



ELSEVIER

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Developmental Review

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

Time and sequence as key developmental dimensions of joint action

Valentina Fantasia^{a,*}, Jonathan Delafield-Butt^b

^a Department of Philosophy, Faculties of Humanities and Theology, Box 192, 221 00, Lund University, Lund, Sweden

^b Laboratory for innovation in autism, Faculty of Humanities and Social Sciences, University of Strathclyde, Glasgow, Scotland, UK

ARTICLE INFO

Keywords:

Joint actions
 Infancy
 Embodied social cognition
 Time
 Sequentiality
 Prospectivity

ABSTRACT

Joint action, generally defined as working together towards a common purpose, has become an important concept in many areas of cognitive science, from philosophical appraisal of its core concepts to empirical mapping of its psychological development. Within mainstream cognitive accounts, to engage in a joint action requires an inferential process of representing the other's intentions and plans to enable social coordination for a shared goal. However, growing endorsement of a contrasting view from embodied and situated accounts of social cognition proposes that joint action is better understood as a dynamic, situated interactional process where participants "roll into" joint action without requiring reflective or representational awareness of it. This work proposes a rethinking of how we conceive the nature of action and its development as joint action early in human life. With particular reference to developmental studies, we advance a rationale for the conceptual framework of joint action to include its temporal and sequential structures, and their intrinsic prospective qualities of human action, solitary or shared, as key analytical aspects for the study of how infants understand and share meaning with another, in joint *interaction*.

Introduction

Human activity is organised to take others into account.

Fasulo, A. & Pontecorvo, C. (1999)

Joint action, defined as working together towards a common purpose, has become an important concept in many areas of cognitive science, from its philosophical implications to its role in learning and social cognition in developmental and comparative psychology. At the root of joint action is its structure as both a psychological and physical phenomenon. On the one hand it is structured and directed by the private knowledge, interests and intentions of the individual, and on the other it is publicly manifest in motor behaviour and available for sharing within a social context. How the nature of action is changed by social engagements is a fundamental question for understanding the development of social cognition, our exceptional human capacity to know other minds and to

* Corresponding author.

E-mail addresses: Valentina.fantasia@lucs.lu.se (V. Fantasia), Jonathan.delafield-butt@strath.ac.uk (J. Delafield-Butt).

<https://doi.org/10.1016/j.dr.2023.101091>

Received 4 October 2022; Received in revised form 22 June 2023;

Available online 29 July 2023

0273-2297/Crown Copyright © 2023 Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

share in their interest and intentions.

There are, however, very different notions of what jointly acting means and what forms of jointness are required for what kinds of actions¹. Within mainstream cognitive accounts, joint action has been often framed as an individual mental activity, which is then shared in communication with another. That is, as a two-step process where interactants rely first on their internal representational mechanisms to make sense of what the other partner intends to do, then to design their own actions to match those of their partner, for example, to direct action toward a common goal. However, this view has been criticised by embodied and situated researchers in philosophy, psychology, and neuroscience as inescapably individualistic, mentalistic, and static, failing to take into account the moment-by-moment dynamic unfolding of inter-acting between individuals (see, for instance, works by Bråten, 2003; Costall, 2013; De Jaegher & Di Paolo, 2007; De Jaegher, Di Paolo, & Gallagher, 2010; Fantasia, Fasulo, Costall, & López, 2014a; Gallagher, 2008a; Leudar & Costall, 2009; Reddy, 2008; Reddy & Morris, 2009; Trevarthen & Delafield-Butt, 2013; Trevarthen, 2016). In this paper, we address this subtle but significant difference of perspective to advance a redefinition of joint action useful for researchers whose work is framed within a continuum of perspectives from mainstream cognitive to embodied accounts. This redefinition is supported by attention to early human engagements in shared, co-created sense-making practices, with particular focus on social routines in infancy. Social routines are temporally organised activities situated in specific time and space contexts. They serve as unique contexts for the generation of social meaning through active participation of both adult and infant, presenting the earliest cases of meaning-making in human life. Importantly, these interactional routines are always embodied and as such, are necessarily structured by the temporal and sequential order of action. This organisational structure is important for understanding the psychological and physical nature of joint action.

Although there is significant overlap between cognitively-focused approaches and more embodied or situated views of social cognition – and intentionality in particular – attributes of how humans make sense and share intentions vary significantly according to different theories and positions. They usually include a joint commitment to a goal, mutual responsiveness, and support in pursuing that commitment, understanding, and willingness for role-reversal (Bratman, 1992; Tomasello & Moll, 2010). Heesen and colleagues (2017) recently proposed to define shared intentionality “as a transient state of collective being that participants in joint action strive to attain and maintain, or in the terminology of conversation analysis, as an interactional achievement” (Schegloff, 1995, p. 392). A mainstream cognitive view of joint action frames it as an inferential process with emphasis on its underlying cognitive processes and their constitutive mechanisms at both the individual and the collective level (Bratman, 1992; Tuomela, 2005; for a recent overview, see Sebanz & Knoblich, 2021). Research within this view has investigated the emergence of children’s capacity to coordinate and cooperate in joint actions as an index of developing mind-reading abilities (for a comprehensive discussion, see Tomasello, 2019).

By contrast, studies adopting ecological, dynamic and processual views of inter-actions have broadened the definition of joint action to a form of acting together, an interactional process where participants can “roll into” joint action without necessarily having reflective awareness, or an explicit intent towards it (Fantasia et al., 2014b). In this view, the meaning of one’s own and others’ behaviours are negotiated along the ‘interactional way’ through a participatory processes of sense- and meaning-making (De Jaegher & Di Paolo, 2007; De Jaegher, Perakyla & Stevanovic, 2016; Trevarthen, & Aitken, 2001; Trevarthen, & Delafield-Butt, 2015). Under this conceptual framework, some developmental researchers have looked at the daily ecologies of infants and young children’s lives to understand how individual, sociocultural, and environmental dimensions shape the developmental pathways towards becoming competent participants in social activities or in everyday human ecologies (e.g., Gratier & Apter-Danon, 2009; Reddy, 2008).

Our work pursues three aims. First, to revisit mainstream cognitively-oriented developmental accounts of joint action as individual endeavours based on children’s representational capacities (such as the ability to infer others’ intentions). Second, to counterbalance this standard view with an embodied, dynamic and ecological view of ‘second person’ perspectives in joint action. Third, to advance time and sequence as essential elements for investigating how infants understand and share meaning with others in joint action¹. This final step is achieved by attending to a fundamental aspect of action not previously considered in mainstream cognitive approaches and seldom explicitly addressed in embodied ecological accounts: that of its prospective quality and its necessary sequential organisation. These key elements of joint action require consideration in future research, and empirical measurement and analysis.

Cognitively-oriented accounts of joint action

Mainstream developmental research in joint action has focused on children’s ability to understand, share, and coordinate with others’ intentions towards the accomplishment of a shared task. This view is grounded in philosophical accounts concerned with concepts such as shared or collective agency and so-called ‘we intentionality’ (Gilbert, 2013; Tuomela, 2005). These accounts maintain that sharing intentions is a necessary pre-requisite for a type of joint action where interactants share a commitment to perform a certain task with a clear end-goal, by formulating predictions upon the other agent’s intention to achieve the same goal (Bratman, 1992; Gilbert, 2013; Pacherie, 2006; Tomasello & Carpenter, 2007, Tomasello, Carpenter, Call, Benhe & Moll, 2005; Schweikard & Schmid, 2013). Jointly acting is therefore based on the premise that each agent must be able to take an intentional stance on her own and her partner’s plan to perform one (or more) action(s), upon engaging in the shared action, so that “in addition to representing single actions, joint action partners also represent other aspects of each other’s tasks, such as the order in which actions are to be performed”

¹ For the purposes of the present work, the terms shared, joint or collective are used interchangeably in referring to actions and/or intentions. Well aware of the philosophical arguments that maintain conceptual differences between these terms, a debate defending differences within those terms is beyond the scope of this work. Please refer to Schweikard and Schmid (2013), for a recent historical overview of those terms (collective intentionality, collective emotions, joint actions).

(Sebanz & Knoblich, 2021, p. 141).

Representational mechanisms grounding joint action development

Developmental research has extensively committed towards pinning down the origins of intentional engagements in joint actions and their shaping factors. Parental scaffolding is one key element: the strategies and approaches used by parents to encourage children's cooperative participation in joint tasks differ according to the children's age: at 18-month-olds, parents emphasised concrete, action-oriented task-based strategies, whereas, at 24-month-olds, they used more abstract reasoning strategies (Waugh, Brownell, & Pollock, 2015).

The emergence of joint action capacities has endorsed an equally inferential view proposing that children's capacity to participate in shared action is grounded in the development of basic representational mechanisms, such as mindreading and perspective-taking, and is tightly connected with imitative and cooperative skills. Selected abilities have been tested through experimental paradigms at specific developmental time points as indexes of this emerging capacity. Typically adopted paradigms include tasks aimed at eliciting and testing how children engage in structured, pre-set joint activities, such as those involving simultaneous coordination of movements and role-exchange or adapted versions of the prisoner's dilemma (Brownell and Carriger, 1990; Warneken, Chen & Tomasello, 2006). The underlying scope shared by these paradigms is to track the development of the cognitive underpinnings that enable young children (1) to share attentional focus (Conte, Grazzani & Pepe, 2018) to coordinate co-actions towards a shared goal (Satta et al., 2017); (2) to construct a meaningful mental representation of the other's intention based on a reading of their behaviours (Brownell & Carriger, 1990; Brownell, Ramani & Zerwas, 2006; Butterfill, 2012); and (3) to become cognisant of, and to coordinate with another's point of view in a functional task (Butterfill, 2012; Milward, Kita & Apperly, 2014; Tomasello, Carpenter, Call, Behne, & Moll, 2005). Warneken and colleagues (Warneken & Tomasello 2007; Warneken, Chen & Tomasello, 2006), for instance, investigated infants' capacity to coordinate actions in a series of shared but individually performed joint tasks, such as retrieving an object from a vertically movable cylinder embedded in a platform. They presented evidence that fourteen-months-old infants reliably helped a partner unable to achieve a goal, but only engaged successfully in tasks demanding low co-ordination; eighteen-months-old infants, instead, showed better coordination capacities, bringing evidence for an early emerging mental capacity [to] "understand another person's unfulfilled goal and altruistically help him or her to achieve it" (2007, p. 290). This feature was already observed at fourteen months, but "coordinating joint actions toward a joint goal seem[ed] to require cognitive and behavioral skills that [were] only beginning to emerge at this age" (p. 291).

More specifically, the capacity to distinguish and (distinctively) represent one's own intentional behaviour from those of others has been regarded as the hallmark of acting towards a joint goal (and, eventually, sharing intentions to do so). In children, according to the cognitive model, this is believed to emerge around the second year of age, indexed by behaviours indicating a capacity "to represent, monitor and regulate both their own and the partner's behaviour relative to their relation to a single, common goal" (Brownell & Carriger, 1990, p.1165). Brownell and colleagues (Brownell & Carriger, 1990; Brownell, Ramani & Zerwas, 2006) designed a series of tasks to pinpoint at what age young children display "... the ability to cooperate with peers, becoming a true social partner". They observed that 24-month-old children (but not younger) successfully engaged in a shared problem-solving task where they had to simultaneously or sequentially pull one handle of a wooden box to activate a musical toy with another child. Authors interpreted the successful accomplishment of the joint task as evidence that (1) this type of task implied "being able to represent and to share goals and intentions with a partner to engage in shared cooperative activities" (p. 806); and (2) mastering of representational abilities (of the kind described) is key in children's emerging ability to successfully engage in joint actions. Using a conceptually similar paradigm (cooperation in a problem-solving task), Hamann, Warneken Tomasello (2012) proposed that a developmental shift in intention understanding, and attribution may be reflected in how children understand and respond to the normative dimensions of a collaborative activity between 2 and 3 years of age.

While there is little doubt that mainstream developmental research has provided us with a rich understanding of how and when, in developmental time, young children can engage in joint coordination of interdependencies with peers or adults, including movements, gaze patterns, or nonverbal signs, we maintain that their contribution is limited in several aspects. One above all is circularity: on developmental grounds, the proposal of high-level (representational) cognitive abilities as necessary prerequisites for engaging in shared actions is clearly problematic. As Butterfill (2012) argued, under these accounts, "engaging in joint action involves sharing intentions and sharing intentions requires much of the understanding of minds whose development is supposed to be explained by appeal to joint action. ... So, if the leading [*cognitive, our term*] account were the whole truth about joint action, engaging in joint action would presuppose, and therefore could not explain, much of the development of reasoning about others' mental states" (p. 28). Infants' and children's "performance" in studies where joint or collaborative capacities are measured, such as those designed by mainstream studies (see above) may be equally determined and shaped by an array of factors escaping the here-and-now of the interaction, as well as their inferential reasoning skills. Below we discuss a few of what we see as major issues present in current developmental accounts of joint actions.

a. Inter-actional abilities are context- and experience-dependant

Or, as de Barbaro and Fausey (2022) have put it, "what you do is what you learn". Experience matters, even (or maybe even more so) in the relatively short life of an infant. Everyday experiences vary within and across educational communities and families, and this variation has implications for infants' and children's acquisition of skills and opportunities over time. For instance, the daily experience of being addressed as a speaker through infant-directed speech increases infants' expressive vocabularies by 24 months, with

important differences related to the extent (e.g., in terms of time during the day) to which children were exposed (Weisleder & Fernald, 2013). Playing with older siblings impacts the quality of social pretend play, resulting in longer and richer play sequences during preschool age (Fantasia & Nomikou, 2022).

Motor skills inextricably shape and are shaped by bodily affordances and constraints within a given environment and its action possibilities (Thelen, 1994). For instance, longer, increased opportunities to move independently accelerate sitting and walking (Rachwani et al., 2020). Engaging in new sequences of actions is context- and experience-dependent (Adolph, 2019) as it entails learning through practising new motor behaviours. Practising the same actions every day brings flexibility in how the action sequence is performed and refines the motor patterns behind it, e.g., when infants' outstretching arm movement gradually turns into a successful reaching pattern towards an object (Thelen, 1995; Adolph and Hoch, 2019). Infants' learning pathways vary significantly in the timing of when and how motor, language or cognitive capacities appear (Iverson, 2021). Joint action research in infants and young children has hardly taken this variability into account: they adopted pre-designed interactional tasks aimed at assessing joint problem-solving activities and their underlying abilities (such as role-reversal, joint attention, or perspective-taking), with little or no reference to infants' daily experiences of sharing or witnessing activities around them (Fawcett & Liszkowski, 2012). Task structure and materials employed might be unfamiliar, difficult to perform or even odd for infants, for instance, having to simultaneously pull two opposite handles on a tube to retrieve an object or activate a toy: or coordinating movements to make a ball jump on a trampoline. Home experience with objects, for instance, might differ from that required in experimental tasks, and motor skills or repertoire might not transfer to other types of action sequences (Fantasia, De Jaegher & Fasulo, 2014b). New tasks may need more time to be competently mastered, as research indicates that infants do not perceive affordances when they first acquire a new skill in development (Adolph & Berger, 2015).

Caregiving styles and the quality of parent-child relationships also have different impacts on infants' social learning experiences and competencies. Individual differences in how caregivers structure the physical space at home (Rachwani et al., 2020), or their attitude towards encouraging or restraining their infants' movements, affect infants' opportunities to sample their environment (see Flóres et al., 2019, for a review). For instance, infants' initial reluctance to touch plants encountered in a public garden decreased as parents approached those plants with exploratory gestures (e.g., by pinching or smelling their leaves), which infants then imitated (Fantasia et al., 2021). Other "contextual" aspects concerning how interactional dynamics are learned and shaped are too often dismissed or overlooked, for instance, the impact of situated socio-cultural practices or norms around involving infants in shared activities or the societal-based beliefs and practices about how to raise and educate its young members. The extensive works by Barbara Rogoff on the cultural nature of human social cognition development (Rogoff, 2003) or the invaluable ethnographic studies by Elinor Ochs (Schieffelin & Ochs, 1986) and Akira Takada (2020) provide evidence of the imprint that social-cultural practices and beliefs have in shaping the participatory possibilities and restrictions offered to infants and young children in their daily interactional environments. Such cultural differences are now flagged as important considerations in developmental research, but the extent to which these "contextual" aspects, such as social experience, engagement history, and the types of activities valued by a culture, influence infants' learning in joint actions over developmental time is an open question worthy of further investigation.

Endorsing an ecological stance to developmental investigations, that is, taking the situated context (or environment) in which interactions are experienced seriously into account - including learning processes of how to coordinate actions and goals - may shed light on the possibilities and restrictions of participation in social contexts and their impact on joint action capacities. Gibson's early observations that the ecological relationship of any organism, animal, or human being with its environment structured its affordances (possibilities) for action (Gibson, 1988) is as relevant to understanding the psychological development of joint action today as it was for Gibson's original work on the role of action in perception. Karen Adolph discussed the importance of this self-environment coupling from a developmental standpoint:

"Animals and their environments share a reciprocal relationship; together they form a behavioural ecosystem (Adolph & Berger, 2006). The *animal* in the system is a specific example of some type of animal at a specific point in its life history (e.g., a particular 12-month-old human infant with a particular body and brain, and particular skills and experiences). Reciprocally, an *environment* is the accessible surroundings for a particular animal in a particular place and time The environment includes surfaces, places, objects, other active agents, and the medium that encompasses it all (air for humans and birds, water for fish, etc.) (2019, p 121)."

b. Jointly acting with others is a learning process rather than an end-product

Developmental studies on joint action are typically committed to pinning down the presence or absence of a particular individual ability at a certain developmental time point, such as being able to coordinate actions toward a shared goal or its underlying components, such as inferring the other's intentions and coordinating one's own action with that intention. Gerson and Woodward (2010) defined this as a "snapshots" approach, "seeking evidence for focal abilities at a particular point in time but not considering how these abilities may change as a function of other events in development."

Forms and possibilities for participating in co-constructed sequences of shared actions (and goals) are manifold in the ecology of infants' daily social interactions: from attending to joyful shared book reading experiences, where the infant's and parent's converging attention is scaffolded by what is being read and its enactment through gestures, vocalisations, and postural adjustments (Rossmanith et al., 2014), to anticipating and adjusting to the adult's ongoing course of action when they are approached and then picked up (Reddy, Markova & Wallot, 2013; Fantasia et al., 2016). In all these forms of inter-acting, infants (but adults too!) learn how to

participate in dialogical, fast-changing dynamical processes by flexibly adjusting, coordinating, evaluating, sustaining, or repairing online actions in relation to those of their interactional partner(s).

Shifting our scientific focus from discrete, present-or-absent products of social interactions to process models of interactions (de Barbaro et al., 2013) may support a dialogical view of joint actions as co-constructed: daily interactions at home or at the childcare centre present routines organised around formal structures, which help infants learn when and how to contribute, e.g., by anticipating others' incipient actions (Lerner, Zimmerman & Kidwell, 2011). This is made possible by local and recurrent scaffolding practices by adults enabling learning how to "make use of the actions produced by others as context for the composition and placement of their own contribution" (ibid 2011, 44). Empirical and experimental research can be improved towards this end by taking into account the unfolding co-constructed nature of joint action. For example, Tomalski et al. (2022) analysed the effects of subtle parent-infant interactions (e.g., reciprocal looking, pointing, talking) during a standard 'solitary' infant eye-tracking paradigm and found that these background interactions significantly modulated infant task performance. In other words, what was assumed to be autonomous infant performance was in fact modulated by its dyadic, co-constructed context. Similarly, in a seminal text on infant mind-reading, Reddy (2008) demonstrated how domestic pets sharing in the daily routines of home-life joined the co-construction of a time- and context-appropriate routine with human-oriented signals and behaviours.

It is important for empirical and experimental research to acknowledge this fundamental role of co-creation and sharing in daily practice routines, especially when considering study design and interpretation of data. The regular pattern of daily and weekly habits presents windows of opportunity for infants or children to share their experiences and display their competencies. This is well known by education practitioners, who, for example, establish routines for sharing emotional and social understanding within regular, safe envelopes within the daily and weekly ritual (Lerner et al., 2011). In the United Kingdom, Nurture Groups formalise this principle with established, regular patterns that afford children with social and emotional difficulties safe opportunities for expressing their emotions and interests, and for engaging in academic cognitive tasks within sensitively co-constructed settings to enable reflective awareness and improved affective self- and co-regulation, setting the preconditions for cognitive classroom success (Boxall, 2002; Billington, 2012; Delafield-Butt & Adie, 2016; Trevarthen, & Delafield-Butt, 2015).

One of the major criticisms addressed by Enactivism and Dynamic System approaches to mainstream research looking into social cognition from a mind-centred perspective is their static view of the interactional process as a mere individual intention-reading endeavour, which misses consideration of its unfolding dynamic of inter-action. Intentions, plans, and actions are not pre-given and static but can be generated and transformed in the process of interacting. We agree with Fuchs and de Jaegher's contrasting view on "the idea that intentions are hidden and inner, or static, ready-made, and waiting to pop out of us for another to pick up (...). Rather, we participate with the other in the emergence and transformation of intentions, affects and understandings" (Fuchs & De Jaegher, 2009, p. 482). This is especially true of studies adopting pre-designed joint tasks, including those seeking to pinpoint when inferential capacities are developed enough to showcase in joint goal-ended activities.

Overall, researchers have advanced strong arguments toward adopting dynamical, embodied (Fantasia, De Jaegher & Fasulo, 2014b; Lux, Gredebäck, Non & Krüger, 2022; Maruyama et al., 2014; Reddy, 2008), and ecological (Adolph, 2019) theories and methods to the study of social interactions and social cognitive development. We deem these arguments equally necessary to studying how humans learn to become skilled, engaged participants in shared activities.

c. Joint actions occur along multiple timescales

Social interactions involve different sequences of activities progressing at different interconnected timescales. Within these activities, individual behaviours also unfold progressively, across distinct yet interconnected timing (de Barbaro, Johnson & Deák, 2013; de Barbaro & Fausey, 2022), becoming shared insofar as co-participants orient to each other and coordinate their actions and motives in sequentially organised ways to accomplish their local interactional business. Behaviours at different temporal scales have their own emergent properties. "In order to adequately describe and model the multi-scaled nature of human behaviour, we need hierarchical methods that can identify or integrate shifts in activity across these temporal scales" (Xu, de Barbaro, Abney & Cox 2020). In the natural flow of mundane social activities experienced by infants, joint actions are constructed over sequences of behaviours that lead to inter-acting and accomplishing interactional tasks together. In early face-to-face patterns of social play activities, for instance, mothers displayed preliminary sequences of gestures and vocalisations to invite infants to participate in the forthcoming playing activity (Fantasia et al., 2019). These preliminaries to the play sequence included: arranging the participatory space by moving the infant to face-to-face positions to create a shared attentional space, followed by multimodal sequences of gaze and high-pitch interrogative vocalisations accompanied by gentle movements of the infant's body. Similarly, in the final part of this paper, we describe a first-hand experience of a bedtime routine as a sequence of activities coordinated through a complex set of embodied vocal exchanges and resulting in the mother's telling a story to accompany her infant into sleep (infant's waving goodnight to the rest of the family, brushing their teeth, putting on the pyjamas, getting into the bed, and waiting for the mum to fetch and put on her own pyjamas before reading a story).

d. Moving matters

Embodiment has become an important concept in many areas of cognitive science, including developmental studies. There are, however, very different notions of exactly what embodiment is and how bodily features are considered for what kind of embodied cognitive development. Developmental research on joint action committed to investigating simpler and developmentally affordable forms of acting together (whereby less cognitively sophisticated abilities are involved; see Carpenter, 2009) have often considered

coordination of micro-level aspects, such as gaze, posture, limb movements as macro-level behavioural unities underlying individual capacities, such as for instance joint action, joint attention, social referencing, or mindreading. (a very common research practice in developmental literature, see [de Barbaro & Fausey, 2022](#)). Yet, bodily expressions have a right to exist on their own. With some differences in the tasks employed, the body is often considered as a medium for the acquisition (and enactment) of higher cognitive achievements; in other words, a tool in service of emerging, high-order cognitive skills devoted to understanding others' mental states through inferentially driven explanations for their actions. Theoretically, this means that bodily actions are regarded in service of emerging cognitive capacities (e.g., mind-reading representational capacities, later to develop into higher inferential mechanisms) and only rarely considered as a driving force shaping infants' cognition per se. It also means that infants' motor development trajectories are rarely considered.

Focussing the study of joint actions in infancy as emerging mental structures and their motor expressions overlooks the role of motor development as a learning process and how this enables and constitutes cognitive development instead of simply supporting it (see [Thelen, 1995](#); [Thelen & Smith, 1994](#); and [Adolph, 2019](#), for a review). Through exploring the environment, observing their own and others' behaviours, and testing and failing new actions, infants and children learn to act, explore, understand inter-action possibilities and flexibly adapt to the action affordances around them. Designing fixed and pre-structured situations for jointly acting overlooks the many rich ways in which the body adapts or learns to fit with the affordances and the limitations provided by interactional situations. For instance, how actions are adapted online to coordinate synchronous pulling, lifting, or pushing movements with that of a partner; or how dynamic system aspects are managed, e.g., the amount of force employed for pulling, lifting, and moving objects; or the maintenance of postural balance needed to simultaneously use hands and standing up. Only by experimentally varying and recording changes in those aspects (in terms of affordances or constraints for action) research may assess the extent to which infants flexibly tailor their behaviour to the demands of different interactional situations. As a wealth of research has already shown, there are many ways to put this into practice: for instance, by modifying the initial condition (e.g., the physical environment) in which the display of certain skills appears to be facilitated; or, on the contrary, in which previously possessed skills seemed to disappear or become impoverished (the so-called u-shaped progress). This is the case of the many studies on newborns' varying performance in walking when positioned on a flat surface compared to when immersed in a pool (see [Smith & Thelen, 2003](#) for a review). Or by manipulating the context in which spontaneous reactions may emerge, as in the many alternative paradigms for testing infants' perception of affordances and depth with different apparatuses (see [Adolph, Kaplan & Kretch, 2021](#), for a review). Or alternatively, by using tech-advanced devices to grasp selected, meaningful aspects of spontaneously occurring behaviours, which can expand current knowledge of the existing developmental variability in certain micro-patterns or capacities. For instance, in a study on infants' anticipatory bodily adjustments to being picked up, Vasu Reddy and her colleagues ([Reddy, Markova, & Wallot, 2013](#)) combined behavioural observations with data from a sensor mat to capture changes in infants' bodily pressure while lying down and waiting to be picked up. In this way, non-visible micro-movements were recorded over time, showcasing infants' variability in responding to the incipient action unfolding as a variable of the adults' "picking up style" (e.g., performing fast or slow movements; displaying a more or less clear approach by using the hands and/or voice; staying upright or bending down towards the infant).

Under a mainstream view, learning to act jointly (in the ways exploited by mainstream studies) is a primarily cognitive activity whose investigation entails how young children learn to *cognise about*, rather than *act within*, social contexts. Everyday joint activities in infancy are varied and include different types of participatory possibilities, objects, more or less structured spaces and timing. All these allow infants to learn through and by different ways of doing things together. To conclude, standard accounts consider participating in shared actions primarily as an encounter of individual minds whose main activity is to achieve sufficient and correct inferential knowledge of each other's intentions and goals. While such a cognitive and inferential approach is at risk, at the very least, of being reductive, the problem of how cognitively-immature creatures (such as infants) gain knowledge about others' intentions in interactions remains open, for it escapes the changing and dynamical dimension of knowing as an engaged process of sense-making, with its unpredictability, inconsistencies, ambiguity, particularly important when taking into account the rapid learning and development of infants and young children. Viewing successful joint actions in childhood as resulting from an increasing mastering of cognitive mechanisms underpinning social knowledge raises questions about what this knowledge is made of, its natural developmental progress – and, un-trivially, how we can investigate it.

Far from being new, these concerns have been raised by researchers from ecological and embodied approaches arguing for a serious rethinking of how human social cognition development is conceived of and investigated, starting from an engaged and *engaging* epistemological reflection on the dynamic and situated nature of human social interactions ([De Jaegher, 2019](#)). As Karen Adolph said: "Unfortunately, the widespread practice in developmental psychology is to side-step the foundational descriptive questions and to jump straight to testing hypotheses about presumed mental abilities or neural underpinnings" (2019, p. 184).

Participants don't need theories²: an ecological, situated, dynamical and embodied perspective

"Long before we can make explicit attributions of intentionality, our embodied responses involve us in interintentionality" ([Stern, D. 1985](#)).

Our main purpose in this work is to reclaim time and sequence as structuring analytical dimensions to understand how humans

² The title refers to the work by Reddy and Morris: Reddy, V. & Morris, P. (2009). Participants don't need theories: Knowing minds in engagement. In I. Leudar & A. Costall (Eds.), *Against Theory of Mind* (pp. 91-107). London Palgrave Macmillan.

learn to engage in social activities from birth. We do so by emphasising the prospective nature of action and its constitutive visibility. Which means: actions (as movements) not only are always projecting our intentions and plans forward, but their meaning is visible and perceivable in the larger situated context and timescales in which they unfold. Such a perspective places action at the centre of joint action research when situated, ecological, and embodied methods and interpretations are valued. Such a shift, as we argue later in this work, may promote a view of social interactions as co-constructed, intrinsically cooperative joint endeavours with different degrees of complexity and engagement, yet always prospectively looking (i.e., goal-oriented).

A participatory view of joint actions: lessons from infancy

The idea that joint actions require participants to define joint tasks, set out shared goals and interlock intentions is rather new in developmental psychology. Since the early 1970s, observational studies in developmental research have provided increasing evidence of the intrinsically shared and coordinated nature of early mother-infant interactions (Bruner, 1985; Fogel & Thelen, 1987; Hubley, 1983; Hubley & Trevarthen, 1979; Stern, 1974; Trevarthen, 1979), yielding a substantial wealth of knowledge on the early inter-subjective co-construction of meaning between infants and adults, and shedding light on infants' social capacities, especially their participatory and agentic contribution to shared activities. Knowledge of this interactional dynamic has impacted clinical practice in neonatology and infant care, improving developmental and psychosocial health outcomes (e.g., Brazelton, 2006; Lester & Sparrow, 2010). By adopting micro-analytical methods for observing contingencies, pauses and timing of participants' affective engagements and dis-engagements, these studies pioneered evidence of how early interactive exchanges with adults shape infants' participatory capacity to tune in and coordinate turn-taking moments of engagement with pauses and silences (Brazelton, Koslowski, & Main, 1974; Kaye, 1979; Stern, 1971), or progressively more sophisticated participation in playful exchanges (Bruner & Sherwood, 1976; Ross & Lollis, 1987). With their intrinsically dialogical nature, early dyadic exchanges have been defined as emerging *cooperative* activities wherein "each of the subjects is taking account of the other's interests and objectives in some relation to the extra personal context and is acting to complement the other's response" (Hubley & Trevarthen, 1979, p. 58). Dynamical Systems theories have consistently shown that development is an accumulation of decentralised and local interactions occurring in real time (Smith & Thelen, 2003, p. 343).

Drawing on research that considers social interaction as a distributed and dynamic system, home-based micro-analytical studies have used a variety of continuous (second-by-second or frame-by-frame) behavioural coding methods to sample multimodal sequences of interactions, e.g. action bouts or patterns, shedding lights on infants' emergent capacity to participate in structured daily routines – routines that are usually seen as relying solely on the adult's guidance, but that are now recognised as dynamically co-created by sides, from the infant and from the caregiver. For instance, ecological observations of mother-infant interaction during breastfeeding at home revealed that from 3 months, infants show an increasing capacity of co-regulating affective expressions and social behaviours while feeding, which also contributed to longer breastfeeding duration (Tsubanaki, Kokkinaki, Triliva, & Karademas, 2022). Later on, at 6 months, infants complement and support joint feeding activities by opening their mouths in coordination or anticipation with the caregiver's approaching hand holding a spoonful of food (Negayama et al., 2021).

Early coordinated joint activity set the stage for the complex configurations of goal-ended joint interactions with objects that mainstream studies on joint actions seek to investigate. Observational studies looking at the dyadic coordination as well as the micro-dynamics of mother-infant activity have adopted continuous, quantitative behavioural measures, such as posture, proxemics and gestures, to investigate how experience with formal structures of joint activities support infants' coordinated contribution to the ongoing action, e.g., in social games (Fantasia et al., 2014a; Nomikou et al., 2017), or when they are being picked up from laying on their back (Reddy, Markova & Wallot, 2013; Fantasia et al., 2016; Negayama et al., 2015).

Studies adopting continuous, fine-grained multimodal measures have captured changes in the developmental trajectories of how infants triadically attend to objects and persons in interactions. As early as 3 months, infants and caregivers engage in spontaneous joint object activities at home, such as shared book reading (Rossmanith, Costall, Reichelt, Lopez, & Reddy, 2014). A gradual transition emerges during the first year of life in patterns of triadic activity with adults, with implications on the quality of shared engagements and the activities carried within them: although infants and their mothers manipulate and attend to objects together, between 4 and 9 months, a joint looking and touching activity decrease, along with diminished exclusive attention to the object manipulated by the adult (de Barbaro, Johnson, Forster & Deak, 2016). These results confirm that early patterns of shared activities are visible when developmental progress is considered in their multi-dimensional and dynamic nature, setting the stage for later, more complex joint interactions.

Ecological, dynamical and embodied views

Although occupying a substantial part of infants' daily time in the first months of life, routine activities have been understudied and overlooked as developmentally relevant places to make sense with and of others' behaviours, emotions, intentions "creating the substrate for continuing cooperation" (Rossmanith et al., 2014, p. 18). Infant studies just presented are in line with evidence from research showing how the early development of perspectival knowledge and control is grounded in infants' neuromotor development and its intrinsic ability to anticipate and gain information from one's own actions, giving the system structure for understanding other's actions, too (von Hofsten, 2009).

Moving away from purely cognitive views, developmental researchers have increasingly adopted less sophisticated and behaviourally bodily focussed methods, e.g., proposing that some basic forms of action, affective dynamics, and attentional mechanisms may support young co-agents in establishing mutual knowledge and coordination. A recent study by Suarez-Rivera and colleagues (Suarez-

Rivera et al., 2022) systematically explored the rates, timing, and modalities of mothers' and infants' spontaneous joint engagement in activities at home. Their findings highlighted that joint engagements have tailored temporal structures and occur in bursts (rather than in a continuous or predictable flow), and interactants' physical proximity sets the stage for when and how joint engagements will occur. The implications for research are noteworthy, as the authors suggest "(...) developmental phenomena hypothesized to be key for infant learning should be investigated in laboratory and everyday contexts alike to understand real-world inputs that propel learning, and to model those inputs to further elucidate potential mechanisms of change" (ibid, 2022, p. 250).

Dynamic Systems research has looked at how bodily sensorimotor processes underlying co-agents' individual movements structure individuals' social behaviours through mutuality of influence in terms of temporally coordinated and superordinate sequences (e.g., van der Wel & Fu, 2015). Joint attentional processes, for instance, have been regarded as essential elements in lower-level, interactionally focused forms of joint actions, such as the capacity to coordinate and share attention with others. (Carpenter & Liebal, 2011; Siposova, Tomasello & Carpenter, 2018). Carpenter and Liebal (2011) proposed that co-interactants establish shared knowledge about or express commitment to perform a joint goal through ostensive communicative processes that can take the form of a reciprocal, timely coordinated gaze exchange or through imitative behaviours. By *sharing looks*, interactants use their gaze to share "the experience of sharing" (Hobson & Hobson, 2008, p. 79). Philosopher Debora Tollefsen (2005) has proposed that attending to each other's actions provides participants with a *shared perceptual space* co-constructed through joint attentional dynamics. In this shared space, actions and their intentions are perceptually overt and identifiable so that even young children without a "robust theory of mind" (ibid, 2005, p. 81) can engage in shared activities. In this view, reaching awareness of others' intentions would not necessarily require an inferential stance insofar as co-agents can implicitly track each other's intentions in their actions.

A few yet important differences characterise cognitively-focused approaches and those endorsing ecological, dynamical, embodied stances. One crucial difference entails which role is assigned to the body and its development in relation to the environment, or the extent to which it is conceived as a primary source of development in other domains (Costall & Leudar, 2010; De Jaegher, 2009, 2019; Gallagher, 2008a; Reddy, 2008, 2015). The body – one's own and others' – and its manifestations as constitutive of how humans learn and gain knowledge through and about the surrounding world (Adolph, 2019; Reddy, 2008; Thelen, 1995). Quoting Adolph and Hoch:

"Psychology is the study of mind and behaviour, and all behaviour is motor behaviour. Thus, the study of motor development is really the study of behavioural development. As such, motor behaviour both requires and reveals the workings of the mind" (2019, p.142).

Under this view, motor development (as specific actions and bodily behaviours) drives learning and acquisition of new knowledge as it affords possibilities for concrete, repeated experiences, e.g., hindering or supporting what can be perceived, explored, encountered, or socialised.

Additionally, dynamic and situated approaches have traditionally considered developmental trajectories as interconnected rather than causally-related systems (Forster, 2002; Isabela Granic & Patterson, 2006; Hutchins, 1995; Oyama, 1985/2000; Taylor, 2000), that is: not searching for traditional causal relationships but rather trying to identify multimodal behavioural patterns, their variability over time and contexts and the conditions which lead to a shift in the trajectory (de Barbaro et al., 2016). As discussed earlier in this work, infants learn to move, act, coordinate or resist others' actions in the ecology of their everyday life, participating in activities that extend through different timescales. "Intentions and goals are not searched before or behind the communicative action as its "cause," but manifest in speakers' behaviour, shaped and adjusted as the interaction unfolds." (Fantasia et al., 2014b). Experiencing others' actions and intentions is "apprehended through the senses, which, in the 'natural attitude', are always in combination with our empathetic co-presence with other living bodies, whose actions reveal their intentionalities (i.e., their being 'about' something, oriented toward some place as a physical and social goal)" (Duranti, 2010, p. 27).

Studying how infants learn to become participants in social activities should start with close observations of how infants and young children spontaneously do things in their most familiar environment(s), e.g., their early social responses, such as vocalising out loud when they have a nappy changed or displaying coyness as they turn the face away with a shy smile when someone addresses them directly. But also, it should consider what they make (or not) of the many affordances offered by those environments and the extent to which others' behavioural sequences are perceived as affordances for their own actions, e.g., by producing vocalisation or movements at specific relevant points within an interactional sequence. This means identifying what is relevant and meaningful for infants in the ecology of their daily experiences, or habitus (Bourdieu, 1995), and how it changes as a function of time and learning.

An ecological and situated approach does not necessarily imply doing research outside labs in the real world. Answering developmental questions requires both naturalistic studies and controlled laboratory experiments, as long as researchers stay true to the phenomenon under study, acknowledging its complexity and its function within infants' or young children's ecology of daily experiences and developmental paths (see, for instance, a naturalistic semi-controlled study on infants' approach to plants by Fantasia et al., 2021). In practical terms, this means combining observational and tech-based tools (e.g., technologies for sensing and activity recognition; de Barbaro, 2019) in order "to collect multiple, precisely time-locked and synchronized dimensions of organism-internal and external factors repeatedly and at scale. Such data will allow developmentalists to begin to disentangle complex reciprocal interactions between these factors and quantify how they drive developmental trajectories, a truly radical proposition for developmental science" (de Barbaro et al., 2016).

The prospective nature of action: an embodied foundation of joint action

"Actions carry one into the future. Prospectively controlled actions bring one from "where one is" to "where one wants to be"."
Delafield-Butt et al. (2018, p.1)

Human movement, from the beginning of life, is never simply the execution of a motor command from higher cerebral centres of activity. Rather, it is inherently future-oriented, necessarily prospectively organised with anticipation of its consequent effects (von Hofsten, 2007), from the beginning of a human life (Delafield-Butt & Gangopadhyay, 2013; Delafield-Butt et al., 2018). Physical forces of inertia and momenta require an action to be intrinsically organised ahead-in-time so that it reaches its target in an efficient and effectual manner (Lee, 2009). Too much force and a movement can overshoot. Too little, and it has no effect. The lawful order of parsimony necessitates action must be organised prospectively. Even before birth, careful kinematic analyses of ultrasound recordings reveal human foetal movements to be anticipatory, expectant of their sensory consequences from as early as the fourteenth week of gestation (Castiello et al., 2010; Zoia et al., 2007). For example, in the case of twin pregnancies, twin foetuses have been observed to make special twin-directed movements distinct from movements toward inanimate objects in their uterine environment, such as the placenta wall or umbilical cord. Computational analysis of these movements' kinematics reveals differentiation at the start of the action, contingent on its end effect: action patterns used to touch the twin were the same as those used to touch one's own face, but distinctly different from those that touched the inanimate objects of the environment, indicating not only that action end-effects were known ahead of time, but that this also suggested very early 'social' awareness of an animate 'other'. Similarly, Zoia et al. (2007) found that in singleton pregnancies, motor planning was evident in the action pattern of the arms structured by their final position, or 'goal', by at least twenty-two weeks of gestational age. These data indicate an early prospective awareness of the sensory consequences of self-generated action evident in the actions of the second-trimester foetus.

Prospective awareness in early human development *in utero* is also evident in a number of studies of the anticipatory movements of foetal hands and fingers to a target part of the body. For example, Myowa-Yamakoshi & Takeshita (2006) and Reissland, Francis, Aydin, Mason & Schaal (2014) observed the mouth of the foetus opened before actual contact, during mouth-directed movements of the arm that carried a finger to it. This indicates an early, embodied awareness of the self-generated future manifest in sensory-motor movement, an anticipation of self-generated sensory consequences. Finally, at birth, detailed high-precision subsecond kinematic analyses demonstrate that newborns' arm movements are prospectively organised with an anticipation of their outcome in body-space, as analysed by a computational theory of prospective motor control (Delafield-Butt et al., 2018).

Altogether, these studies demonstrate that a cardinal sign of anticipatory experience, goal-directed motor control, is evident from the start of the second trimester *in utero* and develops in complexity and precision through foetal life as the skeletomuscular and neural connectivity matures (Delafield-Butt & Gangopadhyay, 2013). This intrinsic prospective nature of human movement is the basic psychomotor foundation on which a young human life experiences the world. Embodied, lived experiences are powered by the interest and intentions of the individual. They are exploratory and sensory seeking, testing the effects of one's interest and agency on the objects, events and people that make up the child's world (Baldwin, 1895; Piaget, 1953). These experiences are parcelled and organised by single, simple actions such as a reach-to-touch or a gesture of the hands to another person in affective expression, constituting the foundation on which joint action is engaged and learned. In human infant life, our gregarious nature and close adult-infant care often nest these simple actions into an intersubjective system of shared engagements, giving the basis for the co-construction of patterns, rituals, and games that the developmental science of joint action attends to (Trevarthen & Delafield-Butt, 2013).

Humans' action-perception loops power and generate prospectively coordinated and synchronised social movement, structured by an early form of pre-reflective intentionality intrinsic to its prospective nature (Delafield-Butt & Gangopadhyay, 2013). This basic motor intentionality, technically identified as 'primary sensorimotor intentionality' (*ibid.*) or as 'intention-actions' following Searle (1980), is a primary form of intentionality that is pre-reflective and embodied in the act of moving, setting the foundation for later capacities of reflective, conceptually-backed intentions more familiar to adult experience. In development, the infant's capability progresses from these primary sensorimotor intentions intrinsic to the prospective organisation of movement to a developed capacity for abstract reflection and the ability to imaginatively plan actions with distal intentions. This is a development from what Searle identified as intention-actions, where the intentional aspect is part-and-parcel of the structure of the movement, to the capability for an intention-to-act, its abstract representational form that affords offline reflection before the act of moving (Searle, 1983). This latter form is typically associated with adult volitional intention and, in lay terms, is conflated as simply 'intention'. However, the prospective nature of human movement necessitates structured action in anticipation of its consequent effects, a powerful template for understanding the effects of one's own agency on the world that "remains fundamentally unchanged and continuous throughout life, driving and shaping development, learning, and cognition" (*ibid.*, p. 404).

This ontogenetic primary embodied intentional form does not only enable one's own actions to be recognised and experienced as goal-directed and intentional. Crucially, it informs the perception of the other's bodily actions as also agentic and purposeful. The concept of *pairing*, as proposed by researchers in developmental phenomenology, supports this idea of a primarily bodily-centred experience of knowing and understanding others in interactions. *Pairing* refers to a perceptual and experiential process of self-other similarity based on, for instance, vocal or proprioceptive-visual domains (Vincini & Gallagher, 2021; Vincini & Fantasia, 2022) acquired during early social experiences. Within this perceptually pairing process, "having experienced characteristic features of embodied animacy, emotions, intentions, and perceptions in myself, in so far as the other's body presents the same or similar features, I tend to perceive it as embodying animacy, emotions, intentions, and perceptions" (Vincini & Gallagher, 2021, p. 190). In other words, understanding others becomes possible through paired proprioceptive and exteroceptive bodily experiences of feeling and perceiving others as co-partners in interaction.

A similar proposal has been advanced by neuroscientists working on the role played by mirror-neurons mechanisms to explain human intersubjective engagement (or the Direct-Matching hypothesis, see, for instance, the works by Vittorio Gallese). Direct perception of another's intention is available without inference or reflection by way of direct perception of the other's intention manifest in its intentional action form, the shared understanding of which is thought to be underpinned by the mirror neuron system

responsive to the action's intention, and not its biomechanical representation (Gallagher, 2008a; Gallese, 2001). The major claim advanced by embodied neuroscience approaches is that basic neuromotor resonance of intention enables an implicit knowledge of others' intent, allowing one to understand and participate in purposeful action sequences. Higher-level mental processes of reflective awareness would then emerge later in development, informed by earlier mechanisms of action coordination (Tollefsen & Dale, 2012; Trevarthen & Delafield-Butt, 2015). Notwithstanding, Reddy and Uithol (2016) have proposed a compelling contrasting argument to the idea that a *pairing or mirroring mechanism* can fully explain how individuals intersubjectively understand others, for "it facilitates a fundamentally spectatorial stance, ignoring engagement and dialogue; it focuses on similarity between self and other and neglects difference; and it succumbs to the static terminology of mechanism rather than a dynamic analysis of the process (Reddy & Uithol, 2016, p. 101).

Studying joint actions in early infancy and childhood is to investigate how humans learn to become competent participants in shared social actions. Whatever theoretical standing point one endorses, limiting this investigation to the analysis of how individuals represent intentions and goals in pre-defined situations with scripted endpoints means having access only to a small part story. As proposed by situated and embodied approaches, analysing daily human behaviour can provide a greater phenomenological understanding and investigation of the nature of social cognition, including joint actions. The third part of the present work considers the sequential structure of early social interactions and their intrinsic prospective and perspectival character as key analytical lenses through which investigating how infants and adults make use of their body, voice, and affective expressions to understand, anticipate and coordinate with each other in daily situations. Such an approach is amenable to elucidating cultural differences in the timing, tempo, and expectations that develop in early infant-caregiver interactions (Frank & Trevarthen, 2012; Gratier, 2003; Gratier & Trevarthen, 2008; Negayama et al., 2015, 2021).

Temporality as key for understanding joint actions in infancy

"The present, and by that is meant not the point which indicates from time to time in our thought merely the conclusion of "finished" time, the mere appearance of a termination which is fixed and held, but the real, filled present, exists only in so far as actual presentness, meeting, and relation exist. The present arises only in virtue of the fact that the Thou becomes present."

M. Buber, (1923), *I and Thou*

Alternatively to being a measure of something standing between different events, time can be conceived as a relational entity, binding together different parts of an interactional event in "a relationship that is constitutive of the organisation, and hence the meaning, of those parts" (Rawls, 2005, p. 167). An interactional view of temporality may frame it as the lived experience of doing something with another person in situated time, encompassing different forms of *experiencing* time in relation to another person: the intrinsic phenomenal sense of being in relation *with*, embedded in the temporality of situated, present engagements; the narrative quality of shared, temporally defined sequences of bouts of reciprocal actions; and finally, the forward-looking, dynamical and *incomplete* quality of interactions where movements and actions can be expected, invited for, anticipated or complemented by co-participants moments-by-moments.

Infant observation studies have shown infants' extreme sensitivity to the synchronised contingency of reciprocal behaviours in early shared routines (de Barbaro & Fausey, 2022). A special kind of temporality is embedded in the sequential unfolding of caregiver-infant daily meaningful activities, such as bathing, eating, and playing, in which participants are oriented towards understanding their reciprocal actions through a continuous alignment and synchronisation of individual temporal trajectories of gestures, vocalisations, posture, affects (Fasulo 2015). This takes shape in temporally-organised sequences of gesture and voice patterns (Fantasia et al., 2014b), producing a narrative in their pattern and form (Malloch & Trevarthen, 2009; Trevarthen & Delafield-Butt, 2013). At the micro-level, early interactional experiences are organised and shaped by temporal dimensions of *prospectivity and sequentiality*. We will now present both, framed within a developmental perspective.

Prospective knowledge structures narrative stories of the infant's action

On the surface, human experience appears continuous and regular as a steady 'stream of consciousness' (James, 1890). But it is made up of discreet events that unfold in time, with beginnings, middles, and ends. These discreet events are made in active engagement with the world by a physical body expressive in action of the interests, knowledge, and intentions of the Self-as-Agent (MacMurray, 1957). From the beginning of life, infants' movements are self-generated, prospectively organised actions anticipating their consequences (see above). Rhythmic structure, the moment-by-moment coordination of verbal and non-verbal rhythmic patterns, is fundamental because most of our motor behaviours are rhythmically organised (Lashley, 1954).

As embodied agents, these participatory experiences become structured by the spatiotemporal properties of the body, and critical to this is its purposeful, future-oriented nature that seeks to create meaning in social interactions with the surrounding world (Reed, 1996; Trevarthen & Delafield-Butt, 2017; Delafield-Butt, 2018). Each action is a self-generated movement of the human body that is necessarily organised prospectively from the beginning of life, *i.e.*, organised with an anticipation of the future it is actively structuring and creating (Delafield-Butt, 2018). These actions and their consequences are affectively evaluated by the agent for their success or failure, benefit, or threat, as are those of an observed other (Bråten, 1998). The future-oriented nature of human motor control – both solitary or social – asserts a powerful prospective structure to knowledge as anticipatory and structured by the self-generated bodily experience, driven by the agent's intentions and interest, and excited by the affects that motivate them.

How do early bodily-grounded and structured experiences and social cognition development relate to each other? Seminal

psychologist Jerome Bruner (1986, 1990) identified two forms of cognition. On the one hand, he proposed there exists a *logico-scientific* mode of cognition, able to form concepts with knowledge of their lawful properties. This is a timeless, abstract knowledge that enables a technical intelligence with its rich knowledge of facts that can be approached from multiple imagined perspectives to work with the generation of plans for desired futures and solutions to present problems. A static, timeless mode of cognition is the focus of educational systems. Cherished by technological society, its capability is what we normally identify as intelligence. Its development in early life is the focus of the cognitive and social cognitive paradigm, which attends to it almost exclusively. On the other hand, the prospective structure grounded in human motor control delivers a necessary unit that opens, then gives closure to activities and engagements to draw experience to a purposeful conclusion, and to make sense of them. In this way, human agency seeks to parcel what seems to be continuous experience on the surface into the embodied experiences of purposeful events. In other words, into stories.

Bruner (1986, 1990) identified another form of cognition, what he defined as *narrative* cognition. Narrative cognition places the lawful properties of objects and persons into an animated temporal order, making sense of their events and processes as they relate one to another through lived, situated time. The logico-scientific knowledge of objects and persons becomes contextualised through this type of intelligence, or way of cognising as their relations, psychological motives, and the feelings that power their interactions become known and meaningful within the context of lived embodied experience. That experience is held in memory as a narrative, the story of its past experience made available as knowledge for recalling in action or re-telling in expressive gesture or words (Delafield-Butt & Trevarthen, 2015; McGowen & Delafield-Butt, 2022; Gallagher, 2008b).

Solms and Panksepp (2012) proposed a model of the nested hierarchy of mind-brain processes, which takes into account the development of experience from its ontogenetic primary, embodied, and direct experience through to higher cognitive, reflective processes. We draw on this model to propose an integrated model (Figure 1) for social engagement, supported by detailed neuropsychological study (see Panksepp and Biven, 2012, also Trevarthen & Delafield-Butt, 2017). At the ontogenetically basic level, primary (1°) process affects and intentions are predominantly brainstem mediated, embodied, and expressed in direct intention-actions. Secondary (2°) level processes store these experiences as memories of routines of motor sequences and their responses and discriminate perceptions with greater affective associations. But importantly, it is the cortically mediated tertiary (3°) processes that afford reflective thought, abstract conceptual considerations, off-line planning, and verbal language. This tertiary level enables Bruner’s logico-scientific thought, but it is through narrative and sequences of bodily motor acts that this thought is communicated.

Communication between infant and adult happens directly at the primary level within embodied and situated routines, supported by personal knowledge and memories stored in secondary processes. The primary levels always directly communicate with each other between persons, while the brainstem is rich in sensory, evaluative, and associative learning, enabling conscious awareness (Merker, 2007). Further, the logico-scientific cognitions and reflective awareness available in tertiary processes must be fed back through the actions of the body, which are primary processes. As the adult possesses excellent tertiary resources, s/he can reflect linguistically and conceptually on those primary purposeful, expressive actions and their effects. This tertiary reflective self enables elaboration and conscious reflection of primary embodied actions, intentions, their affects and effects. And as such, one can regulate the interaction occurring through primary processes with the additional tertiary reflective self.

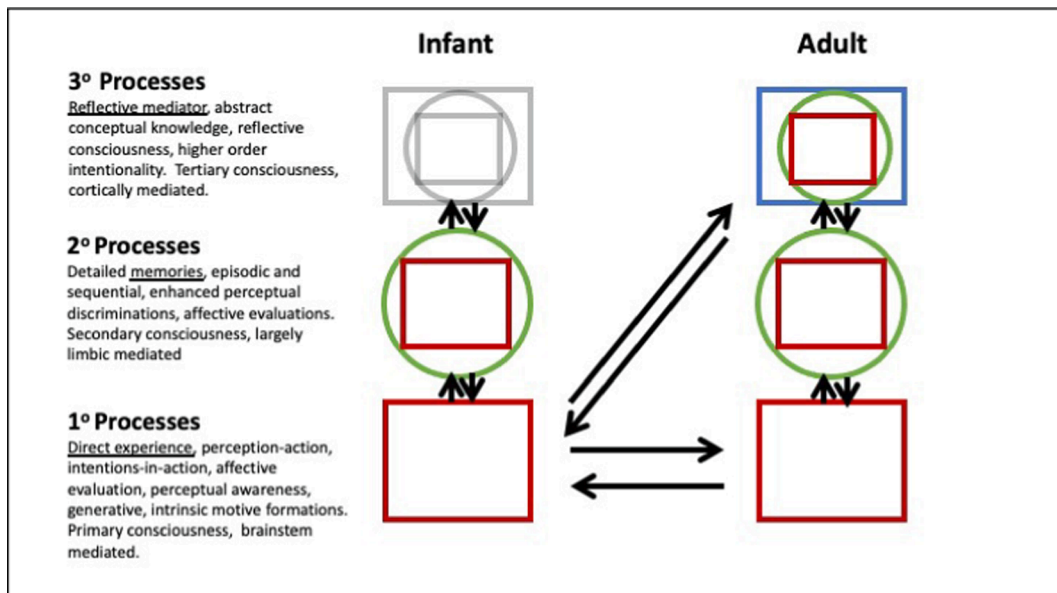


Figure 1. Schematic of interactions between infants and adults, organised by their nested mind-brain hierarchies. Primary processes mediated by brainstem systems directly communicate, enhanced with secondary process memories and affective evaluations. Tertiary processes mediated by cortical systems enable reflection and conceptual knowledge, particularly well developed in adults but not in infants (grey). These regulate primary process engagements, and in adults they serve as a ‘reflective mediator’, powered by language and logical abstract thought. Importantly the primary mode of communication is always through the primary process regulations of the body. (see Solms and Panksepp, 2012).

As lived experience necessarily progresses through time, all abstract, reflective logico-scientific intelligence must be expressed back through the animated temporal order of the motor expressions of the body, through its meaningful sequences generated in narrative – whether verbal or not. The two intelligences work hand-in-hand, informing and structuring each other, while knowledge – particularly social knowledge, is always contextualised within these stories, with a temporal sequence of events:

“a narrative is composed of a unique sequence of events, mental states, happenings involving human beings as characters or actors. These are its constituents. But these constituents do not, as it were, have a life or meaning of their own. Their meaning is given by their place in the overall configuration of the sequence as a whole.” (Bruner, 1990, p. 43)

From an interactional framework, prospectivity can be viewed as the future-oriented and forward-looking character of actions and their awareness by others in social encounters. That is: interactants can reciprocally anticipate, predict and partake in the ongoing unfolding of their action trajectories, envisioning what possibilities lie ahead. It is akin to what Gallagher identifies as ‘direct perception’ of the others’ intentions (Gallagher, 2008a,b). But because engagement in joint, social actions is a dynamic and participatory process of sense-making, the prospective nature of action structures social interactions as future-oriented, expectant, yet situated and grounded in the *here* and *now* of interactional time and space. A similar concept is found in ethnomethodology, which proposes *retrospection* and *projection* as two constitutive aspects of the lived experience of time in interactions. While retrospection is fundamental for interactants to display mutual understanding and accomplish intersubjectivity (Deppermann & Günthner, 2015; Deppermann & Streeck, 2018), projection works towards coordinating, anticipating and complementing the other’s intentions in action.

Within developmental literature, the essential role of prospectivity for developmental processes such as motor development and action-perception systems for detecting affordances (Gibson, 1979) is well established, although somehow limited. The pioneering works by Esther Thelen and Linda Smith (Smith & Thelen, 2003; Thelen & Smith, 1994) shed light on prospective control as a key mechanism in developing motor capacities (walking in particular) and, consequently, for infants’ interactional and cognitive affordances and possibilities to explore the surrounding world. Along the same line, Reddy (2015) proposed that early meaningful (intentional) actions are prospectively oriented as possessing an intrinsic quality of incompleteness:

“Intentional actions are largely about effecting change in the world – whether pushing a door shut or looking around for food or calling someone to come. This quality of expectation or ‘incompleteness’ – of waiting for completion or effect or response - not only allows the intentionality of actions to be perceived by others but more crucially, it has the potential to invite others to engage with the incompleteness.” (Reddy, 2015, p. 2)

Incompleteness here has to do with not-knowing but also with learning-to-know as an engaged epistemic process of understanding and knowing others (their actions, intentions, emotions) by responding to their invitation to deal with, and address the here-and-now interactional incompleteness in different ways. It is precisely within this incompleteness that development becomes possible and takes place, as infants (and caregivers) make experience of an unfulfilled space between the other’s action and their own. And learn how to reciprocally fill that space; a space left open to experimenting and learning how the presence of the others - their moves, speech, affective displays - can make a difference, for better or worse.

Sequentiality

As systematically observed in ethnomethodology research, sequentiality is an ordering principle in mundane conversations in adults (Garfinkel, 1988), constitutive of the public and shared dimension of any social interaction where “each party builds upon structure provided by others” cooperatively (Goodwin, 2013, p. 17). Sequentiality is intrinsic in the system of turns in interactions, warranting visibility and accountability to the sequential interactive order (Goffman, 1983). In other words, social interactions (from the spontaneous meeting of friends on the streets to formally coded social events) are culturally-shaped and situated practices where co-participants understand and complement each other’s actions in a sequence-organised format where “actors share some set of expectations – or practices – that provide guidelines for how to use sounds and movements to create mutual understanding in particular contexts of use” (Rawls, 2005, p. 167).

Recently, a growing number of studies in psycholinguistics applied sequential analysis methods to look at how interactional meanings (and their linguistic enactment) are shared and co-constructed in joint daily interactions with infants and young children (Fantasia et al., 2019; Fasulo et al., 1999; Forrester, 2009; Nomikou, Schilling, Heller & Rohlfing, 2016; Ochs & Schieffelin, 1995). They shed light on the sequentially-locked structuring of early joint activities, as built around turn-shaped multimodal *sequences of units of actions* (or “action arcs”, Rossmannith et al., 2014), including turn shifts, pauses (Hilbrink et al., 2014), coordination of verbal and non-verbal aspects, synchrony or dis-synchrony in the temporality of each participant’s expressive movement (Berducci, 2010; Trevarthen, 1998). These sequences are the building blocks through which early participation is co-constructed (Duranti, 2000), as infants gradually recognise typical elements and progression of routine activities (e.g., opening and closings), learning to position their own expressions of interest and feeling (vocalisations, movements, looks) within them. In this sense, sequencing helps participants’ expressions to be organised, enabling their understanding and predictability. Interestingly, early dyadic exchanges seem to present similar sequential organisational structures to those regulating adults’ conversations, such as turn-alternation and turn-transition (Berducci, 2010; Hilbrink et al., 2014) and the use of preliminaries (Fantasia et al., 2019; Fantasia, forthcoming).

A first-hand experience of sharing a joint routine activity by the first author with her own child shows how an ordered, repeated sequence of actions makes participants’ actions recognisable, affording interactional complementarity.

“At about 12 months of age, D. started to make clear, intentional attempts to help me in what was, by then, our bedtime routine. He would sit on my bed, attentively looking at me whilst I was putting on my pyjamas before lying next to him to read a book.

One night, while I was searching for my pyjama pants all around the bed, he fetched it from the pillow behind him and passed it on to me. I stared at him cluelessly, convinced it was a fortunate coincidence. Only, the day after, he did it again. There was no doubt that D. really was handing me over what I was looking for, anticipating my action, and at the very least, displaying some knowledge of the sequential structure embedded within our nightly routine. He understood my intention nested within its action and interaction sequence, and in helping me achieve my immediate purpose to find and put my pyjamas on, he may also have been collaborating to move quickly to story-time, his favourite moment.”

D.'s actions appeared as meaningful and intentionally directed to help the mother's search, nested within a larger context, known to D. through the repetition of routine. The sequential structure of this bedtime example is imbued with a narrative dimension: a co-created story whose plot was known and expected at detailed levels by both participants (Delafield-Butt & Trevarthen, 2015). And yet, when the story was disrupted and at risk of becoming incomplete, when the pyjama trousers were missing, the visibility and situatedness of his mother's searching actions – although new – supported an embodied, situated understanding of her intentions (finding the trousers), and invited a complementary response by D. In other words, the early interactional formats of actions unfolded as interpersonally, coordinated motor intentions co-created and enacted into meaningful and organised sequences that took on a shared, recognised and anticipated narrative quality (Delafield-Butt, 2018). Repeated, changed, or modified over time, these formats become increasingly conventionalised, progressing in less idiosyncratic ways and more recognisable and predictable for both the infant and the adult. As co-actor in a daily repeated routine, D. had learned to make sense of the sequential unfolding of actions and how to complement them as part of a shared storyline developing at different timescales (including a preparatory phase of teeth-brushing, putting on his own pyjamas, waving goodnight to other family members and going to bed).

Human minds operate in nested layers of meaning, affectively imbued within the expectation of a culture, an intimate culture in this example; to reduce this multidimensional nesting of social, affective and instrumental meaning to only the last is to lose the meaning of the life of a human mind altogether. This 'hidden' knowledge of how things unfold between agents and actors is exactly the purpose of Bruner's identification of the 'narrative' mode of cognition (Bruner, 2004). One's intentions are not only private states but publicly available and part of a shared, situated interactional dynamic.

The sequential character of interactions plays a role also in the development of an embodied and situated sense of self-agency in relation to the surrounding social world. Interactants simultaneously perceive the other's expression (a gesture, a movement, an utterance) as well as their own movements as sequentially tied by means of “a complicated mechanism of retentions and anticipations [which] serves to connect one element of his [...] speech with whatever preceded and what will follow until the unit of what he wants to convey has been grasped” (Rawls, 2005, p. 181). Such a mechanism is grounded in the forward-backwards character of interactional temporality, defined here as its *prospective* nature.

Conclusions

In sum, we present the case for an expanded consideration of the nature of joint action from its current focus in cognitivist accounts on its representational and inferential aspects occurring in single heads to an inclusive perspective that additionally accounts for its embodied, situated, and prospectively structured nature that incorporates implicit knowledge within the flow of interaction learned from previous situated experiences. By doing so, we sought to deepen appreciation of cognitive theories for private humans' mental and cognitive experience to include its fundamental embodied, psycho-motor nature and its relevance for social interactions.

This new foundation for understanding joint action impacts on developmental science in complementary domains. For example, the study of infant imitation has remained polarised and controversial for almost fifty years, stuck between being considered a first sign of social communication (Meltzoff et al., 2018) or the simple, non-mental heightening of arousal in response to another's actions (Oostenbroek et al., 2016). The debate has hinged on whether the imitations observed in such experimental paradigms are statistically relevant to confirm or not infant imitation as a phenomenon. Such debate has been ongoing, with evidence on both sides, since its first observation (Meltzoff & Moore, 1977). However, by considering the temporal nature of acting together, the imitative phenomenon can be better described as nested within a sequence of ongoing expressive, intentional actions between participants, unfolding at different timescales at the service of multiple intersubjective purposes. Such a view presents an improved theoretical perspective that more comprehensively accounts for the infant's lived experience, enacted in time with another. Empirical analysis of the temporal sequence of expressive action in the paradigm is a logical next step to further elucidate this early form of joint attention, action, and potential for generating shared meaning (Heimann, 2022; McGowan & Delafield-Butt, 2023).

Similarly, when interpreting the behaviour of a child or young person with a neurodevelopmental or psychiatric condition such as autism spectrum disorder, intellectual disability, or Down's syndrome, superficial interpretation of an expressive action that does not take into account its context, and without attention to its timing, sequence, and narrative purpose within its situated routine and present circumstance, could lead to misunderstanding the person's action. In the case of autism, such superficial interpretation can miss the meaning of an actions as non-social, when in fact it may be socially intended and specifically socially meaningful (Jaswal & Akhtar, 2019). Education and care practitioners have adopted an axiom to improve their careful consideration of the potential of all non-verbal expressive action, no matter its idiosyncratic peculiarity, 'all behaviour is communication'.

The present work proposes a rethinking of how we conceive joint action and its development by considering the nature of human action as prospective, from the beginning of life; of what it takes to share a task in infancy to a more satisfactory redefinition of shared routines as early embodied interactional contexts, temporally and sequentially grounded, of which infants actively co-create participatory experience in, and gain knowledge of, through its shared active generation. On the side, it also offers a reflection on the different implicit ways of conceiving inter-action, or joint action, embedded in our practice of investigating social cognition, and the

deep theoretical as well as methodological implications these forms of conceiving joint action have. Doing things *with* others, understanding others' actions, and learning how to be understood is not a straightforward process. It requires a continuous engagement in sequentially relevant interactions that develop shared experiences, where actions (and the intentions and feelings behind or stemming from them) acquire narrative meaning as infants and caregivers attend to each other's interests, feelings, and intentions to share the focus of their attention in embodied, situated ways. Over time – developmental time and daily lived time – repeated formats of actions co-create a shared historical sequence, a time frame, which participants have “built... together out of their collaborative work and it would not exist except for that collaborative work and does not exist unless that work is successful” (Rawls, 2005, p. 122).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the Wallenberg AI-Humanities and Society (WASP-HS) project “Social interaction with autonomous systems” of which VF is PI.

References

- Adolph, K. E. (2019). An ecological approach to learning in (not and) development. *Human Development*, 63(3–4), 180–201.
- Adolph, K. E., & Berger, S. E. (2015). Physical and motor development. In *Developmental Science* (pp. 269–342). Psychology Press.
- Adolph, K. E., & Hoch, J. E. (2019). Motor development: Embodied, embedded, enculturated, and enabling. *Annual Review of Psychology*, 70, 141–164. <https://doi.org/10.1146/annurev-psych-010418-102836>
- Adolph, K. E., Kaplan, B. E., & Kretch, K. S. (2021). *Infants on the edge: Beyond the visual cliff* (pp. 51–72). Revisiting the classic studies: Developmental psychology.
- Baldwin, J. M. (1895). *Mental Development in the Child and the Race*. New York: The Macmillan Company.
- Berducci, D. (2010). From infants' reacting to understanding: Grounding mature communication and sociality through turn-taking and sequencing. *Psychology of Language and Communication*, 14(1), 3–28.
- Billington, T. (2012). ‘When they’re making breakfast they’ll talk...’: Narrative approaches in the evaluation of Nurture Groups. *Journal of Early Childhood Research*, 10(3), 318–331.
- Bourdieu, P. (1995). Structures, habitus, practices. In *Rethinking the Subject* (1st edition., pp. 31–45). Routledge.
- Boxall, M. (2002). *Nurture groups in school: Principles & practice*. Sage.
- Bratman, M. E. (1992). Shared cooperative activity. *Philosophical Review*, 101, 327–341.
- Bråten, S. (Ed.). (1998). *Intersubjective communication and emotion in early ontogeny* (No. 3. Cambridge University Press.
- Bråten, S. (2003). Participant perception of others' acts: Virtual otherness in infants and adults. *Culture Psychology*, 9, 261–276.
- Brazelton, T. B. (2006). *Touchpoints: Birth to Three*. Da Capo Press.
- Brazelton, T. B., Koslowski, B., & Main, M. (1974). The origins of reciprocity: The early mother-infant interaction. In M. Lewis, & L. Rosenblum (Eds.), *The effect of the infant on its caregiver* (pp. 49–76). Oxford, England: Wiley.
- Brownell, C. A., & Carriger, M. S. (1990). Changes in cooperation and self-other differentiation during the second year. *Child Development*, 61(4), 1164–1174.
- Brownell, C. A., Ramani, G. B., & Zerwas, S. (2006). Becoming a social partner with peers: Cooperation and social understanding in one- and two-year-olds. *Child Development*, 77(4), 803–821.
- Bruner, J. (1985). The role of interaction formats in language acquisition. In J. P. Forgas (Ed.), *Language and social situations* (Vol. 9, pp. 31–46). New York: Springer. Springer Series in Social Psychology.
- Bruner, J. (1986). Play, thought and language. *Prospects: Quarterly Review of Education*, 16(1), 77–83.
- Bruner, J. (1990). Culture and human development: A new look. *Human development*, 33(6), 344–355.
- Bruner, J. (2004). Life as narrative. *Social research: An international quarterly*, 71(3), 691–710.
- Bruner, J. S., & Sherwood, V. (1976). Early rule structure: The case of peekaboo. *Life sentences: Aspects of the social role of language*, 55–62.
- Butterfill, S. (2012). Joint action and development. *The Philosophical Quarterly*, 62(246), 23–47. <https://doi.org/10.2307/41426867>
- Carpenter, M. (2009). Just how joint is joint action in infancy? *Topics in Cognitive Science*, 1(2), 380–392.
- Carpenter, M., & Liebal, K. (2011). Joint attention, communication, and knowing together in infancy. In A. Seemann (Ed.), *Joint Attention: New Developments in Psychology, Philosophy of Mind, and Social Neuroscience* (pp. 159–182). Cambridge, MA: MIT Press.
- Castiello, U., Becchio, C., Zoia, S., Nelini, C., Sartori, L., Blason, L., D'Ottavio, G., Bulgheroni, M., & Gallese, V. (2010). Wired to Be Social: The Ontogeny of Human Interaction. *PLoS One*, 5(10), Article e13199. <https://doi.org/10.1371/journal.pone.0013199>
- Conte, E., Grazzani, I., & Pepe, A. (2018). Social cognition, language, and prosocial behaviors: A multitrait mixed-methods study in early childhood. *Early Education and Development*, 29(6), 814–830.
- Costall, A. (2013). Things that help make us what we are. In G. Sammut, P. Daanen, & F. M. Moghaddam (Eds.), *Understanding the self and others: Explorations in intersubjectivity and interobjectivity* (pp. 66–76). Oxon: Routledge.
- de Barbaro, K. (2019). Automated sensing of daily activity: A new lens into development. *Developmental psychobiology*, 61(3), 444–464.
- de Barbaro, K., & Fausey, C. M. (2022). Ten lessons about infants' everyday experiences. *Current Directions in Psychological Science*, 31(1), 28–33.
- de Barbaro, K., Johnson, C. M., & Deák, G. O. (2013). Twelve-month “social revolution” emerges from mother-infant sensorimotor coordination: A longitudinal investigation. *Human Development*, 56(4), 223–248.
- de Barbaro, K., Johnson, C. M., Forster, D., & Deák, G. O. (2016). Sensorimotor decoupling contributes to triadic attention: A longitudinal investigation of mother-infant-object interactions. *Child development*, 87(2), 494–512.
- De Jaegher, H. (2009). Social understanding through direct perception? Yes, by interacting. *Consciousness and Cognition*, 18(2), 535–542.
- De Jaegher, H. (2019). Loving and knowing: Reflections for an engaged epistemology. *Phenomenology and the Cognitive Sciences*, 1–24.
- De Jaegher, H., & Di Paolo, E. (2007). Participatory sense-making: An enactive approach to social cognition. *Phenomenology and the Cognitive Sciences*, 6(4), 485–507.
- De Jaegher, H., Di Paolo, E. A., & Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Sciences*, 14(10), 441–447.
- De Jaegher, H., Perakyla, A., & Stevanovic, M. (2016). The co-creation of meaningful action: Bridging enaction and interactional sociology. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 371(1693). <https://doi.org/10.1098/rstb.2015.0378>
- Delafield-Butt, J. (2018). The emotional and embodied nature of human understanding: Sharing narratives of meaning. In C. Trevarthen, J. Delafield-Butt, & A. W. Dunlop (Eds.), *The Child's Curriculum: Working with the Natural Values of Young Children*, 59. Oxford: Oxford University Press.
- Delafield-Butt, J. T., & Adie, J. (2016). The embodied narrative nature of learning: Nurture in school. *Mind, Brain, and Education*, 10(2), 117–131.

- Delafield-Butt, J., & Gangopadhyay, N. (2013). Sensorimotor intentionality: The origins of intentionality in prospective agent action. *Developmental Review*, 33(4), 399–425.
- Delafield-Butt, J., & Trevarthen, C. (2015). The ontogenesis of narrative: from moving to meaning. *Frontiers in Psychology*, 6, 1157.
- Duranti, A. (2010). Husserl, intersubjectivity and anthropology. *Anthropological theory*, 10(1–2), 16–35.
- Duranti, A. (2000). *Antropologia del Linguaggio*. Roma: Meltemi Editore.
- Depperman, A. & Günthner, S. (2015). (Eds.) *Temporality in Interaction*. John Benjamins B.V.
- Deppermann, A. & Streeck, J. (2018) (Eds.), *Time in Embodied Interaction: Synchronicity and Sequentiality of Multimodal Resources*, John Benjamins Publishing Company, Amsterdam, Philadelphia.
- Fantasia, V. Some features of language socialisation in early interactions with infants. Forthcoming.
- Fantasia, V., De Jaegher, H., & Fasulo, A. (2014b). We can work it out: An enactive look at cooperation. *Frontiers in Psychology*, 5(874). <https://doi.org/10.3389/fpsyg.2014.00874>
- Fantasia, V., Fasulo, A., Costall, A., & López, B. (2014a). Changing the game: Exploring infants' participation in early play routines. *Frontiers in Psychology*, 5(522). <https://doi.org/10.3389/fpsyg.2014.00522>
- Fantasia, V., Galbusera, L., Reck, C., & Fasulo, A. (2019). Rethinking Intrusiveness: Exploring the Sequential Organization in Interactions Between Infants and Mothers. *Frontiers in Psychology*, 10, 1543. <https://doi.org/10.3389/fpsyg.2019.01543>
- Fantasia, V., Markova, G., Fasulo, A., Costall, A., & Reddy, V. (2016). Not just being lifted: Infants are sensitive to delay during a pick-up routine. *Frontiers in Psychology*, 6, 2065.
- Fantasia, V., & Nomikou, I. (2022). The intersubjective roots of pretend play. *Journal of Applied Psycholinguistics*, 22(2), 45–60.
- Fantasia, V., Oña, L. S., Wright, C., & Wertz, A. E. (2021). Learning blossoms: Caregiver-infant interactions in an outdoor garden setting. *Infant Behavior and Development*, 64, Article 101601.
- Fasulo, A. (1999). L'acquisizione della competenza conversazionale. In R. Galatolo, & G. Pallotti (Eds.), *La Conversazione. Un' introduzione allo studio dell'interazione verbale* (pp. 287–318). Milano: Raffaello Cortina Editore.
- Fasulo, A. (2015). The value of conversation analysis for child mental health. In M. O'Reilly, & J. Lester, (Eds.). *The Palgrave handbook of child mental health* (pp. 3-24). Basingstoke: Palgrave Macmillan.
- Fasulo, A., & Pontecorvo, C. (1999). *Come si dice. Linguaggio e apprendimento in famiglia e a scuola [How do you say? Language and learning in family and school]*. Rome: Carocci.
- Fawcett, C., & Liszowski, U. (2012). Observation and initiation of joint action in infants. *Child Development*, 83(2), 434–441.
- Flóres, F. S., Rodrigues, L. P., Copetti, F., Lopes, F., & Cordovil, R. (2019). Affordances for Motor Skill Development in Home, School, and Sport Environments: A Narrative Review. *Perceptual and Motor Skills*, 126(3), 366–388. <https://doi.org/10.1177/0031512519829271>
- Fogel, A., & Thelen, E. (1987). The development of early expressive and communicative action: Re-interpreting the evidence from a dynamic systems perspective. *Developmental Psychology*, 23, 747–761.
- Forrester, M. A. (2009). The emergence of self-repair. A case study of one child during the early preschool years. *Research on Language and Social Interaction*, 41(1), 99–128.
- Frank, B., & Trevarthen, C. (2012). Intuitive meaning: Supporting impulses for interpersonal life in the sociosphere of human knowledge, practice, and language. In A. Foalen, U. Lüdtke, T. P. Racine, & J. Zlatev (Eds.), *Moving Ourselves, Moving Others: Motion and emotion in intersubjectivity, consciousness, and language* (pp. 261–303). John Benjamins.
- Fuchs, T., & De Jaegher, H. (2009). Enactive intersubjectivity: Participatory sense-making and mutual incorporation. *Phenomenology and the Cognitive Sciences*, 8, 465–486.
- Gallagher, S. (2008a). Inference or interaction: Social cognition without precursors. *Philosophical Explorations*, 11, 163–174.
- Gallagher, S. (2008b). Direct perception in the intersubjective context. *Consciousness and Cognition*, 17, 535–543.
- Gallese, V. (2001). The 'Shared Manifold' Hypothesis: From mirror neurons to empathy. *Journal of Consciousness Studies*, 8(5–7), 33–50.
- Garfinkel, Harold 1988 'Evidence for locally produced naturally accountable phenomena of order', logic, reason, meaning, method, etc. in and as of the essential haecceity of immortal ordinary society'. *Sociological Theory* 6/1: 103–109.
- Gerson, S., & Woodward, A. L. (2010). Building intentional action knowledge with one's hands. *Neoconstructivism: The new science of cognitive development*, 295–313.
- Gibson, E. J. (1988). Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. *Annual Review of Psychology*, 39(1), 1–41. <https://doi.org/10.1146/annurev.ps.39.020188.000245>
- Gibson, J. J. (1979). *The Theory of Affordances*. In *The ecological approach to visual perception*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Gilbert, M. (2013). *Joint commitment: How we make the social world*. Oxford University Press.
- Goffman, E. (1983). The interaction order: American Sociological Association, 1982 presidential address. *American sociological review*, 48(1), 1–17.
- Goffman, E. (1983). Footing. In E. Goffman (Ed.), *Forms of talk* (pp. 124–159). Philadelphia: University of Pennsylvania Press.
- Goodwin, C. (2013). The co-operative, transformative organisation of human action and knowledge. *Journal of Pragmatics*, 46, 8–23. <https://doi.org/10.1016/j.pragma.2012.09.003>
- Gratier, M. (2003). Expressive timing and interactional synchrony between mothers and infants: Cultural similarities, cultural differences, and the immigration experience. *Cognitive Development*, 18, 533–554.
- Gratier, M., & Apter-Danon, G. (2009). The musicality of belonging: Repetition and variation in mother-infant vocal interaction. In S. Malloch, & C. Trevarthen (Eds.), *Communicative Musicality*. Oxford University Press.
- Gratier, M., & Trevarthen, C. (2008). Musical Narratives and Motives for Culture in Mother-Infant Vocal Interaction. *Journal of Consciousness Studies*, 15, 122–158.
- Heesen, R., Genty, E., Rossano, F., Zuberbühler, K., & Bangenter, A. (2017). Social play as joint action: A framework to study the evolution of shared intentionality as an interactional achievement. *Learning & behavior*, 45(4), 390–405.
- Heimann, M. (2022). *Imitation from Infancy Through Early Childhood*. Springer.
- Hilbrink, E., Casillas, M., Bobb, S.C., Clark, E.V., Gattis, M. & Levinson, S.C. (2014). *Turn-timing in naturalistic mother-child interactions: A longitudinal perspective*. Poster presented at the International Conference on Infant Studies (ICIS), Berlin, Germany.
- Hobson, P., & Hobson, J. A. (2008). Engaging, sharing, knowing. In J. Zlatev (Ed.), *The shared mind: Perspectives on intersubjectivity* (pp. 67–88). John Benjamins Publishing.
- Hubley, P. (1983). *The development of cooperative action in infants*. University of Edinburgh. Unpublished doctoral dissertation.
- Hubley, P., & Trevarthen, C. (1979). Sharing a task in infancy. *New Directions for Child and Adolescent Development*, 4, 57–80.
- Iverson, J. M. (2021). Developmental Variability and Developmental Cascades: Lessons From Motor and Language Development in Infancy. *Current Directions in Psychological Science*, 30(3), 228–235.
- James, W. (1890). *The principles of psychology*. Dover.
- Jaswal, V. K., & Akhtar, N. (2019). Being versus appearing socially uninterested: Challenging assumptions about social motivation in autism. *Behavioral and Brain Sciences*, 42, 82.
- Kaye, K. (1979). The maternal role in developing communication. In M. Bullowa (Ed.), *Before speech. The beginning of interpersonal communication* (pp. 191–206). Cambridge, MA: Cambridge University Press.
- Lerner, G. H., Zimmerman, D. H., & Kidwell, M. (2011). Formal structures of practical tasks: A resource for action in the social life of very young children. In J. Streeck, C. Goodwin, C. LeBaron (Eds.) *Embodied interaction: Language and body in the material world*, pp. 44–58.
- Lashley, K. S. (1954). Dynamic processes in perception. *Brain mechanisms and consciousness* (pp. 422–443).
- Lee, D. N. (2009). General Tau Theory: Evolution to date. *Perception*, 38, 837–858.
- Lester, B., & Sparrow, J. D. (2010). *Nurturing young children and families: Building on the legacy of T.B. Brazelton*. Wiley-Blackwell.
- Leudar, I., & Costall, A. (2009). *Against Theory of Mind*. New York: Palgrave Macmillan.

- Lux, V., Gredebäck, G., Non, A. L., & Krüger, M. (2022). The next step in developmental embodiment research: Integrating concepts and methods. *Frontiers in Systems Neuroscience*, 34.
- Macmurray, J. (1957). *The Self as Agent*. Faber and Faber.
- Malloch, S., & Trevarthen, C. (2009). Musicality: Communicating the vitality and interests of life. In S. Malloch, & C. Trevarthen (Eds.), *Communicative Musicality: Exploring the basis of human companionship* (pp. 1–12). Oxford University Press.
- McGowan, T., & Delafield-Butt, J. (2022). Narrative as co-regulation: A review of embodied narrative in infant development. *Infant Behavior and Development*, 68, Article 101747. <https://doi.org/10.1016/j.infbeh.2022.101747>
- McGowan, T., & Delafield-Butt, J. (2023). Neonatal participation in neonatal imitation: narrative in newborn dialogues. *Human Development*.
- Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, 198(4312), 75–78. <https://doi.org/10.1126/science.198.4312.75>
- Meltzoff, A. N., Murray, L., Simpson, E., Heimann, M., Nagy, E., Nadel, J., Pedersen, E. J., Brooks, R., Messinger, D. S., Pascalis, L. De, Subiaul, F., Paukner, A., & Ferrari, P. F. (2018). Re-examination of Oostenbroek et al. (2016): evidence for neonatal imitation of tongue protrusion. *Developmental Science*, 21(4), 1–8. Doi: 10.1111/desc.12609.
- Merker, B. (2007). Consciousness without a cerebral cortex: A challenge for neuroscience and medicine. *Behavioral and Brain Sciences*, 30(1), 63–81.
- Milward, S. J., Kita, S., & Apperly, I. A. (2014). The development of co-representation effects in a joint task: Do children represent a co-actor? *Cognition*, 132(3), 269–279.
- Myowa-Yamakoshi, M., & Takeshita, H. (2006). Do human fetuses anticipate self-oriented actions? A study by four-dimensional (4D) ultrasonography. *INFANCY*, 10(3), 289–301. https://doi.org/10.1207/s15327078in1003_5
- Negayama, K., Delafield-Butt, J. T., Momose, K., Ishijima, K., & Kawahara, N. (2021). Comparison of Japanese and Scottish Mother-Infant Intersubjectivity: Resonance of Timing, Anticipation, and Empathy During Feeding. *Frontiers in Psychology*, 12, Article 724871. <https://doi.org/10.3389/fpsyg.2021.724871>
- Negayama, K., Delafield-Butt, J. T., Momose, K., Ishijima, K., Kawahara, N., Lux, E., et al. (2015). Embodied Intersubjective Engagement in Mother-infant Tactile Communication: A cross-cultural study of Japanese and Scottish mother-infant behaviours during infant pick-up. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.00066>
- Nomikou, I., Leonardi, G., Radkowska, A., Rączaszek-Leonardi, J., & Rohlfing, K. J. (2017). Taking up an active role: Emerging participation in early mother–infant interaction during peekaboo routines. *Frontiers in psychology*, 8, 1656.
- Nomikou, I., Schilling, M., Heller, V., & Rohlfing, K. J. (2016). Language-at all times. *Interaction Studies*, 17, 120–145. <https://doi.org/10.1075/is.17.1.06nom>
- Ochs, E., & Schieffelin, B. B. (1995). The impact of language socialization on grammatical development. In P. Fletcher, & B. MacWhinney (Eds.), *The handbook of child language* (pp. 73–94). Oxford: Blackwell.
- Oostenbroek, J., Suddendorf, T., Nielsen, M., Redshaw, J., Kennedy-Costantini, S., Davis, J., et al. (2016). Comprehensive longitudinal study challenges the existence of neonatal imitation in humans. *Current Biology*, 26(10), 1334–1338. <https://doi.org/10.1016/j.cub.2016.03.047>
- Pacherie, E. (2006). Towards a dynamic theory of intentions. In S. Pockett, W. P. Banks, & S. Gallagher (Eds.), *Does consciousness cause behavior? An investigation of the nature of volition* (pp. 145–167). Cambridge, MA: MIT Press.
- Panksepp, J., & Biven, L. (2012). *The Archaeology of Mind: Neuroevolutionary Origins of Human Emotions*. Norton.
- Piaget, J. (1953). How children form mathematical concepts. *Scientific American*, 189(5), 74–79.
- Rachwani, J., Hoch, J. E., & Adolph, K. E. (2020). Action in development: Variability, flexibility, and plasticity. In J. J. Lockman, & C. S. Tamis-LeMonda (Eds.), *The Cambridge handbook of infant development: Brain, behavior, and cultural context* (pp. 469–494). Cambridge University Press.
- Rawls, A. W. (2005). Garfinkel's conception of time. *Time & Society*, 14(2–3), 163–190.
- Reed, E. S. (1996). *Encountering the World: Toward an ecological Psychology*. Oxford University Press.
- Reddy, V. (2008). *How infants know minds*. Harvard University Press.
- Reddy, V. (2015). Joining intentions in infancy. *Journal of Consciousness Studies*, 22(1-2), 24–44.
- Reddy, V., Markova, G., & Wallot, S. (2013). Anticipatory adjustments to being picked up in infancy. *PLoS one*, 8(6), e65289.
- Reddy, V., & Morris, P. (2009). Participants don't need theories: Knowing minds in engagement. In I. Leudar, & A. Costall (Eds.), *Against Theory of Mind* (pp. 91–107). London Palgrave Macmillan.
- Reddy, V., & Uithol, S. (2016). Engagement: Looking beyond the mirror to understand action understanding. *British Journal of Developmental Psychology*, 34(1), 101–114.
- Reissland, N., Francis, B., Aydin, E., Mason, J., & Schaal, B. (2014). The development of anticipation in the fetus: A longitudinal account of human fetal mouth movements in reaction to and anticipation of touch. *Developmental psychobiology*, 56(5), 955–963.
- Rogoff, B. (2003). *The cultural nature of human development*. Oxford university press.
- Ross, H. S., & Lollis, S. P. (1987). Communication within infant social games. *Developmental psychology*, 23(2), 241.
- Rossmann, N., Costall, A., Reichelt, A. F., López, B., & Reddy, V. (2014). Jointly structuring triadic spaces of meaning and action: Book sharing from 3 months on. *Frontiers in Psychology*, 5, 1390.
- Satta, E., Ferrari-Toniolo, S., Visco-Comandini, F., Caminiti, R., & Battaglia-Mayer, A. (2017). Development of motor coordination during joint action in mid-childhood. *Neuropsychologia*, 105, 111–122.
- Schegloff, E. A. (1995). Discourse as an interactional achievement III: The omnirelevance of action. *Research on language and social interaction*, 28(3), 185–211.
- Schieffelin, B. B., & Ochs, E. (1986). Language socialization. *Annual review of anthropology*, 15(1), 163–191.
- Schweikard, D. P., & Schmid, H. B. (2013). Collective intentionality.
- Searle, J. R. (1980). The intentionality of intention and action. *Cognitive science*, 4(1), 47–70.
- Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge university press.
- Sebanz, N., & Knoblich, G. (2021). Progress in joint-action research. *Current Directions in Psychological Science*, 30(2), 138–143. <https://doi.org/10.1177/0963721420984425>
- Siposova, B., Tomasello, M., & Carpenter, M. (2018). Communicative eye contact signals a commitment to cooperate for young children. *Cognition*, 179, 192–201.
- Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *Trends in cognitive sciences*, 7(8), 343–348.
- Solms, M., & Panksepp, J. (2012). The “Id” knows more than the “Ego” admits: Neuropsychoanalytic and primal consciousness perspectives on the interface between affective and cognitive neuroscience. *Brain Sciences*, 2(2), 147–175.
- Stern, D. (1971). A Microanalysis of mother–infant interaction. *Journal of the American Academy of Child Psychiatry*, 10, 501–517.
- Stern, D. (1974). The goal and structure of mother–infant play. *Journal of American Academic Child Psychiatry* 13, 402–422. doi:10.1016/S0002-7138.
- Stern, D. (1985). *The Interpersonal World of the Infant* New York Basic Books. Winne, LC and Singer, MT (1963), *Thought disorder and family relations of schizophrenics*, 1, 9–191.
- Suarez-Rivera, C., Schatz, J. L., Herzberg, O., & Tamis-LeMonda, C. S. (2022). Joint engagement in the home environment is frequent, multimodal, timely, and structured. *Infancy*, 27(2), 232–254.
- Takada, A. (2020). *The Ecology of Playful Childhood*. Springer International Publishing.
- Thelen, E. (1995). Motor development: A new synthesis. *American Psychologist*, 50(2), 79.
- Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. MIT press.
- Tollesfens, D. (2005). Let's pretend! Children and joint action. *Philosophy of the Social Sciences*, 35, 75–97. <https://doi.org/10.1177/0048393104271925>
- Tollesfens, D., & Dale, R. (2012). Naturalizing joint action: A process-based approach. *Philosophical Psychology*, 25(3), 385–407.
- Tomalski, P., López Pérez, D., Radkowska, A., & Malinowska-Korczak, A. (2022). Dyadic interactions during infant learning: Exploring infant–parent exchanges in experimental eye-tracking studies. *Infant Behavior and Development*, 69, Article 101780. <https://doi.org/10.1016/j.infbeh.2022.101780>
- Tomasello, M. (2019). *Becoming human: A theory of ontogeny*. Harvard University Press.
- Tomasello, M., & Carpenter, M. (2007). Shared intentionality. *Developmental Science*, 10, 121–125.

- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28, 675–735.
- Tomasello, M., & Moll, H. (2010). The gap is social: Human shared intentionality and culture. In *Mind the gap* (pp. 331–349). Berlin, Heidelberg: Springer.
- Trevarthen, C. (1979). Communication and cooperation in early infancy: A description of primary intersubjectivity. In M. Bullowa (Ed.), *Before speech. The beginning of interpersonal communication* (pp. 321–348). New York: Cambridge University Press.
- Trevarthen, C. (1998). The concept and foundation of infant intersubjectivity. In S. Bråten (Ed.), *Intersubjective communication and emotion in early ontogeny* (15–46), Cambridge, MA: Cambridge University Press.
- Trevarthen, C. (2016). The spiritual nature of the infant self: an imaginative actor in relations of affection. *Journal of Consciousness Studies*, 23(1–2), 258–282. <http://www.ingentaconnect.com/content/imp/jcs/2016/00000023/F0020001/art00017>.
- Trevarthen, C., & Aitken, K. J. (2001). Infant Intersubjectivity: Research, theory, and clinical applications. *Journal of Child Psychology and Psychiatry*, 42, 3–48.
- Trevarthen, C., & Delafield-Butt, J. T. (2013). Biology of shared meaning and language development: Regulating the life of narratives. In M. Legerstee, D. Haley, & M. Bornstein (Eds.), *The Infant Mind: Origins of the Social Brain* (pp. 167–199). Guildford Press.
- Trevarthen, C. & Delafield-Butt, J. T. (2015). The Infant's Creative Vitality, In Projects of Self-Discovery and Shared Meaning. In S. Robson & S. F. Quinn (Eds.), *International Handbook of Young Children's Thinking and Understanding* (pp. 3–18). Routledge.
- Trevarthen, C. & Delafield-Butt, J. T. (2017). Development of Consciousness. In B. Hopkins, E. Geangu, & S. Linkenauer (Eds.), *Cambridge Encyclopedia of Child Development* (pp. 821–835). Cambridge University Press.
- Tsabanaki, A., Kokkinaki, T., Triliva, S., & Karademas, E. (2022). Intersubjectivity in interactions between breastfeeding infants and their mothers: A longitudinal observational study in the first year of life. *European Journal of Developmental Psychology*, 1–20.
- Tuomela, R. (2005). We-intensions revisited. *Philosophical Studies*, 125(3), 327–369.
- van der Wel, R. P. R. D., & Fu, E. (2015). Entrainment and task co-representation effects for discrete and continuous action sequences. *Psychonomic Bulletin Review*, 22, 1685–1691. <https://doi.org/10.3758/s13423-015-0831-6>
- Von Hofsten, C. (2007). Action in development. *Developmental science*, 10(1), 54–60.
- von Hofsten, C. (2009). Action, the foundation for cognitive development. *Scandinavian Journal of Psychology*, 50, 617–623.
- Vincini, S., & Fantasia, V. (2022). Rich or lean? A phenomenological alternative for explaining early social cognition. *Rivista Internazionale di Filosofia e Psicologia*, 13(2), 108–124.
- Vincini, S., & Gallagher, S. (2021). Developmental phenomenology: Examples from social cognition. *Continental Philosophy Review*. <https://doi.org/10.1007/s11007-020-09510-z>
- Warneken, F., & Tomasello, M. (2007). Helping and cooperation at 14 months of age. *Infancy*, 11(30), 271–294.
- Warneken, F., Chen, F., & Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Development*, 77(3), 640–663.
- Waugh, W., Brownell, C., & Pollock, B. (2015). Early socialization of prosocial behavior: Patterns in parents' encouragement of toddlers' helping in an everyday household task. *Infant Behavior and Development*, 39, 1–10.
- Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, 24(11), 2143–2152. <https://doi.org/10.1177/0956797613488145>
- Xu, T. L., De Barbaro, K., Abney, D. H., & Cox, R. F. (2020). Finding structure in time: Visualizing and analyzing behavioral time series. *Frontiers in Psychology*, 11, 1457.
- Zoia, S., Blason, L., D'Ottavio, G., Bulgheroni, M., Pezzetta, E., Scabar, A., & Castiello, U. (2007). Evidence of early development of action planning in the human foetus: a kinematic study. *Experimental brain research*, 176, 217–226.