Rio, K.W., Dachner, G.C., Warren, W.H., 2018. Local interactions underlying collective motion in human crowds. *Proc. R. Soc. B Biol. Sci.* 285, 20180611. <https://doi.org/10.1098/rspb.2018.0611>.
- Rivas, J.J., Orihuela-Espina, F., Sucar, L.E., Bianchi-Berthouze, N., 2021. Dealing with a missing sensor in a multilabel and multimodal automatic affective states recognition system. *Proc. Affect. Comput. Intelligen Interact.* 2021 in press.
- Rizzolatti, G., Craighero, L., 2004. The MIRROR-NEURON System. *Annu. Rev. Neurosci.* 27, 169–192. <https://doi.org/10.1146/annurev.neuro.27.070203.144230>.
- Roloff, M.E., Janiszewski, C.A., McGrath, M.A., Burns, C.S., Manrai, L.A., 1988. Acquiring resources from intimates when obligation substitutes for persuasion. *Hum. Commun. Res.* 14, 364–396. <https://doi.org/10.1111/j.1468-2958.1988.tb00161.x>.
- Ros, R., Baroni, I., Demiris, Y., 2014. Adaptive human-robot interaction in sensorimotor task instruction: From human to robot dance tutors. *Rob. Auton. Syst.* 62, 707–720. <https://doi.org/10.1016/j.robot.2014.03.005>.
- Rouaix, N., Retru-Chavastel, L., Rigaud, A.-S., Monnet, C., Lenoir, H., Pino, M., 2017. Affective and engagement issues in the conception and assessment of a robot-assisted psychomotor therapy for persons with dementia. *Front. Psychol.* 8, 950. <https://doi.org/10.3389/fpsyg.2017.00950>.
- Russell, J., 1980. A circumplex model of affect. *J. Personality Social Psychol.* 39 (6), 1161–1178. <https://doi.org/10.1037/h0077714>.
- Rychlowska, M., Canadas, E., Wood, A., Krumhuber, E.G., Fischer, A., Niedenthal, P.M., 2014. Blocking mimicry makes true and false smiles look the same. *PLoS One* 9, e90876. <https://doi.org/10.1371/journal.pone.0090876>.

- Salmela, M., Nagatsu, M., 2017. How does it really feel to act together? Shared emotions and the phenomenology of we-agency. *Phenomenol. Cogn. Sci.* 16, 449–470. <https://doi.org/10.1007/s11097-016-9465-z>.
- Sansonetti, R., Pierpaoli, C., Ferrante, L., Fabri, M., NARDI, B., 2020. Imitation strategies in subjects with schizophrenia: a behavioural approach. *Arch. Ital. Biol.* 158, 3–16. <https://doi.org/10.12871/00039829202011>.
- Scarpa, A., Ashley, R.A., Waldron, J.C., Zhou, Y., Swain, D.M., Dunsmore, J.C., Bell, M. A., 2018. Side by side: modeling dyadic physiological linkage in strangers. *Emotion* 18, 615–624. <https://doi.org/10.1037/emo0000340>.
- Scharoun, S.M., Bryden, P.J., 2016. Anticipatory planning in children with autism spectrum disorder: an assessment of independent and joint action tasks. *Front. Integr. Neurosci.* 10 <https://doi.org/10.3389/fnint.2016.00029>.
- Scherer, K.R., 2005. What are emotions? And how can they be measured? *Soc. Sci. Info.* 44, 695–729. <https://doi.org/10.1177/0539018405058216>.
- Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., Voegele, K., 2013. Toward a second-person neuroscience. *Behav. Brain Sci.* 36, 393–414. <https://doi.org/10.1017/S0140525X12000660>.
- Schilbach, L., Derntl, B., Aleman, A., Caspers, S., Cios, M., Diederer, K.M.J., Gruber, O., Kogler, L., Liemburg, E.J., Sommer, I.E., Möller, V.I., Cieslik, E.C., Eickhoff, S.B., 2016. Differential patterns of dysconnectivity in mirror neuron and mentalizing networks in schizophrenia. *Schizophr. Bull.* 42 <https://doi.org/10.1093/schbul/sbw015>.
- Schulte-Rüther, M., Otte, E., Adigüzel, K., Firk, C., Herpertz-Dahlmann, B., Koch, I., Konrad, K., 2017. Intact mirror mechanisms for automatic facial emotions in children and adolescents with autism spectrum disorder. *Autism Res.* 10, 298–310. <https://doi.org/10.1002/aur.1654>.
- Sebanz, N., Bekkering, H., Knoblich, G., 2006. Joint action: Bodies and minds moving together. *Trends Cogn. Sci.* <https://doi.org/10.1016/j.tics.2005.12.009>.
- Sebo, S., Stoll, B., Scassellati, B., Jung, M.F., 2020. Robots in groups and teams. *Proc. ACM Human-Comput. Interact.* 4, 1–36. <https://doi.org/10.1145/3415247>.
- Sequeira, J.S., Ferreira, I.A., 2016. Lessons from the MONARCH project. ICINCO 2016 - Proceedings of the 13th International Conference on Informatics in Control, Automation and Robotics. SciTePress 241–248. <https://doi.org/10.5220/0005998102410248>.
- Shamay-Tsoory, S.G., Aharon-Peretz, J., Perry, D., 2009. Two systems for empathy: A double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain* 132, 617–627. <https://doi.org/10.1093/brain/awn279>.
- Shamay-Tsoory, S.G., Saporta, N., Marton-Alper, I.Z., Gvirts, H.Z., 2019. Herding brains: a core neural mechanism for social alignment. *Trends Cogn. Sci.* <https://doi.org/10.1016/j.tics.2019.01.002>.
- Shu, L., Xie, J., Yang, M., Li, Ziyi, Li, Zhenqi, Liao, D., Xu, X., Yang, X., 2018. A review of emotion recognition using physiological signals. *Sensors (Switzerland)*. <https://doi.org/10.3390/s18072074>.
- Siegel, E.H., Sands, M.K., Van den Noortgate, W., Condon, P., Chang, Y., Dy, J., Quigley, K.S., Barrett, L.F., 2018. Emotion fingerprints or emotion populations? A meta-analytic investigation of autonomic features of emotion categories. *Psychol. Bull.* 144, 343–393. <https://doi.org/10.1037/bul0000128>.
- Stowiński, P., Zhai, C., Alderisio, F., Saless, R., Guegnon, M., Marin, L., Bardy, B.G., Di Bernardo, M., Tsaneva-Atanasova, K., 2016. Dynamic similarity promotes interpersonal coordination in joint action. *J. R. Soc. Interface* 13. <https://doi.org/10.1098/rsif.2015.1093>.
- Stowiński, P., Alderisio, F., Zhai, C., Shen, Y., Tino, P., Bortolon, C., Capdevielle, D., Cohen, L., Khoramshahi, M., Billard, A., Saless, R., Guegnon, M., Marin, L., Bardy, B.G., Di Bernardo, M., Raffard, S., Tsaneva-Atanasova, K., 2017. Unravelling socio-motor biomarkers in schizophrenia. *NPJ Schizophr.* 3 <https://doi.org/10.1038/s41537-016-0009-x>.
- Smirnov, D., Saarimäki, H., Glerean, E., Hari, R., Sams, M., Nummenmaa, L., 2019. Emotions amplify speaker–listener neural alignment. *Hum. Brain Mapp.* 40, 4777–4788. <https://doi.org/10.1002/hbm.24736>.
- Soleymani, M., Lichtenauer, J., Pun, T., Pantic, M., 2012. A multimodal database for affect recognition and implicit tagging. *IEEE Trans. Affect. Comput.* 3, 42–55. <https://doi.org/10.1109/T-AFFC.2011.25>.
- Solomon, R.L., Corbit, J.D., 1974. An opponent-process theory of motivation: I. Temporal dynamics of affect. *Psychol. Rev.* <https://doi.org/10.1037/h0036128>.
- Soriano, M., Cavallo, A., D'Ausilio, A., Becchio, C., Fadiga, L., 2018. Movement kinematics drive chain selection toward intention detection. *Proc. Natl. Acad. Sci. U. S. A.* 115, 10452–10457. <https://doi.org/10.1073/pnas.1809825115>.
- Srinivasan, S.M., Kaur, M., Park, I.K., Gifford, T.D., Marsh, K.L., Bhat, A.N., 2015. The effects of rhythm and robotic interventions on the Imitation/Praxis, interpersonal synchrony, and motor performance of children with autism Spectrum disorder (ASD): a pilot randomized controlled trial. *Autism Res. Treat.* 2015, 1–18. <https://doi.org/10.1155/2015/736516>.
- Srinivasan, S.M., Eigsti, I.M., Gifford, T., Bhat, A.N., 2016. The effects of embodied rhythm and robotic interventions on the spontaneous and responsive verbal communication skills of children with Autism Spectrum Disorder (ASD): A further outcome of a pilot randomized controlled trial. *Res. Autism Spectr. Disord.* 27, 73–87. <https://doi.org/10.1016/j.rasd.2016.04.001>.
- Stellar, J.E., Gordon, A.M., Piff, P.K., Cordero, D., Anderson, C.L., Bai, Y., Maruskin, L.A., Keltner, D., 2017. Self-transcendent emotions and their social functions: compassion, gratitude, and awe bind us to others through prosociality. *Emot. Rev.* 9, 200–207. <https://doi.org/10.1177/1754073916684557>.
- Strogatz, S., 2004. *Syn: the Emerging Science of Spontaneous Order*.
- Su, W.C., Culotta, M., Mueller, J., Tsuzuki, D., Pelphrey, K., Bhat, A., 2020a. Differences in cortical activation patterns during action observation, action execution, and interpersonal synchrony between children with or without autism spectrum disorder (ASD): an fNIRS pilot study. *PLoS One* 15. <https://doi.org/10.1371/journal.pone.0240301>.
- Su, W.C., Culotta, M.L., Hoffman, M.D., Trost, S.L., Pelphrey, K.A., Tsuzuki, D., Bhat, A. N., 2020b. Developmental differences in cortical activation during action observation, action execution and interpersonal synchrony: an fNIRS study. *Front. Hum. Neurosci.* 14 <https://doi.org/10.3389/fnhum.2020.00057>.
- Subramanian, R., Wache, J., Abadi, M.K., Vierli, R.L., Winkler, S., Sebe, N., 2018. Ascertain: emotion and personality recognition using commercial sensors. *IEEE Trans. Affect. Comput.* 9, 147–160. <https://doi.org/10.1109/TAFFC.2016.2625250>.
- Takagi, A., Hirashima, M., Nozaki, D., Burdet, E., 2019. Individuals physically interacting in a group rapidly coordinate their movement by estimating the collective goal. *Elife* 8. <https://doi.org/10.7554/eLife.41328>.
- Terada, K., Takeuchi, C., 2017. Emotional expression in simple line drawings of a robot's face leads to higher offers in the ultimatum game. *Front. Psychol.* 8, 724. <https://doi.org/10.3389/fpsyg.2017.00724>.
- Terzioglu, Y., Mutlu, B., Sahin, E., 2020. Designing social cues for collaborative robots: the role of gaze and breathing in human-robot collaboration. In: *ACM/IEEE International Conference on Human-Robot Interaction*. IEEE Computer Society, New York, NY, USA, pp. 343–357. <https://doi.org/10.1145/3319502.3374829>.
- Thelen, E., Smith, L.B., 1994. *A Dynamic Systems Approach to the Development of Cognition and Action*.
- Thompson, E.L., Bird, G., Catmur, C., 2019. Conceptualizing and testing action understanding. *Neurosci. Biobehav. Rev.* 105, 106–114. <https://doi.org/10.1016/j.neubiorev.2019.08.002>.
- Ting, L.H., Chiel, H.J., Trumbower, R.D., Allen, J.L., McKay, J.L., Hackney, M.E., Kesar, T.M., 2015. Neuromechanical principles underlying movement modularity and their implications for rehabilitation. *Neuron*. <https://doi.org/10.1016/j.neuron.2015.02.042>.
- Todorov, E., Jordan, M.I., 2002. Optimal feedback control as a theory of motor coordination. *Nat. Neurosci.* 5, 1226–1235. <https://doi.org/10.1038/nn963>.
- Tognoli, E., Lagarde, J., DeGuzman, G.C., Kelso, J.A.S., 2007. The phi complex as a neuromarker of human social coordination. *Proc. Natl. Acad. Sci. U. S. A.* 104, 8190–8195. <https://doi.org/10.1073/pnas.0611453104>.
- Toichoa Eyam, A., 2019. *Emotion-driven Human-robot Interaction Based on EEG Industrial Applications*.
- Toma, C.L., 2014. Towards conceptual convergence: an examination of interpersonal adaptation. *Commun. Q.* 62, 155–178. <https://doi.org/10.1080/01463373.2014.890116>.
- Tomasello, M., 2011. *Human Culture in Evolutionary Perspective*. *Adv. Cult. Psychol. Vol. 1, Advances in Culture and Psychology*, pp. 5–51.
- Torres, E.P., Torres, E.A., Hernández-Álvarez, M., Yoo, S.G., 2020. EEG-based BCI emotion recognition: a survey. *Sensors* 20, 5083. <https://doi.org/10.3390/s20185083>.
- Trevisan, D.A., Enns, J.T., Birmingham, E., Iarocci, G., 2021. Action coordination during a real-world task: evidence from children with and without autism spectrum disorder. *Dev. Psychopathol.* 33, 65–75. <https://doi.org/10.1017/S0954579419001561>.
- Troje, N.F., 2002. Decomposing biological motion: a framework for analysis and synthesis of human gait patterns. *J. Vis.* 2, 371–387. <https://doi.org/10.1167/2.5.2>.
- Tsai, J.C.C., Sebanz, N., Knoblich, G., 2011. The GROOP effect: groups mimic group actions. *Cognition* 118, 135–140. <https://doi.org/10.1016/j.cognition.2010.10.007>.
- Tunçgenç, B., Pacheco, C., Rochowiak, R., Nicholas, R., Rengarajan, S., Zou, E., Messenger, B., Vidal, R., Mostofsky, S.H., 2021. Computerized assessment of motor imitation as a scalable method for distinguishing children with autism. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 6, 321–328. <https://doi.org/10.1016/j.bpsc.2020.09.001>.
- Tuomela, R., 2011. Cooperation as joint action. *Anal. Und Krit.* 2011, 65–86. <https://doi.org/10.1515/auk-2011-0106>.
- Ugander, J., Backstrom, L., Marlow, C., Kleinberg, J., 2012. Structural diversity in social contagion. *Proc. Natl. Acad. Sci. U. S. A.* 109, 5962–5966. <https://doi.org/10.1073/pnas.1116502109>.
- Ukita, N., Moriguchi, Y., Hagita, N., 2016. People re-identification across non-overlapping cameras using group features. *Comput. Vis. Image Underst.* 144, 228–236. <https://doi.org/10.1016/j.cviu.2015.06.011>.
- Valdesolo, P., Ouyang, J., DeSteno, D., 2010. The rhythm of joint action: synchrony promotes cooperative ability. *J. Exp. Soc. Psychol.* 46, 693–695. <https://doi.org/10.1016/j.jesp.2010.03.004>.
- Valentí Soler, M., Agüera-Ortiz, L., Olazarán Rodríguez, J., Mendoza Rebolledo, C., Pérez Muñoz, A., Rodríguez Pérez, I., Osa Ruiz, E., Barrios Sánchez, A., Herrero Cano, V., Carrasco Chillón, L., Felipe Ruiz, S., López Álvarez, J., León Salas, B., Cañas Plaza, J. M., Martín Rico, F., Martínez Martín, P., 2015. Social robots in advanced dementia. *Front. Aging Neurosci.* 7, 133. <https://doi.org/10.3389/fnagi.2015.00133>.
- van de Perre, G., van Damme, M., Lefeber, D., Vanderborcht, B., 2015. Development of a generic method to generate upper-body emotional expressions for different social robots. *Adv. Robot.* 29, 597–609. <https://doi.org/10.1080/01691864.2015.1031697>.
- Van den Stock, J., Righart, R., de Gelder, B., 2007. Body expressions influence recognition of emotions in the face and voice. *Emotion* 7, 487–494. <https://doi.org/10.1037/1528-3542.7.3.487>.
- Van der Donck, S., Dzhelyova, M., Vettori, S., Mahdi, S.S., Claes, P., Steyaert, J., Boets, B., 2020. Rapid neural categorization of angry and fearful faces is specifically impaired in boys with autism spectrum disorder. *J. Child Psychol. Psychiatry Allied Discip.* 61, 1019–1029. <https://doi.org/10.1111/jcpp.13201>.
- Van Der Veer, R., 1996. Henri Wallon's theory of early child development: the role of emotions. *Dev. Rev.* 16, 364–390. <https://doi.org/10.1006/drev.1996.0016>.

- Varlet, M., Marin, L., Raffard, S., Schmidt, R.C., Capdevielle, D., Boulenger, J.P., Del-Monte, J., Bardy, B.G., 2012. Impairments of social motor coordination in schizophrenia. *PLoS One* 7. <https://doi.org/10.1371/journal.pone.0029772>.
- Varni, G., Volpe, G., Camurri, A., 2010. A system for real-time multimodal analysis of nonverbal affective social interaction in user-centric media. *IEEE Trans. Multimed.* 12, 576–590. <https://doi.org/10.1109/TMM.2010.2052592>.
- Varni, G., Mancini, M., Fadiga, L., Camurri, A., Volpe, G., 2019. The change matters! Measuring the effect of changing the leader in joint music performances. *IEEE Trans. Affect. Comput.* <https://doi.org/10.1109/taffc.2019.2951368>, 1–1.
- Veltmeijer, E.A., Gerritsen, C., Hindriks, K., 2021. Automatic emotion recognition for groups: a review. *IEEE Trans. Affect. Comput.* <https://doi.org/10.1109/TAFFC.2021.3065726>.
- Verduyn, P., Delaveau, P., Rotgé, J.-Y., Fossati, P., Van Mechelen, I., 2015. Determinants of emotion duration and underlying psychological and neural mechanisms. *Emot. Rev.* 7, 330–335. <https://doi.org/10.1177/1754073915590618>.
- Vesper, C., Van Der Wel, R.P.R.D., Knoblich, G., Sebanz, N., 2011. Making oneself predictable: reduced temporal variability facilitates joint action coordination. *Exp. Brain Res.* 211, 517–530. <https://doi.org/10.1007/s00221-011-2706-z>.
- Viana, K.M.P., Zambrana, I.M., Karevold, E.B., Pons, F., 2020. Emotions in motion: impact of emotion understanding on children's peer action coordination. *Cogn. Emot.* 34, 831–838. <https://doi.org/10.1080/02699931.2019.1669535>.
- Vicaria, I.M., Dickens, L., 2016. Meta-analyses of the intra- and interpersonal outcomes of interpersonal coordination. *J. Nonverbal Behav.* 40, 335–361. <https://doi.org/10.1007/s10919-016-0238-8>.
- Vicary, S., Sperling, M., Von Zimmermann, J., Richardson, D.C., Orgs, G., 2017. Joint action aesthetics. *PLoS One* 12. <https://doi.org/10.1371/journal.pone.0180101>.
- von Zimmermann, J., Richardson, D.C., 2016. Verbal synchrony and action dynamics in large groups. *Front. Psychol.* 7 <https://doi.org/10.3389/fpsyg.2016.02034>.
- Wagner, J., Lingensfelder, F., André, E., Kim, J., 2011. Exploring fusion methods for multimodal emotion recognition with missing data. *IEEE Trans. Affect. Comput.* 2, 206–218. <https://doi.org/10.1109/T-AFFC.2011.12>.
- Wallbott, H.G., 1998. Bodily expression of emotion. *Eur. J. Soc. Psychol.* 28, 879–896 [https://doi.org/10.1002/\(SICI\)1099-0992\(199811\)28:6<textless>879::AID-EJSP901<textgreater>3.0.CO;2-W](https://doi.org/10.1002/(SICI)1099-0992(199811)28:6<textless>879::AID-EJSP901<textgreater>3.0.CO;2-W).
- Wallot, S., Mitkidis, P., McGraw, J.J., Roepstorff, A., 2016a. Beyond synchrony: joint action in a complex production task reveals beneficial effects of decreased interpersonal synchrony. *PLoS One* 11, e0168306. <https://doi.org/10.1371/journal.pone.0168306>.
- Wallot, S., Mitkidis, P., McGraw, J.J., Roepstorff, A., 2016b. Beyond synchrony: joint action in a complex production task reveals beneficial effects of decreased interpersonal synchrony. *PLoS One* 11, e0168306. <https://doi.org/10.1371/journal.pone.0168306>.
- Wang, S., He, M., Gao, Z., He, S., Ji, Q., 2014. Emotion recognition from thermal infrared images using deep Boltzmann machine. *Front. Comput. Sci. China* 8, 609–618. <https://doi.org/10.1007/s11704-014-3295-3>.
- Wang, Q., Han, Z., Hu, X., Feng, S., Wang, H., Liu, T., Yi, L., 2020. Autism symptoms modulate interpersonal neural synchronization in children with autism Spectrum disorder in cooperative interactions. *Brain Topogr.* 33, 112–122. <https://doi.org/10.1007/s10548-019-00731-x>.
- Wang, C., Gao, Y., Mathur, A., De, C., Williams, A.C., Lane, N.D., Bianchi-Berthouze, N., 2021. Leveraging activity recognition to enable protective behavior detection in continuous data. *Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol.* 5, 1–27. <https://doi.org/10.1145/3463508>.
- Warren, W.H., 2006. The dynamics of perception and action. *Psychol. Rev.* 113, 358–389. <https://doi.org/10.1037/0033-295X.113.2.358>.
- Warreyn, P., Roeyers, H., 2014. See what I see, do as I do: Promoting joint attention and imitation in preschoolers with autism spectrum disorder. *Autism* 18, 658–671. <https://doi.org/10.1177/1362361313493834>.
- Wass, S.V., Smith, C.G., Clackson, K., Gibb, C., Eitzenberger, J., Mirza, F.U., 2019. Parents mimic and influence their infant's autonomic state through dynamic affective state matching. *Curr. Biol.* 29, 2415–2422. <https://doi.org/10.1016/j.cub.2019.06.016> e4.
- Williams, J.H.G., Massaro, D.W., Peel, N.J., Bosseler, A., Suddendorf, T., 2004. Visual-auditory integration during speech imitation in autism. *Res. Dev. Disabil.* 25, 559–575. <https://doi.org/10.1016/j.ridd.2004.01.008>.
- Williams, K.E., Savage, S., Eager, R., 2020. Rhythm and Movement for Self-Regulation (RAMSR) intervention for preschool self-regulation development in disadvantaged communities: a clustered randomised controlled trial study protocol. *BMJ Open* 10, e036392. <https://doi.org/10.1136/bmjopen-2019-036392>.
- Wilquin, H., Delevoye-Turrell, Y., Dione, M., Giersch, A., 2018. Motor synchronization in patients with schizophrenia: preserved time representation with abnormalities in predictive timing. *Front. Hum. Neurosci.* 12 <https://doi.org/10.3389/fnhum.2018.00193>.
- Wiltermuth, S.S., Heath, C., 2009. Synchrony and cooperation. *Psychol. Sci.* 20, 1–5. <https://doi.org/10.1111/j.1467-9280.2008.02253.x>.
- Witkower, Z., Tracy, J.L., 2019. Bodily Communication of Emotion: Evidence for Extrafacial Behavioral Expressions and Available Coding Systems. *Emot. Rev.* 11, 184–193. <https://doi.org/10.1177/1754073917749880>.
- Wolpert, D.M., Doya, K., Kawato, M., 2003. A unifying computational framework for motor control and social interaction. *Philos. Trans. R. Soc. B Biol. Sci.* <https://doi.org/10.1098/rstb.2002.1238>.
- Wood, A., Rychlowska, M., Korb, S., Niedenthal, P., 2016. Fashioning the face: sensorimotor simulation contributes to facial expression recognition. *Trends Cogn. Sci. (Regul. Ed.)*. <https://doi.org/10.1016/j.tics.2015.12.010>.
- Xavier, J., Magnat, J., Sherman, A., Gauthier, S., Cohen, D., Chaby, L., 2016. A developmental and clinical perspective of rhythmic interpersonal coordination: from mimicry toward the interconnection of minds. *J. Physiol. Paris*. <https://doi.org/10.1016/j.jphysparis.2017.06.001>.
- Yamashita, Y., Tani, J., 2008. Emergence of functional hierarchy in a multiple timescale neural network model: a humanoid robot experiment. *PLoS Comput. Biol.* 4, e1000220. <https://doi.org/10.1371/journal.pcbi.1000220>.
- Yang, F., Yin, W., Inamura, T., Björkman, M., Peters, C., 2020. Group behavior recognition using attention- and graph-based neural networks. *Frontiers in Artificial Intelligence and Applications*. IOS Press BV, pp. 1626–1633. <https://doi.org/10.3233/FAIA200273>.
- Yoshida, 2002. The application of entrainment to musical ensembles. II International Conference on Music and Artificial Intelligence (ICMAI), Edinburgh, Scotland.
- Yücel, Z., Zanlungo, F., Ikeda, T., Miyashita, T., Hagita, N., 2013. Deciphering the crowd: modeling and identification of pedestrian group motion. *Sensors (Switzerland)* 13, 875–897. <https://doi.org/10.3390/s130100875>.
- Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., Chapman, M., 1992. Development of concern for others. *Dev. Psychol.* 28, 126–136. <https://doi.org/10.1037/0012-1649.28.1.126>.
- Zampella, C.J., Csusmitta, K.D., Simon, E., Bennetto, L., 2020. Interactional synchrony and its association with social and communication ability in children with and without autism Spectrum disorder. *J. Autism Dev. Disord.* 50, 3195–3206. <https://doi.org/10.1007/s10803-020-04412-8>.
- Zhang, M., Dumas, G., Kelso, J.A.S., Tognoli, E., 2016. Enhanced emotional responses during social coordination with a virtual partner. *Int. J. Psychophysiol.* 104, 33–43. <https://doi.org/10.1016/j.ijpsycho.2016.04.001>.
- Zhang, M., Beetle, C., Kelso, J.A.S., Tognoli, E., 2019. Connecting empirical phenomena and theoretical models of biological coordination across scales. *J. R. Soc. Interface* 16, 20190360. <https://doi.org/10.1098/rsif.2019.0360>.
- Zhao, Z., Salesse, R.N., Gueugnon, M., Schmidt, R.C., Marin, L., Bardy, B.G., 2015. Moving attractive virtual agent improves interpersonal coordination stability. *Hum. Mov. Sci.* 41, 240–254. <https://doi.org/10.1016/j.humov.2015.03.012>.
- Zhao, S., Gholamnejad, A., Ding, G., Gao, Y., Han, J., Keutzer, K., 2019. Personalized emotion recognition by personality-aware high-order learning of physiological signals. *ACM Trans. Multimed. Comput. Commun. Appl.* 15, 1–18. <https://doi.org/10.1145/3233184>.
- Zhao, Z., Salesse, R.N., Qu, X., Marin, L., Gueugnon, M., Bardy, B.G., 2020. Influence of perceived emotion and gender on social motor coordination. *Br. J. Psychol.* 111, 536–555. <https://doi.org/10.1111/bjop.12419>.
- Zheng, W.L., Liu, W., Lu, Y., Lu, B.L., Cichocki, A., 2019. EmotionMeter: a multimodal framework for recognizing human emotions. *IEEE Trans. Cybern.* 49, 1110–1122. <https://doi.org/10.1109/TCYB.2018.2797176>.
- Zheng, W., Yu, A., Fang, P., Peng, K., 2020. Exploring collective emotion transmission in face-to-face interactions. *PLoS One* 15, e0236953. <https://doi.org/10.1371/journal.pone.0236953>.