

Ezequiel Di Paolo¹

Participatory Object Perception

Abstract: *Social factors have so far been neglected in embodied theories of perception despite the wealth of phenomenological insights and empirical evidence indicating their importance. I examine evidence from developmental psychology and neuroscience and attempt an initial classification according to whether social factors play a contextual, enabling, or constitutive role in the ability to perceive objects in a detached manner, i.e. beyond their immediate instrumental use. While evidence of cross-cultural variations in perceptual styles and the influence of social cues on visual attention could not be said to play more than a contextual role, other factors such as the intricate developmental links between dyadic and triadic interactions in infancy, as well as episodes of peer-learning in children, play enabling roles. A common element in these factors is the presence and resolution of interpersonal conflict. Detached object perception could not develop without these social factors. I argue, in addition, that social skills such as managing partial social acts which are addressed to and completed by others, linguistic mediation, make-believe play, and the ability to control perspectival switches are constitutive — i.e. are of the essence — of the ability to see objects as*

Correspondence:

Ezequiel Di Paolo, Department of Logic and Philosophy of Science, IAS-Research Centre for Life, Mind, and Society, University of the Basque Country, San Sebastián, Spain. Email: ezequiel.dipaolo@ehu.es

¹ Ikerbasque, Basque Foundation for Science, Bilbao, Spain; Department of Logic and Philosophy of Science, IAS-Research Centre for Life, Mind, and Society, University of the Basque Country, San Sebastián, Spain; Department of Informatics, Centre for Computational Neuroscience and Robotics, and Centre for Research in Cognitive Science, University of Sussex, Brighton, UK.

present with a detached attitude. I discuss the prospects of incorporating such social elements into dynamical interpretations of the sensorimotor approach through the enactive notion of participatory sense-making.

1. Sensorimotor Solipsism

Sensorimotor approaches to perception differ from traditional computational approaches in their commitment to the central explanatory role of embodied action. But they are similar in accepting the explanatory sufficiency of the individual cognitive agent for the phenomena they purport to understand. They move beyond the brain to include the body in the world, but, like cognitivist theories of perception, sensorimotor perspectives are for the most part, to use Shaun Gallagher's apt description, *philosophically autistic* (Gallagher, 2009; 2014). The assumption in traditional cognitive science — which sensorimotor approaches do not question — is that of a transparent, literal, and accessible objective reality that contrasts with an opaque, complex, and half-hidden social world. This assumption applies pretty much across the board in the primary literature on sensorimotor approaches (e.g. Noë, 2004; O'Regan and Noë, 2001; O'Regan, Myin and Noë, 2005) as well as in subsequent discussions. It is also implicit in the supposition that one must first investigate how we perceive objects and then move on to investigate how we perceive the social world (e.g. Engel *et al.*, 2013).

The question I wish to raise is that of the social determinants of perceptual attitudes, and whether there is a way to classify, from the perspective of an embodied approach such as sensorimotor contingencies theory (or extended versions of it), the scale and variety of social influences on the sensorimotor processes that constitute human perception. Like in representational theories of perception, it may seem that the simplest starting point is a situation involving a single perceiver confronted with a single static and neutral object and that more complex cases (multiple objects, in motion, available for different concrete uses, or naturalistic scenes and social situations) must follow this prototypical case as so many degrees of increasing complexity (e.g. Marr, 1982). Everyday experience, however, as well as phenomenological insights and evidence from developmental psychology, tell a different story. I am hardly ever confronted with abstract objects, which I regard and inspect in their perceptual presence in a manner devoid of interest or social value. Most of the time, I work my

way around a world of social and biological purposes and norms, for which objects are seldom present to me as the bottles or tomatoes that populate the examples of perceptual presence in the sensorimotor literature (O'Regan and Noë, 2001, p. 945; Noë, 2004, p. 62). I don't normally inspect a bottle with my fingers to perceive in it a Gestalt shape as opposed to disjointed sensations; I pour some beer out of it. I don't regard the red tomato and consider its unseen sides and its voluminous presence; I cut it and add it to the salad. The problem of how the object maintains its identity despite its manifold of appearances is not a problem for an agent whose attitude is pragmatic, but only for one whose attitude is abstract. Once I pick up *this* hammer to drive a nail into the wall, it is clear to me that it remains the same hammer between blows, as its identity is verified by its being part of an organized sensorimotor scheme that I enact.

The issue is precisely that of the origins of the abstract perceptual attitude that seems unquestionably primordial, at least implicitly, in the sensorimotor and the traditional literatures. How do I ever develop such an attitude (exploring the shape of the bottle or regarding the tomato as a soft red volume)? Is there anything that sensorimotor theory can say about the distinction between this attitude and that of everyday involved coping? Is one a precondition (logically and/or developmentally) for the other?

The answer, Gallagher (2008; 2012) suggests, involves the social world (see also Froese and Di Paolo, 2009). One hint that points in this direction is the fact that, were it not for my co-inhabiting a world in which a myriad of perspectives, motivations, and interests flourish, I would never learn that an object has a presence beyond my own solipsistic concerns. Thus, the object would not be abstract first; a chunk of the physical world upon which each person bestows his or her meanings (for you this tomato is good for a salad, for me it is good for throwing at certain politicians). On the contrary, it is because I can *somehow* see that I am not the only one for whom the object is meaningful, that I can also appreciate what remains at the intersection of those interested perspectives — which 'cancel out' like the summation of so many randomly oriented vectors — as an abstract object.

We can formulate the same point differently. The sensorimotor approach has been criticized as insufficient for a theory of perception because it does not offer an account of the subjective, first-personal aspects of perceptual experience (Thompson, 2005) or, in other words, an account of agency (Buhrmann and Di Paolo, 2015). This problem is one that can be better approached by adopting a wider enactive

perspective and providing operational notions of sensorimotor contingencies and mastery (Buhrmann, Di Paolo and Barandiaran, 2013; Di Paolo *et al.*, 2014). According to this account, sensorimotor agency is itself constituted as a self-sustaining network of mutually equilibrated sensorimotor schemes (Buhrmann and Di Paolo, 2015). Successful acts contribute to the reassertion of this precarious network while obstacles and failures are directly evaluated by the stress they induce in the mutually equilibrating processes that make up the agent herself. This expanded theoretical base has good chances of resolving the problem of why the perceptual experience that is constituted by the enacted know-how of the laws of sensorimotor contingencies is, given the right conditions for this enactment, also the experience *of* a subjective agent, with her own perspective on the world.

But that leaves us still with a problem lying, so to speak, at the opposite end: if perception is so constituted by *my* enactment of sensorimotor schemes, *my* deployment of *my* skilful mastery of the laws of the sensorimotor contingencies that correspond to *my* body, why is it that I perceive a world that transcends my activity? Why, in other words, do I perceive objects as being out there, publicly present to me and to others?

One could attempt to answer this question from a sensorimotor perspective based on how a perceiver learns to integrate different sensorimotor schemes which are involved in her access to an object, and how, through overcoming obstacles and lacunae, and equilibrating between sensorimotor schemes, she learns about the object as an integrated whole, one to be looked at from different angles, manipulated, smelled, tasted, and utilized in various ways.

There is mileage to this notion. Its most famous version is Piaget's theory of the progressive decentring of the child's attitude towards the world (Piaget, 1936; Piaget and Inhelder, 1956). Accordingly, during the sensorimotor stage in the first two years of life, the child progresses in integrating (mutually accommodating) sensorimotor schemes belonging to the different modalities into schemes for object manipulation. Objects are increasingly treated as having enduring properties and physical permanence and as sustaining lawful spatio-temporal relations. This leads to the pre-operational stage lasting until about 7 years of age, in which language and pretence dominate interactions and the child progressively improves her ability to decentre, understand perspectives, and moral relations.

To be sure, there is room for nuances in how to interpret the social elements in the Piagetian approach, which are not exactly absent (see

e.g. Kitchener, 1981; Müller and Carpendale, 2000; Piaget, 1995; Piaget and Inhelder, 1956, van de Voort, 1983). But let us consider the individualistic version of the Piagetian framework, which is the most widespread. Such an account of decentring seems to end up in a developed perceiver who very much confronts objects in the way that has been described by proponents of the sensorimotor approach. As I see the object, my perception is also constituted by the various virtual sensorimotor possibilities that are afforded by the object, which I have learned to master in combination with my currently enacted sensorimotor scheme. I see the cup and I perceive its various features as well as its graspability apart from other pragmatic aspects. My visual scanning scheme has assimilative linkages with my grasping scheme as well as with others, and if I execute a particular one of these I do it also in a way that is coloured by the other schemes at my disposal, whether simultaneously enacted or inhibited. The whole of my sensorimotor expertise regarding cups can have an influence on how I enact a particular action such as grasping it or simply looking at it. It is this whole that constitutes my appreciation of the cup as a present, full object.

There are various problems with this individualistic account. For one, its 'adultist' view of 'undeveloped' perception in the child as incomplete or deficient (Merleau-Ponty, 2010). Before the infant reaches sufficient mutual accommodation between various ways of dealing with objects, she would seem to live in a disorganized perceptual world, whose parts are gradually put together by equilibration and association. This particular shape and these tactile sensations go together with this kinaesthetic activity and these sounds, as the infant learns to handle a toy. For Merleau-Ponty, however, the idea of disjointed sensations is never verified phenomenologically. Even if it were, the assumption that the child must bring these supposedly isolated sensations together relies ultimately on some pre-existing conception of what counts as a good organization. Meaning is imposed onto meaningless sensations by anterior schemas. In this sense, Piaget is a sophisticated intellectualist.

However, it is conceivable that one could address such deep worries by reinterpreting the process of the construction of decentred perception in the child as one in which perceptual wholes get progressively differentiated without at any point losing a meaningful relation between child and world, as opposed to being built out of pre-existent meaningless sensations. One could imagine a more trans-actional account in which progressive forms of equilibration are not

directed from the subject's side but are also shaped by encounters with an organized (and organizing) world, i.e. a revision of Piaget, which nevertheless remains largely individualistic.

This would still leave us with our problem unsolved. At the end of such an account, I, the perceiver, would access the external object in multiple ways: the confluences of several modes of engagement between my body and the object. But is this sufficient for the object to have an existence that transcends my perceptual activity? The object has a synergy of manifestations that belong to it (its surfaces change with its orientation thus, its shades of colour in this way, and if I slide my fingers on one of its faces it is smooth or rough, hard or soft). This particular synergy with my body belongs to this object and not to another one. It is a synergy of my body's modes of seizing the world, a set of organized relations in my sensorimotor schemes. So I can see two objects as different, but I cannot see either as other than as *mine*. The object, distinct from other objects and in relation to them, remains inescapably oriented towards my interests and values. This could not be otherwise since its distinction as the object that it is has been achieved by the equilibration of *my* body's own sensorimotor schemes, a process whose norms are those of my sensorimotor agency (Buhrmann and Di Paolo, 2015). This 'pragmatic angle' of the object is challenged during breakdowns, but these are breakdowns precisely to my projects. The recovery from breakdown during equilibration can only return to a new, possibly more sophisticated but nevertheless agent-centred view on the object. Objects therefore achieve only a relative independence, i.e. independence from each other, but not from my sensorimotor acts. I still inhabit a solipsistic world.²

To some extent, these problems are a consequence of what is good about the sensorimotor approach to perception, *viz.* its reliance on the situated, active agent, as constituting perceptual objects in a structured

² This would apparently still leave the possibility that I could discover the object's transcendence in the case of *unrecoverable* breakdowns, but to achieve this I should already be in possession of the abstract perceptual attitude that we are attempting to explain. Instead of simply giving up, I should be able to question the persistent breakdown and therefore enquire as to the un-assimilated properties of the object that put its practical use beyond my reach. The intellectualist position assumes that this attitude is available by some pre-given organizing principle and all that is required is to put it to the test by gathering concrete knowledge about the world. For this, direct embodied experience with the world would suffice. But if we reject intellectualism, with its pre-given principles of organization of experience, we are forced to examine the question of how the abstract perceptual attitude itself ever arises, which is what we are trying to do here.

coupling with the environment. On the enactive interpretation of this approach we cannot rely on the availability of internal representations that may be conveniently manipulated in order to extract from them hypotheses about the externality of objects, to be later empirically verified by the senses, as if it were not just an homunculus inside our heads, but a professional scientist. Representational strategies meet with their own problems, which I do not intend to discuss here. But for the enactivist, the problem, as sketched above, comes much earlier and one available option, possibly the only one, is to go social.

The problem I am exploring is an old one. Albeit described in the register and vocabulary of sensorimotor approaches to perception, phenomenologically speaking the issue of object presence is closely related to the problem of the transcendence of the world with respect to subjective experience. The argument sketched above regarding the insufficiency of a single developing subject undergoing progressive equilibration of her sensorimotor organization for constituting an abstract, objective perceptual attitude parallels Husserl's argument against the solipsistic constitution of the objective world through the variations between normal and pathological or disturbed experiences (Husserl, 1989; see also Gallagher, 2008). Like Mead (1934; 1938) who viewed the meaning and objectivity of the world as a consequence of the socially acquired expertise at adopting various roles, the Husserl of *Ideen II* sees the transcendent as constitutively intersubjective; as does Merleau-Ponty (e.g. 1964; 2010). Each of these philosophers (despite important differences) seems to appeal to a similar solution involving the transformative consequences for the subjective grasp on the world, which depend on effectively coordinating perspectives with other people.

2. Empirical Evidence

Before exploring possible enactive responses to the problem of how an abstract perceptual attitude can develop, I discuss in this section the rich empirical evidence linking social interaction practices and other social phenomena with the development, shaping, and maintenance of 'objective' and 'decentred' perceptual stances.

The evidence suggesting various kinds of influences of the social world on perceptual development is disparate. It has been obtained in attempts to answer very different questions, and it is not at all pre-ordered regarding its significance. It is also abundant. In this section, I will by no means be exhaustive nor will I attempt to organize this

evidence in any definitive manner, except for signposting some very broad categories.

I will use the terminology introduced by De Jaegher, Di Paolo and Gallagher (2010) to attempt to pre-classify some of this evidence in general terms. This will be a provisional classification since in many cases the underlying processes involved are not fully understood.

According to this classification, it is possible to ask whether social factors play a *contextual*, an *enabling*, or a *constitutive* role in phenomena in which we observe some kind of social influence on perceptual development. At the least involved end, a contextual role simply implies some sort of influence or modulation on the properties of a phenomenon. Thus, we may find that the development of object perception co-varies with some factor such as the visual richness of inhabited spaces, but that this covariation is insufficient to establish the necessity of this factor for this development to occur. In that case, such a factor plays a contextual role.

If we establish a causal necessity between a process and a particular phenomenon, then that process plays an enabling role. That is, we need not only demonstrate a covariation but also provide some explanation according to which the phenomenon could not occur without the factor in question. Enabling roles are, of course, context-dependent. Proving that a factor plays an enabling role in one particular chain of events does not imply that it will play the same or any enabling role in each possible history. But in general, and tacitly, if we postulate a factor as enabling, we understand that we mean this under a range of typical circumstances.

Finally, a factor or process is constitutive of a given phenomenon if it is established that the occurrence of the phenomenon conceptually (not just causally) implies that factor or process. Thus, for instance, sustained, powered displacement without contact with the ground is constitutive of flying, while using wings is an enabling factor (as flying without wings is at least logically conceivable). Of course, to determine this constitutive status one must have a clear and precise description of the phenomenon (if someone decided to define flying as sustained aerial displacement using *wings*, in that case wings would be constitutive and not merely enabling of flying). It may be that under a coarse first approximation the phenomenon is not yet sufficiently understood to establish whether something plays a constitutive role or not. And as conceptual categories change, constitutive relations may change too. Empirical evidence cannot by itself serve as proof of constitutive status, since the relation is conceptual. However, it can

help fine-tune categorical distinctions and refine their relations such that constitutive status may be more clearly revealed.

Let us now consider some of the evidence for social influences on the development of perceptual attitudes, moving roughly from the contextual to the constitutive categories.

2.1. Sociocultural practices

One of the broadest senses in which perception is influenced by the social world is that objects themselves, the situations they are embedded in, and the practices they involve are largely social in origin (e.g. Costall, 1995; Heft, 2007). Most infants play with toys, wear clothes, etc., which are produced according to sociocultural standards. This ubiquity simply hides from view the network of social relations involved in producing these objects, both as material entities as well as in terms of the practices, relations, and norms associated with them. But this 'hiding from view' belongs to the adult perspective from which these aspects are second nature. For the infant, the linkage between objects and the associated practices in the adult world is probably mysterious and adult guidance is required in different ways, via demonstrations or directly manipulating the infants' bodies (getting them dressed, covering them up with a blanket, etc.). This is very much also the case for 'natural' objects encountered by the infant: they are subject to socially established norms and practices (you don't put pebbles in your mouth, you don't yank out the pretty flowers, etc.).

The social origins of most objects encountered by the infant should not be seen as trivial regarding the effects on the development of infant perception and cognition. To this we should add the social origins of various sensory standards (music scales, prevalence of certain geometrical shapes and colour palettes, etc.). These topics are difficult to investigate and demand sophisticated cross-cultural studies. However, it seems true that if social influences on perception were limited to the fact that the production, standards, and practices associated with the perceived objects are social in origin, it would be conceivable that, once these relations are taken into account, the rest of the story could remain individualistic. The infant could learn on her own to understand this socially constructed world, which would only play a contextual role in influencing the development of perceptual styles.

Cultural factors have indeed been shown to co-vary with perceptual styles, even if we exclude overtly social aspects of perception such as interacting with others, perceiving signs, utterances, social norms, dealing with institutions, etc. These covariations have been found in the perception of visual illusions (Segall, Campbell and Herskovits, 1963), depth and orientation perception, and pictorial representations of space (Deregowski, 1989), among aspects including other modalities and motor factors (see Werker, Maurer and Yoshida, 2009; Adolph, Karasik and Tamis-LeMonda, 2009).

One of the most studied cultural variations in visual perceptual attitudes involves the contrast between holistic attention to context in East Asian cultures and analytic attention to parts in Westerners (Nisbett and Miyamoto, 2005). Some aspects of this contrast are studied in terms of so called field-(in)dependence. Accordingly, depending on the style of attention, the disambiguation of a figure from its background is more or less effective as figures and background vary. Holistic viewers show greater skill in allocating attention to background and relations between objects while analytic viewers excel at discriminating figures from backgrounds (e.g. Witkin and Berry, 1975). Witkin and Goodenough (1977) specifically linked these differences to interpersonal behaviours, suggesting that more field-dependence is connected with stronger interest in others, in social relations, and with emotional openness.

More recent work has verified this difference in visual style. For instance, Ji, Peng and Nisbett (2000) tested Asian and North American participants in the classical rod-and-frame task. The objective is to vertically align a rod, ignoring the tilted frame that surrounds it. North Americans can more readily ignore the frame than Asians, suggesting a different attentional style. In tests aimed at taking the background into account (not just ignoring it), Kitayama *et al.* (2003) found further evidence of this difference. In the frame-and-line test, participants are shown a square frame with a short vertical line inside it touching the top of the square. Then they are given a blank square of different size and asked to draw a vertical line either identical in absolute length to the first line shown or keeping the same relative proportion to the square. Japanese participants make fewer errors at the relative task while North Americans are better at the absolute task.

These differences in visual styles have been investigated by examining eye movements. While viewing images with a complex background and a central object, East Asians tend to fixate less, and

less quickly, on the object than North Americans and make instead more saccades into the background (Chua, Boland and Nisbett, 2005).

Interestingly, the differences in visual style are manifested from a certain age onwards, but not before. Using the frame-and-line task, Duffy *et al.* (2009) found that differences in attentional style begin to be manifested around the age of 6, while younger children do not show significant differences. Similar developmental courses are documented in the perception and production of phonemes in relation to exposure to different languages (e.g. Kuhl *et al.*, 2006). The general suggestion in these cases is to attribute the development of these differences to socialization patterns.

Of course, cultural variations in perceptual styles do not immediately imply an interpersonal influence on development. It is also possible that other factors are at play, for instance, the structure of typical visual scenes (Miyamoto, Nisbett and Masuda, 2006). Other relevant factors tend to co-vary with cultural ones (climate, diet, landscapes, etc.) and these could also have effects on perceptual development. While they are suggestive of the influence of social practices, cultural variations are not definitive proof. But we can at least claim that they play a contextual role in modulating perceptual styles. Conversely, lack of cultural variation is not by itself sufficient to discard social influences on perceptual development since these influences could play similar roles across cultures.

2.2. *Social cues*

A series of studies reveal that social cues (e.g. the gaze or facial expression of others) can influence attention, object evaluation, and affective salience. Presented with a static image of a person shifting their gaze towards one side, adult participants tend to respond faster when a visual target appears in a position that is congruent with the gaze in the image or video clip (e.g. Frischen, Bayliss and Tipper, 2007). Visual attention is drawn towards affectively salient images just by the belief that the other person in the room is looking at a similar series of images (Richardson *et al.*, 2012).

Gaze cues also influence affective judgments with participants finding objects more appealing if they are congruent with perceived gaze direction (Bayliss *et al.*, 2007; Hayes *et al.*, 2008). Similar sensitivities to social cueing are found in young infants. Eye-tracking and event related potentials measurements indicate that 4-month-old infants are sensitive to the difference between directed and averted

gaze with respect to an object (Hoehl *et al.*, 2008) and can be similarly sensitive to other cues such as head movements — the orientation of a ‘non-social’ stimulus (a toy car), however, does not elicit the same effects (Wahl *et al.*, 2013).

The pre-recorded images or clips in these studies are non-interactive social cues. By themselves they do not reveal more than a contextual involvement of social stimuli in modulating object-oriented perception. Still, the evidence of early sensitivity to these cues is congruent with other factors discussed below in ways that enabling roles may not yet be discarded.

2.3. Dyadic and triadic interactions in the first year

During the first year infants become increasingly sensitive to joint-attentional contexts as they move from dyadic interactions with the caregiver to triadic interactions involving objects (Striano and Stahl, 2005). Primary intersubjectivity, i.e. early reciprocal engagements, affective tuning, and mutual regulations, is progressively followed by co-regulated engagements involving shared attention and action on objects, or secondary intersubjectivity (Trevarthen and Hubley, 1978). Dyadic interactions are established early on and infants attempt to re-engage interactions that are paused artificially (still-face paradigm) already in the second month (Tronick *et al.*, 1978). At around 6 months objects dominate the infants’ attention while they sometimes look up at their mother. Between 9 and 12 months they engage in full triadic exploration of objects with attention being given to both the adult and the object and also focused on the relation between the two.³

Elements of joint attention such as gaze following are present as early as 3 months (Scaife and Bruner, 1975), becoming more frequent over the following 5 months (see also Rossmann *et al.*, 2014). Between 9 and 12 months, as the gaze of the adult is combined with other actions, such as grasping and manipulating the object, the infant develops an increased sensitivity not just to the object attended to but also to the relation between the adult and the object (e.g. whether an object is one that the adult has been repeatedly manipulating). Woodward (2003) suggests that the understanding of this relation may be supported by the adult’s behavioural regularities. Witnessing the

³ During this period several aspects of mother–infant engagement, such as interactional synchrony and timing, are also modulated by culture (Gratier, 2003; Keller *et al.*, 2008).

enactment of these behavioural organizations may be for the infant an indication that others also relate to the world as she does.

There are close developmental links between dyadic and triadic competences (Striano and Reid, 2006). Studies show that there is a correlation between 7-month-old infants' response to a sudden break in dyadic interaction (the still-face paradigm) and the degree of activity (attention following, joint engagement) during triadic object exploration (Striano and Rochat, 1999). These links are corroborated in longitudinal studies positively correlating dyadic sensitivities at 6 months and joint attention at 9 and 12 months (Yazbek and D'Entremont, 2006; De Schuymer *et al.*, 2011).

The complex relation between dyadic and triadic competences may be better understood by considering the dyad as a dynamical system (e.g. Hsu and Fogel, 2003). From this perspective and performing longitudinal microanalysis of interactive and sensorimotor patterns during naturalistic triadic interactions, de Barbaro, Johnson and Deàk (2013) observe robust changes in the infant's sensorimotor organization that go together with changes in the organization of social engagement. At 4 months infants tend to engage all the sensory modalities in the exploration of the object in front of them, grasping it, looking at it, taking it to the mouth. This is followed in the second half of the first year by more frequent separation of sensorimotor activities allowing decoupled engagement with the adult and other objects in synchrony with object manipulation. Mothers are active during this transition in engaging the infant's attention to the object (showing it, moving it, in conjunction with smiles and other calls to attention). They also solicit eye contact and attempt to switch the infant's attention from one object to another. These scaffolds result in the infant at 6 months being able to decouple attention to objects and parent through different modalities, holding one object in the hand while looking at another one or attending to the mother. Transitions between objects or holding two objects at once go together with a reduction in the mother's bids for attention to novel objects and more engagement in attending together to the same object, imitating, and facilitating the activity of the infant. At 12 months the infant displays bi-manual coordination in handling objects in elaborate sensorimotor schemes. Mothers at this stage perform or demonstrate more complex manipulations of objects and the infant attends to these demonstrations and alters subsequent manipulations. At this age, interaction patterns organize in turns of activity and involve partial acts such as giving or requesting.

These and other naturalistic observations (e.g. Nomikou, Rohlfing and Szufnarowska, 2013; Rossmanith *et al.*, 2014) suggest that the infant's attention and sensorimotor skills are 'educated' by context-sensitive scaffolding resulting in a socially guided mastery of attention and object manipulation (lab studies are compatible with this view, e.g. Parrinello and Ruff, 1988; Mendive *et al.*, 2013). These changes go together with elaborate affective co-regulation, activity sharing, and the organization of social interactive structures such as turn-taking, including the beginnings of reciprocal social acts. Under this interpretation, social interaction patterns, in particular the contingent bids to either diversify or focus the infant's attention in increasingly complex ways, play an enabling role in the formation of the flexible schemes involved in sophisticated object perception and manipulation.

2.4. Decentring, perspectives, and conflict

The abstract perceptual attitude, which we would like to explain in enactive terms, seems to rely on such factors as the scaffolded mastery of various, mutually accommodated sensorimotor schemes, involving different modalities. During the Piagetian sensorimotor stage and following into the pre-operational stage, such interactions with objects are marked by an increasingly decentred attitude. The child is progressively aware that objects tend to conserve their existence whether they are visible or not, that their properties are also conserved in spite of the sometimes apparently contradictory partial cues, and that object relations co-vary lawfully with perspectives which may differ from the child's current point of view. To the extent that the initial sensorimotor organizations required for these later developments already rely on the enabling intervention of parental scaffolding, as we have just seen, decentring would also be dependent on such early interactive engagements during the first year. What we find, in addition, is that social interaction as the enabler of a decentred attitude is even more prevalent as the child grows, taking the form of sophisticated awareness of others as autonomous agents with their own interests and perspectives. This awareness is often mediated by the resolution of interactive conflicts at various stages as well as conflicts internal to the child. This opens the possibility of social interactions playing not just additional enabling roles but also constitutive ones: decentring would not only be factually made possible by social practices, it would be inconceivable without them.

Decentring has long been associated with social proficiency (e.g. Feffer and Suchotliff, 1966; Chaplin and Keller, 1974). An empirical hint of this relation was provided by Martin Hughes in the 1970s who showed that children supposed to be too young to solve Piaget's three-mountain task (Piaget and Inhelder, 1956) could easily solve an equivalent task presented in terms of motivated social actors (Donaldson, 1978). Instead of presenting children with an arrangement of abstract static objects differing in size and colour and asking what this arrangement would look like from the perspective of a doll or a person sitting at a different position, as in the original task, Hughes presented them with a socially engaging situation. Using two perpendicular model walls on a table and placing a policeman doll at a certain location, children were asked to hide another doll, so that the 'policeman' could not see it. The test was complicated with the addition of one and even two more 'policemen' at different positions. While children start reliably solving the three-mountain task only at age 7 or 8, Hughes' 3- to 4-year-old participants could solve the 'policemen' problem. According to Donaldson, this is because the task makes human sense, the motives and intentions of the characters are clear as are the relations between concrete actions (escape, chase, search, hide). It may seem as if the three-mountain paradigm is too abstract, but in fact it is not only that: it also relies on a 'demotivated' notion of perspective, which the child has had very little reason and opportunity to contemplate during her history of social engagements. Not being able to adopt such non-action-based, neutral perspective by no means signifies a lack of understanding of the other as having *motivated* perspectives (orientations towards a world they share).

This understanding develops as Piaget suggests, through conflict within the organization of the child's sensorimotor schemes and its subsequent re-equilibration into better-adapted cognitive structures, but also through conflict as it emerges in interaction with adults and peers. This is something that Piaget himself proposed on various occasions. He even developed a formal account of interpersonal equilibration applied to moral values (Piaget, 1995; Kitchener, 1981).

In a series of elegant experiments in the 1970s, a team of psychologists at the University of Geneva studied the effects of peer-interactions while children attempt to solve a problem together (e.g. Doise, Mugny and Perrett-Clermont, 1975; Doise and Mugny, 1984; Mugny and Doise, 1978; Perrett-Clermont, 1980). The problems were variations of typical Piagetian tasks: conservation of length and volumes, rearranging objects to a new location conserving their

relative positions, as well as games that could be played alone or in collaboration.

The experimental protocol is similar in most of these experiments. Children are individually pre-tested on their own on the task in question and ranked as low, intermediate, and high performers. Then they are allocated in pairs to solve a similar problem together, and finally tested on their own once more. After the interaction there are reliable positive improvements in individual performance provided that certain conditions apply. Doise and colleagues observe that individual progress depends on interpersonal conflicts during the joint condition and how they are resolved. Improvements in performance do not always correlate with the observation of the correct solution to the problem, as for instance when pairing a passive low with a dominant high performer. In fact, no improvement is observed if there is a lack of participation in the lower performing child. A demonstration of a solution to a passively observing low performer does not result in robust learning while joint participation does (and in this case learning extends to other tasks, for instance interactive peer-learning of a length conservation problem extends to better performance in volume conservation problems).

Improvement also happens robustly even in the absence of a demonstration of the correct solution. When pairing low and intermediate performers a conflict may arise between two 'wrong' solutions and resolving this conflict helps both children understand aspects of the problem they could not grasp on their own. Thus, Doise and colleagues conclude that there is an optimal distance in abilities that promotes mutual learning, a zone of proximal development to put it in Vygotskian terms. Or to put it in enactive terminology, novel meanings emerge through the recovery of breakdowns in coordination during social interactions in a process of participatory sense-making (De Jaegher and Di Paolo, 2007).

The connection between these experiments and the enactive approach is also striking in terms of how enactivists conceive of social interactions. These are defined as co-regulated processes between autonomous agents whereby relational dynamical patterns acquire an autonomy of their own, without the agents losing their autonomy in the process (*ibid.*, p. 493). Doise and colleagues observe no participatory progression if one party imposes solutions and the other is excluded or must simply obey (loss of engagement, loss of individual autonomy during the interaction). They also observe that socio-cognitive conflict, when it emerges, prompts participants to repair

their interaction. When this happens the interaction is sustained and re-established on a modified ground (autonomy of the interaction).

These experiments show that there is a clear enabling influence between interactive factors (conflict, participation, demonstration, breakdowns, agreement) and progression in sensorimotor/cognitive tasks that require a decentred attitude. This influence does not correlate with the availability of correct information, but with the inter- and intra-personal organization of sensorimotor and cognitive schemes. There is also a suggestion of a possible constitutive link between aspects of social interaction — e.g. the confrontation of perspectives, the resolution of interpersonal and interactive conflicts — and the very skill of perspective taking.

It is interesting to note that while adults are much better at decentring than children, they are not infallible. They achieve a decentred solution to a problem only after having considered and inhibited an egocentric one. For instance, adults instructed to move an object in a shared frame of reference often look at the egocentric solution first (an object they can see but is occluded to the instructor) and so do children (Epley, Morewedge and Keysar, 2004). However, adults inhibit this initial reaction and proceed to reach towards the correct object while children have more difficulty doing this. This is compatible with the notion of decentring involving an active regulation of one's own perspective from the perspective of a real or putative other, rather than an acquired and readily available novel kind of personal 'objective' perspective, in which case one would not expect adult egocentric looking to occur at all. Decentring, it would seem, is an act.

2.5. Language and play

The organizing effects of language on perception are major. According to Vygotsky (1978), linguistic mediation is the key distinguishing factor between human and animal perception. Apart from providing lexical categories that structure the perceptual world (e.g. Brown and Lenneberg, 1954; Luria, 1976; Rosch *et al.*, 1976), language also enables the control of perception by permitting the voluntary shifting of attention, easier retention of past perceptual events, and action planning. These influences are too numerous to describe in detail. Even if we limit ourselves to the not overtly social aspects of perception, evidence shows that there is a deep structural influence of linguistic categories on colour perception (e.g. Thierry *et al.*, 2009),

object recognition (e.g. Boutonnet and Lupyan, 2015), motion detection (e.g. Meteyard, Bahrami and Vigliocco, 2007), and spatial understanding (e.g. Richardson *et al.*, 2003). These linguistic effects modulate early stage sensorimotor processes and they are often unconscious.

Linguistic scaffolding in the context of mother–infant play facilitates the development of flexible relations between objects (feeding the doll with a spoon later leads to feeding the teddy bear with the spoon). These flexible relations underpin object substitutions in make-believe play, which are at the basis of symbolic function (Vygotsky, 1978; Elkonin, 1978). For substitutions to work (for a pencil to become a rocket) the child adopts gestural schemas through which objects afford novel meanings (countdown to lift-off re-enactments, rocket sounds). This capacity becomes progressively abstract during the second year (Watson and Jackowitz, 1984) allowing substitutions to depend less on strict structural similarities (a toy banana for a phone) and increasingly on linguistic/gestural complementary schemas, usually in a context of social interaction. The links between symbolic play and language development are complex: around the end of the first year, language comprehension, more than language production, correlates with episodes of make-believe, while at 20 months there is a stronger correlation with semantic diversity, more than simply the amount and length of utterances (Tamis-LeMonda and Bornstein, 1994). As single make-believe episodes turn into social dramatic play (Starks, 1960; Bodrova and Leong, 1998), children learn not only to self-regulate and adopt different social roles, but also to perceive and use objects as the recipients of novel social meanings, no longer just centred on the immediate sensorimotor grasp (Sinha, 2009).

The developmental relations between language-mediated play, language comprehension and use, the development of malleable object relations, and the self-regulation of attitudes and roles are not entirely understood. But it would seem that several of these processes are of the essence for the capability to regulate perceptual attitudes. To be able to adopt a detached observational attitude towards an object, I must be able to empty it of its most immediate meaning for myself, to temporarily bracket its significance. This capability logically implies the power to manipulate meanings and relations between objects, uses, and perspectives and this suggests that some linguistic abilities play a constitutive role in abstract object perception.

3. Perceiving in the Plural

Let us evaluate, from an enactive perspective, the evidence suggesting that social processes play an enabling role in the development of a decentred perceptual attitude, as well as the evidence that helps reveal constitutive relations in what it means to enact such a skill.

According to the sensorimotor approach, perceptual experience is constituted by the enactments of sensorimotor schemes displaying mastery of the laws of sensorimotor contingencies. Enactivists interpret this proposal by avoiding the need for internal representations and relying on the dynamical processes by which sensorimotor schemes relate to each other and constitute sensorimotor agency (Di Paolo *et al.*, 2014; Buhrmann and Di Paolo, 2015). The perceiving agent is an organized system of precarious sensorimotor schemes forming a corporeal schema (Merleau-Ponty, 1964, pp. 117ff., refers to ‘conducts’ instead of schemes). These organized schemes constantly reassert their relation of mutual dependence through synergies between body and world, and synergies among themselves.

If we wish to keep these dynamical, embodied premises as our starting point, we are forced to conclude that any explanation of the emergence of decentred perceptual attitudes must be based on how the agent’s sensorimotor organization varies through processes of (meta-stable, precarious) equilibration. Dynamically, such processes rely on the emergence and subsequent resolution of conflicts: conflicts within a given scheme (adjusting the strength and precision of the grip upon realizing that the mug is a bit too full), conflicts between schemes (adjusting to driving on the opposite side of the road), and interpersonal conflicts.

The latter seem to be at the root of the most compelling evidence briefly reviewed in the previous section. We see this clearly in the experiments by Doise and colleagues, where interpersonal conflicts that afford some degree of joint participatory resolution are the ones that most influence individual progress.

Conflict is also present (sometimes implicitly) at the phases of qualitative change in mother–infant dyadic and triadic interactions. In the observations by de Barbaro and colleagues we see mothers insisting on contextually diversifying the infant’s acts and attention when these are multi-modally over-focused on a single object and later encouraging refocused engagements as the infant learns to manipulate different objects at the same time. It is as if mothers were providing a context aimed first to open up and diversify the infant’s sensorimotor

schemes, let them accommodate to each other, only to follow this with a context aimed to integrate the now diverse modes of engagement so as to make them work together. This guidance is conflictive (though this is not to be understood necessarily as antagonistic) since the mother's influence at the key stages tends to disequilibrate the infant in specific ways.

As we have mentioned, these are forms of participatory sense-making (De Jaegher and Di Paolo, 2007). According to this notion, novel meaning is jointly achieved through a history of breakdowns and recoveries in interactive coordination. From a systemic perspective, this is a way of resolving the apparent contradiction between the autonomy of the cognitive, living agent and the possibility of socially constituted, incorporated skills. Social processes can intervene, but not in a determining fashion, within the self-constituted identity of the autonomous agent (otherwise she would cease to be autonomous). However, a history of co-specifications can lead to forms of individual agency that would be unreachable through the systemic transformations available only to the isolated agent. The autonomous agent moves through developmental trajectories by herself, but through paths that are jointly constructed.

That enactive theory permits this possibility does not yet mean that we have reached a full explanation (nor will we pretend to reach it in the context of this article). First we would need to understand, from this perspective, what it could mean to say that a certain perceptual attitude is constituted intersubjectively. What aspects of such an attitude (which undoubtedly belongs to a particular, individual embodied agent) can be said to be social in essence?

The first thing to note if we try to sketch an answer is that, as was mentioned, the enactive definition of social interaction involves recognizing the autonomy of interactive patterns themselves. Interactions tend to have a life of their own and involve phenomena that are irreducible to the summation of individual actions and intentions of the participants (De Jaegher and Di Paolo, 2007; De Jaegher, Di Paolo and Gallagher, 2010). The interactive order is what brings people together (there is a pull towards sustaining interaction) but it is also a source of heteronomy for each individual participant. Individual acts during the interaction are subject to a double normative framework: they are, on the one hand, part of individual sense-making, and, on the other, they are moves in the unfolding of the interactive encounter. I may be drawn into looking at someone behind you while we are chatting, fully listening to what you're saying. But my gaze shifting

will affect how you sense my commitment to our conversation. Even if not meant interactively, my gaze shifts impact the interaction because of the double normativity at play.

Cuffari, Di Paolo and De Jaegher (2015) discuss this primordial tension between individual and interactive norms. Their dialectical model explores the notion of participatory sense-making, which moves from general coordination of intentional activity during interaction, through various forms of social agency, to languaging. In this model one form of self-differentiation in participatory sense-making is the emergence of complementary social acts, i.e. acts that require the interpersonal coordination to bring them to completion, such as the act of giving. Each individual enacts a partial act, and to learn to successfully perform social acts implies interpersonal equilibration. The ability to participate in such acts is an irreducibly social skill. Social interaction is not only enabling of this skill (via mutual equilibration) but it is also constitutive since the social act can only be conceived as being bound by the normativity of the interactive order. I may perform an elegant and dexterous gesture offering you the object you have been looking for, but if I do it with bad social timing, e.g. when you turn to address someone else, the social act is awkward or simply fails.

The partial acts that jointly make a social act tend to be reciprocal and not just mutually fitting, but also mutually *addressed*, so that a child's sensorimotor repertoire is at some point endowed with coherent partial schemes that are easy to bring into synergy with one another, but which unlike other sensorimotor schemes are inherently oriented towards an other (like one half of a handshake).⁴ So as the child learns to hand over a toy to the mother and then to take it back, she is not too far from being able to 'hand it over' to a resting place, while doing something else, and then 'accepting' it again. This is different from simply putting the toy down and then picking it up. The attitude towards the toy is tinted with social normativity involving an implied other. The toy is partially objectified by the solitary combination of two complementary partial acts enacted by a same individual. The resulting whole — though solitary — social act implies the object

⁴ See Bakhtin (1986) regarding the addressivity of spoken utterances. This property belongs more generally to gestures, partial acts, and moves that are other-oriented and structured so as to invite a response. Most partial components of social acts belong to this class. A hand extended in front of the body invites the completion of the handshake and therefore is imbued as a gesture with addressivity; the other is always-already implied in it.

is being 'held' and regarded by an absent viewpoint. In this way, the reciprocal sensorimotor skills involved in partial acts invite the shifting of perspectives, which even in solitary play can be enacted, as in the above example, by the child 'giving' the toy to an absent other whom she partially incarnates.

The social nature of object presence can also be appreciated phenomenologically. The abstract perceptual attitude involves perceiving the object with its properties and relations to the perceiver and its surroundings, but neutralized (as far as possible; we never reach a fully abstract attitude) of any instrumental interest. Successful instrumental actions are complete. A contemplative attitude towards the object is open instead. I don't know in principle where it may lead, how long it will keep me occupied, and, if the object is unfamiliar, what I am likely to discover (can this sort of contemplation *fail?*). We may ask how the child develops the elements that lead her to this open perceptual attitude. Where has she experienced such unfinished, prolonged, sensorimotor engagements? Answer: while interacting with others, in particular during play. Actions in the interactive context are hardly ever closed on themselves since they also serve in sustaining the encounter (when they fully close, they signal a pause or a break from the interaction). Playful interactions are especially open-ended, as meanings break loose from direct instrumental concerns.

So one element of the contemplative perceptual attitude is afforded by the infant's experience in expecting action not to finish with a clear goal. But this is not the only phenomenological clue. As I regard an extraneous object, my attitude is one of curiosity, an *interrogative* attitude. The present object is a reduced social other, one that is unable to fully complete my open acts. I interrogate the object with my hands, eyes, mouth, and ears, and I open myself to being surprised (or disappointed or quickly bored).⁵ As I have learned to do when participating in social encounters, I expect answers and solicitations to further action, but in this case I myself have to facilitate them.

Others have reached similar conclusions. For instance, Fuchs (2012) presents an analysis of the development of first-, second-, and third-person perspectives in infancy focusing on the emergence of reflexive, meta-perspectival stances in the child that are constitutive of social

⁵ Erwin Straus argues that the child's question 'What is this?' already implies a form of self-transcendence, a demand 'to know things as they are in themselves and not as they are in terms of a wish-fulfilling relation to himself' (Straus, 1955, p. 70).

and self understanding. He highlights the developmental relevance of conflict during interpersonal interactions. Reddy (2008) suggests that there is a primordially to second-person interactions from which notions of the self and others develop. The infant experiences in herself what it means to be an *object* of attention for others and this experience serves her to understand that others may attend also to other objects and that objects can be made into social actors.

Another way of putting it is this: in the contemplative perceptual attitude, the child approaches objects starting from a second-person perspective, one, however, that never quite reaches the full engagement of this perspective, or does so partially (personalization of toys, animism). This perspective does not have all the elements of a proper second-person attitude as the object shows itself not to be an autonomous other. This invokes complementary partial schemes to sustain the encounter in a quasi-interactive mode, as well as perspective shifting. (*‘Let’s see what happens if I remove the cover and look inside the box’, ‘It’s empty!’’, ‘But where does the rattling noise come from?’*)

The present, non-instrumental object invites and at the same time resists a second-person engagement. By eliciting in the child a second-person stance that cannot succeed, but invites completion through the adoption of multiples roles, the interrogative engagement with the object gives rise to a novel viewpoint. Out of a breakdown in the second-person attitude a third-person perspective emerges.

4. Sensorimotor Perspectives

The degree to which the social world influences human cognition and experience is a topic that tends to be shrouded in controversy. This situation does not benefit from frequent semantic slippages and conflicting vocabularies between the disciplines that could shed some light on this question. In an attempt to clear the ground, I have made use of De Jaegher, Di Paolo and Gallagher’s (2010) distinction between contextual, enabling, and constitutive roles, in order to help us understand the relevant evidence. While some forms of social influence may be declared to play a contextual role in the development of a detached perceptual attitude, evidence suggests that other kinds play enabling roles (e.g. maternal scaffolding during triadic interactions). Other factors still emerge as being precisely the kind of constitutive capacities that are at play while adopting detached attitudes and observing objects by bracketing immediate pragmatic concern.

How do these findings affect the sensorimotor account of perceptual presence? We may come back once more to Noë's (2004) explanation according to which I experience a tomato as having an unseen side and volume, because of my 'knowledge' of how sensory and motor changes would co-vary were I to go around the tomato and look at it from a different perspective. Put in these terms, this unperformed act is merely counterfactual and invites representational explanations. Gallagher (2008) rightly points out that if it were a question of imagining these non-actualized moves, I would not then perceive the hidden aspects of the object as coexisting at this moment with my view of it. I would know the hidden sides 'as a necessary consequence of a certain law of the development of my perception' (Merleau-Ponty, 1964, p. 14). But what are given to me in perception are not necessary truths, but presences (*ibid.*).

Gallagher seeks the solution to the phenomenological conundrum in open intersubjectivity. According to the Husserlian account, 'if the absent profiles cannot be correlated with my possible but non-actualized perceptions, then the absent profiles may be correlated with the possible perceptions that others could currently have' (Gallagher, 2008, p. 172). Possible truths, however, have the same inconvenience as necessary truths: they are intellectual notions that differ from the real experience of presence. Those 'possible perceptions' belonging to others may not be actualized either when I decide to stare at the tomato. And if they were, the right synergy between my body and theirs could still be absent. Failing these actualizations, possibilities must be inferred, imagined, or remembered. If there is an apperceptive horizon of possible experiences subtended by the object which includes mine and those of others, this phenomenological insight does not provide us with an explanation of how this quality of perception could come about (which is Noë's project here) in the absence of, let alone co-present alternative acts, but co-present others.

Gallagher is correct to invoke the intersubjective dimension so as to add degrees of freedom to the problem and aim for a closer phenomenological match. And as we have discussed, empirical evidence backs this move. But he is too quick in ruling out virtual non-actualized sensorimotor schemes as temporally disjointed from the perceptual act. Dynamically speaking, alternative engagements with the object can be virtually co-present if we see perception as time-extended (Buhrmann and Di Paolo, 2015; Froese and Di Paolo, 2009). But this occurs in a concrete context-dependent manner, not as conveniently invoked counterfactual acts. 'The hidden side [of an

object] is present in its own way. It is in my vicinity' (Merleau-Ponty, 1964, p. 14). Virtualities can take the form of sensitivities to adjacent situations (the risks and opportunities I confront) as well as inner tensions as certain evoked sensorimotor schemes are inhibited in favour of others.⁶

The problem is that these virtualities may or may not involve those sensorimotor covariations required to reveal specific aspects of the object, such as having an unseen side. This may often be the case with graspable objects, such as tomatoes and bottles, but not for objects that I engage differently. Do I readily see the keys of my computer keyboard in this way? Unless I am a technician, I rarely see keys outside the context of a keyboard, and I seldom access their hidden side. The virtualities present in my normal engagement with the keyboard as I write these words do in fact create a perceptual horizon that extends beyond the immediate visual and tactile sensations, but they do not suffice to constitute keys as abstractly present, as having an underside (especially — and thankfully — when I am typing).

An enactive explanation is not so much based on how I know what I know, but on how I do what I do. It is not so much about how to obtain the available information, or prove that it exists and is accessible, but about how sense-making is organized as an ongoing, world-involving activity. I do *know* keys have a hidden side. However, following the line we have been pursuing, we do not need to invoke a representational story to account for this aspect should we decide to contemplate keys abstractly, outside the context of typing. Instead, we need a story, like the one we have sketched, that explains how I am able to enact the perspectival shifts involved in adopting an abstract perceptual attitude. We have argued that the skill behind this act is intersubjectively constituted. I can see keys as objects not by mentally disassembling the keyboard, or by invoking a flat imaginary being who lives in my laptop. I do it by manipulating my sensitivities to the virtualities of my situation, i.e. by enacting the skill of perspective shifting that allows me to establish a dialogical, even playful, relation to the object, which I may now regard, no longer instru-

⁶ The latter idea is hardly new. In his essay on *The Physical Thing*, Mead (1932) proposed very much the same view in terms of the inner resistance offered by behaviours (sensorimotor schemes) that are inhibited from being enacted in a given situation. Margaret Floy Washburn (e.g. 1926) saw perception as organized in terms of action possibilities, including those that are prevented from being enacted. At no point is there any representational access or act of imagination involved.

mentally, but by invoking a different sensorimotor repertoire, for instance, the one I use when I look at small flat objects such as buttons or coins, which normally I also manipulate. Such perspectival skills come in degrees and relate to personal experience. Photographers and painters appreciate complex patterns of colour and light in ordinary objects more easily, in general, than other people.

The challenge for the enactive approach is to provide explanations of what is required for achieving the perspectival shifts that make a given environmental situation change in its meaning. Our proposal is that the skill of altering that which we are sensitive to (our field of virtualities) is constituted — not just learned — intersubjectively. A detached perceptual attitude is not so much an individual skill that I learn socially, but a social skill that I enact individually.

Acknowledgments

Thanks to Michael Wheeler, Hanne De Jaegher, and Thomas Buhrmann for comments on previous versions of this article.

References

- Adolph, K.E., Karasik, L.B. & Tamis-LeMonda, C.S. (2009) Motor skills, in Bornstein, M.H. (ed.) *Handbook of Cultural Developmental Science*, pp. 61–88, New York: Taylor & Francis.
- Bakhtin, M.M. (1986) *Speech Genres and Other Late Essays*, Austin, TX: University of Texas Press.
- Bayliss, A.P., Frischen, A., Fenske, M.J. & Tipper, S.P. (2007) Affective evaluations of objects are influenced by observed gaze direction and emotional expression, *Cognition*, **104**, pp. 644–653.
- Bodrova, E. & Leong, D.J. (1998) Development of dramatic play in young children and its effects on self-regulation: The Vygotskian approach, *Journal of Early Childhood Teacher Education*, **19** (2–2), pp. 115–124.
- Boutonnet, B. & Lupyan, G. (2015) Words jump-start vision: A label advantage in object recognition, *Journal of Neuroscience*, **35** (25), pp. 9329–9335.
- Brown, R.W. & Lenneberg, E.H. (1954) A study in language and cognition, *Journal of Abnormal Psychology*, **49**, pp. 454–462.
- Buhrmann, T., Di Paolo, E. & Barandiaran, X.E. (2013) A dynamical systems account of sensorimotor contingencies, *Frontiers in Psychology*, **4**, art. 285.
- Buhrmann, T. & Di Paolo, E. (2015) The sense of agency — a phenomenological consequence of enacting sensorimotor schemes, *Phenomenology and the Cognitive Sciences*, doi: 10.1007/s11097-015-9446-7, online first.
- Chaplin, M.V. & Keller, H.R. (1974) Decentering and social interaction, *The Journal of Genetic Psychology*, **124** (2), pp. 269–275.
- Chua, H.F., Boland, J.E. & Nisbett, R.E. (2005) Cultural variation in eye movements during scene perception, *Proceedings of the National Academy of Sciences USA*, **102** (35), pp. 12629–12633.
- Costall, A. (1995) Socializing affordances, *Theory & Psychology*, **5**, pp. 467–481.

- Cuffari, E.C., Di Paolo, E. & De Jaegher, H. (2015) From participatory sense-making to language: There and back again, *Phenomenology and the Cognitive Sciences*, pp. 1–37, [Online], <http://doi.org/10.1007/s11097-014-9404-9>.
- 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**, pp. 223–248.
- De Jaegher, H. & Di Paolo, E. (2007) Participatory sense-making: An enactive approach to social cognition, *Phenomenology and the Cognitive Sciences*, **6**, pp. 485–507.
- De Jaegher, H., Di Paolo, E.A. & Gallagher, S. (2010) Can social interaction constitute social cognition?, *Trends in Cognitive Sciences*, **14**, pp. 441–447.
- De Schuymer, L., De Groot, I., Striano, T., Stahl, D. & Roeyers, H. (2011) Dyadic and triadic skills in preterm and full term infants: A longitudinal study in the first year, *Infant Behavior & Development*, **34**, pp. 179–188.
- Deregowski, J.B. (1989) Real space and represented space: Cross-cultural perspectives, *Behavioral and Brain Sciences*, **12**, pp. 51–119.
- Di Paolo, E., Barandiaran, X.E., Beaton, M. & Buhrmann, T. (2014) Learning to perceive in the sensorimotor approach: Piaget’s theory of equilibration interpreted dynamically, *Frontiers in Human Neuroscience*, **8**, art. 551.
- Doise, W., Mugny, G. & Perrett-Clermont, A.-N. (1975) Social interaction and the development of cognitive operations, *European Journal of Social Psychology*, **5** (3), pp. 367–383.
- Doise, W. & Mugny, G. (1984) *The Social Development of Intellect*, Oxford: Pergamon.
- Donaldson, M. (1978) *Children’s Minds*, Glasgow: Fontana/Collins.
- Duffy, S., Toriyama, R., Itakura, S. & Kitayama, S. (2009) Development of cultural attention strategies in young children in North America and Japan, *Journal of Experimental Child Psychology*, **102**, pp. 351–359.
- Elkonin, D.B. (1978) *Psikhologiya igrы [Psychology of play]*, Moscow: Pedagogika.
- Engel, A.K., Maye, A., Kurthen, M. & König, P. (2013) Where’s the action? The pragmatic turn in cognitive science, *Trends in Cognitive Sciences*, **17** (5), pp. 202–209.
- Epley, N., Morewedge, C.K. & Keysar, B. (2004) Perspective taking in children and adults: Equivalent egocentrism but differential correction, *Journal of Experimental Social Psychology*, **40**, pp. 760–768.
- Feffer, M. & Suchotliff, L. (1966) Decentering implications of social interaction, *Journal of Personal & Social Psychology*, **4**, pp. 415–423.
- Frischen, A., Bayliss, A.P. & Tipper, S.P. (2007) Gaze cueing of attention: Visual attention, social cognition, and individual differences, *Psychological Bulletin*, **133**, pp. 694–724.
- Froese, T. & Di Paolo, E.A. (2009) Sociality and the life-mind continuity thesis, *Phenomenology and the Cognitive Sciences*, **8** (4), pp. 439–463.
- Fuchs, T. (2012) The phenomenology and development of social perspectives, *Phenomenology and the Cognitive Sciences*, **12**, pp. 655–683.
- Gallagher, S. (2008) Intersubjectivity in perception, *Continental Philosophy Review*, **41** (2), pp. 163–178.
- Gallagher, S. (2009) Two problems of intersubjectivity, *Journal of Consciousness Studies*, **16** (6–8), pp. 289–308.

- Gallagher, S. (2012) From the transcendental to the enactive: Comments on Varga, *Philosophy, Psychology and Psychiatry*, **19** (2), pp. 119–121.
- Gallagher, S. (2014) In the shadow of the transcendental: Social cognition in Merleau-Ponty and cognitive science, in Novotny, K., Rodrigo, P., Slatman, J. & Stoller, S. (eds.) *Corporeity and Affectivity*, Leiden: Brill.
- Gratier, M. (2003) Expressive timing and interactional synchrony between mothers and infants: Cultural similarities, cultural differences, and the immigration experience, *Cognitive Development*, **18** (4), pp. 533–554.
- Hayes, A.E., Paul, M.A., Beuger, B. & Tipper, S.P. (2008) Self produced and observed actions influence emotion: The roles of action fluency and eye gaze, *Psychological Research*, **72** (4), pp. 461–472.
- Heft, H. (2007) The social constitution of perceiver-environment reciprocity, *Ecological Psychology*, **19** (2), pp. 85–105.
- Hoehl, S., Reid, V.M., Mooney, J. & Striano, T. (2008) What are you looking at? Infants' neural processing of an adult's object-directed gaze, *Developmental Science*, **11**, pp. 10–16.
- Hsu, H.C. & Fogel, A. (2003) Stability and transitions in mother–infant face-to-face communication during the first 6 months: A microhistorical approach, *Developmental Psychology*, **39**, pp. 1061–1082.
- Husserl, E. (1989) *Ideas Pertaining to a Pure Phenomenology and a Pure Phenomenological Philosophy, Second Book*, Dordrecht: Kluwer.
- Ji, L.J., Peng, K. & Nisbett, R.E. (2000) Culture, control, and perception of relationships in the environment, *Journal of Personality & Social Psychology*, **78** (5), pp. 943–955.
- Keller, H., Otto, H., Lamm, B., Yovsi, R.D. & Kärtner, J. (2008) The timing of verbal/vocal communications between mothers and their infants: A longitudinal cross-cultural comparison, *Infant Behavior & Development*, **31**, pp. 217–226.
- Kitayama, S., Duffy, S., Kawamura, T. & Larsen, J.T. (2003) Perceiving an object and its context in two cultures: A cultural look at New Look, *Psychological Science*, **14**, pp. 201–206.
- Kitchener, R.F. (1981) Piaget's social psychology, *Journal for the Theory of Social Behavior*, **11**, pp. 255–277.
- Kuhl, P.K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S. & Iverson, P. (2006) Infants show a facilitation effect for native language phonetic perception between 6 and 12 months, *Developmental Science*, **9** (2), pp. F13–F21.
- Luria, A. (1976) *Cognitive Development: Its Cultural and Social Foundations*, Cambridge, MA: Harvard University Press.
- Marr, D. (1982) *Vision: A Computational Approach*, San Francisco, CA: Freeman & Co.
- Mead, G.H. (1932) *The Philosophy of the Present*, Chicago, IL: University of Chicago Press.
- Mead, G.H. (1934) *Mind, Self, and Society from the Standpoint of a Social Behaviorist*, Chicago, IL: University of Chicago Press.
- Mead, G.H. (1938) *The Philosophy of the Act*, Chicago, IL: University of Chicago Press.
- Mendive, S., Marc, H., Bornstein, M.B. & Sebastian, C. (2013) The role of maternal attention-directing strategies in 9-month-old infants attaining joint engagement, *Infant Behavior & Development*, **36**, pp. 115–123.
- Merleau-Ponty, M. (1964) *The Primacy of Perception*, Evanston, IL: Northwestern University Press.

- Merleau-Ponty, M. (2010) *Child Psychology and Pedagogy: The Sorbonne Lectures 1949–1952*, Welsh, T. (trans.), Evanston, IL: Northwestern University Press.
- Meteyard, L., Bahrami, B. & Vigliocco, G. (2007) Motion detection and motion verbs: Language affects low-level visual perception, *Psychological Science*, **18** (11), pp. 1007–1013.
- Miyamoto, Y., Nisbett, R. & Masuda, T. (2006) Culture and the physical environment: Holistic versus perceptual affordances, *Psychological Science*, **17**, pp. 113–119.
- Mugny, G. & Doise, W. (1978) Socio-cognitive conflict and structure of individual and collective performances, *European Journal of Social Psychology*, **8**, pp. 181–192.
- Müller, U. & Carpendale, J.I.M. (2000) The role of social interaction in Piaget's theory: Language for social cooperation and social cooperation for language, *New Ideas in Psychology*, **18**, pp. 139–156.
- Nisbett, R.E. & Miyamoto, Y. (2006) The influence of culture: Holistic vs. analytic perception, *Trends in Cognitive Sciences*, **9**, pp. 467–473.
- Noë, A. (2004) *Action in Perception*, Cambridge, MA: MIT Press.
- Nomikou, I., Rohlfing, K.J. & Szufnarowska, J. (2013) Educating attention: Recruiting, maintaining, and framing eye contact in early natural mother–infant interactions, *Interaction Studies*, **14** (2), pp. 240–267.
- O'Regan, J. & Noë, A. (2001) A sensorimotor account of vision and visual consciousness, *Behavioral and Brain Sciences*, **24**, pp. 939–1031.
- O'Regan, J., Myin, E. & Noë, A. (2005) Sensory consciousness explained (better in terms of 'corporality' and 'alerting capacity', *Phenomenology and the Cognitive Sciences*, **4**, pp. 369–387.
- Parrinello, R.M. & Ruff, H.A. (1988) The influence of adult intervention on infants' level of attention, *Child Development*, **59** (4), pp. 1125–1135.
- Perrett-Clermont, A.-N. (1980) *Social Interaction and Cognitive Development in Children*, London: Academic Press.
- Piaget, J. (1936) *La Naissance de L'intelligence Chez L'enfant*, Neuchâtel, Paris: Delachaux et Niestlé.
- Piaget, J. (1995) *Sociological Studies*, London: Routledge.
- Piaget, J. & Inhelder, B. (1956) *The Child's Conception of Space*, London: Routledge.
- Reddy, V. (2008) *How Infants Know Minds*, Cambridge, MA: Harvard University Press.
- Richardson, D.C., Spivey, M.J., Barsalou, L.W. & McRae, K. (2003) Spatial representations activated during real-time comprehension of verbs, *Cognitive Science*, **27**, pp. 767–780.
- Richardson, D.C., Street, C.N.H., Tan, J.Y.M., Kirkham, N.Z., Hoover, M.A. & Ghane Cavanaugh, A. (2012) Joint perception: Gaze and social context, *Frontiers in Human Neuroscience*, **6**, art. 194.
- Rosch, E., Mervis, C.B., Gray, W., Johnson, D. & Boyes-Braem, P. (1976) Basic objects in natural categories, *Cognitive Psychology*, **8**, pp. 382–439.
- 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**, art. 1390.
- Scaife, M. & Bruner, J.S. (1975) The capacity for joint attention in the infant, *Nature*, **253**, pp. 265–266.

- Segall, M.H., Campbell, D.T. & Herskovits, J.M. (1963) Differences in perception of geometric illusions, *Science*, **139**, pp. 769–771.
- Sinha, C. (2009) Objects in a storied world: Materiality, normativity, narrativity, *Journal of Consciousness Studies*, **16** (6–8), pp. 167–190.
- Starks, E.B. (1960) Dramatic play, *Childhood Education*, **37** (4), pp. 163–167.
- Straus, E.W. (1955) Man, a questioning being, *Tijdschrift voor Philosophie*, **17** (1), pp. 48–74.
- Striano, T. & Rochat, P. (1999) Developmental link between dyadic and triadic social competence in infancy, *British Journal of Developmental Psychology*, **17**, pp. 551–562.
- Striano, T. & Stahl, D. (2005) Sensitivity to triadic attention in early infancy, *Developmental Science*, **8**, pp. 333–343.
- Striano, T. & Reid, V.M. (2006) Social cognition in the first year, *Trends in Cognitive Sciences*, **10** (10), pp. 471–476.
- Tamis-LeMonda, C.S. & Bornstein, M.H. (1994) Specificity in mother–toddler language–play relations across the second year, *Developmental Psychology*, **30** (2), pp. 283–292.
- Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B. & Kuipers, J.R. (2009) Unconscious effects of language-specific terminology on preattentive color perception, *Proceedings of the National Academy of Sciences USA*, **106**, pp. 4567–4570.
- Thompson, E. (2005) Sensorimotor subjectivity and the enactive approach to experience, *Phenomenology and the Cognitive Sciences*, **4** (4), pp. 407–427.
- Trevarthen, C. & Hubley, P. (1978) Secondary intersubjectivity: Confidence, confiding and acts of meaning in the first year, in Lock, A. (ed.) *Action, Gesture, and Symbol*, London: Academic Press.
- Tronick, E.Z., Als, H., Adamson, L., Wise, S. & Brazelton, T. (1978) The infant’s response to entrapment between contradictory messages in face-to-face interaction, *Journal of the American Academy of Child Psychiatry*, **17**, pp. 1–13.
- Van de Voort, W. (1983) Sensorimotor egocentrism, social interaction and the development of self and gesture, in Lee, B. & Noam, G.G. (eds.) *Developmental Approaches to the Self*, pp. 143–188, New York: Springer.
- Vygotsky, L.S. (1978) *Mind and Society: The Development of Higher Mental Processes*, Cambridge, MA: Harvard University Press.
- Wahl, S., Michel, C., Pauen, P. & Hoehl, S. (2013) Head and eye movements affect object processing in 4-month-old infants more than an artificial orientation cue, *British Journal of Developmental Psychology*, **31**, pp. 212–230.
- Washburn, M.F. (1926) ‘Gestalt’ psychology and motor psychology, *American Journal of Psychology*, **37** (4), pp. 516–520.
- Watson, M.W. & Jackowitz, E.R. (1984) Agents and recipient objects in the development of early symbolic play, *Child Development*, **55** (3), pp. 1091–1097.
- Werker, J.F., Maurer, D.F. & Yoshida, K.A. (2009) Perception, in Bornstein, M.H. (ed.) *Handbook of Cultural Developmental Science*, pp. 89–125, New York: Taylor & Francis
- Witkin, M.A. & Berry, J.W. (1975) Psychological differentiation in cross-cultural perspective, *Journal of Cross-Cultural Psychology*, **6**, pp. 4–78.
- Witkin, M.A. & Goodenough, D.R. (1977) Field dependence and interpersonal behaviour, *Psychological Bulletin*, **84**, pp. 661–689.
- Woodward, A.L. (2003) Infants’ developing understanding of the link between looker and object, *Developmental Science*, **6**, pp. 297–311.

Yazbek, A. & D'Entremont, B. (2006) A longitudinal investigation of the still-face effect at 6 months and joint attention at 12 months, *British Journal of Developmental Psychology*, **24**, pp. 589–601.