Why Engagement? A second person take on social cognition

Vasudevi Reddy

There has, in recent years, been a second move in developmental psychology, to place the social at the centre of social cognition. The first occurred in the 1960s with Bruner’s invention of a LASS to constrain Chomsky’s LAD and, in the 1970s both with Donaldson’s invocation of a ‘human sense’ to modulate Piaget’s abstract logic and with the ‘discovery’ of infant sociality by Snow, Trevarthen, Tronick, Shotter and others. In broad terms, these were attempts to turn away from the individualism, internalism and a-contextuality of aspects of Piagetian and Chomskyan approaches. The current move, visible not only in developmental psychology but also in neuroscience and cognitive psychology, has a narrower focus on social cognition, but less clear targets against which it may be reacting. There are several terms in the recent literature which capture it - ‘interaction’, the ‘second-person’, ‘engagement’, ‘child-directedness’, ‘ostensive cues’, ‘shared engagement’, ‘joint engagement’, the ‘interactive brain’ and so on. However, these terms differ from each other in fundamental ways, leaving their role in social cognition unclear, unspecified and contradictory. Part of the problem lies in confusions about what is meant by each of the terms. But a deeper problem may lie in the unwitting continuation of precisely the internalism and acontextuality that the earlier move in the 1960s and 70s was attempting to displace. All current theoretical contenders in infant social cognition would claim sociality as foundational - but often very differently so; and the challenges that characterised the first move have most often fallen by the wayside.

In this paper I will attempt to tease apart some of the meanings of the terms and to argue that we do indeed need to be conscious of the principles of these challenges; that we need to see the social in social cognition not merely as supporting or permitting, but also as
fundamentally constituting it (in the developmental and genealogical, sense, Brinck, 2016). The typical development of social cognition, I will argue, originates in (and is sustained by throughout adulthood, but this is another paper!) second-person engagements which irresistibly involve the infant, changing not only the infant cogniser’s capacity to cognise, but also that which develops to be cognised.¹ The emotional involvement of persons, in particular those most salient of emotional involvements which occur in second-person engagements where the infant is directly addressed or responded to by another, becomes the crucible of cognition. It may be time to reclaim the term cognition and once again re-frame it as a phenomenon involving the person in relation.

The significance of emotional involvement is hardly debatable in our everyday lives – it provides meaning to existence, leading to better quality of life and longer mortality; it enhances school and academic performance in childhood (Furrer & Skinner, 2003); it is an explicit strategic objective in advertising (Heath, 2009); it is crucial according to some, in the therapeutic process (Maroda, 1998); it is one of the central factors driving autobiographical memory (Holland & Kensinger, 2010); and it holds, according to innumerable writings over centuries, transformative power for individual as well as group living. I use the term engagement because it captures these emotional qualities better than the term ‘interaction’, but it needs clarification. In the following sections I ask first, how we can conceptualise engagement (in terms of its structure, its contexts and its manifestation) and differentiate second and third person relations. Current explanations of social cognition, although emphasising joint engagement and direct interactions, don’t quite grasp the nettle of emotional involvement in their explanations. Lastly, I discuss evidence for the developmental significance of the You and of second person engagements using evidence of attentional and

¹ If the infant herself is changed by engagements, then what the infant needs to understand about the social is also different: not only are new potential engagements afforded, but the new aspects of social being created in the engagements arise to be understood.
intentional engagements in the first year, whose presence (as well as absence in atypical
development) can only be explained through recognising mutuality and emotional
involvement.

What is engagement?

This crucial representative image of engagement - emotional involvement - implies
that ‘mere’ interaction without acknowledgement of, or response to, the person-ness of the
other – e.g., ‘merely’ paying for petrol at a counter or showing the guard your ticket on a
train or absent-mindedly saying ‘thank you’ to someone ahead of you who holds a door open
– does not typify engagement. Interaction must contain something else: a smile, a joke,
gratitude, surprise, dislike, attraction, interest - for us to say that we are engaged. It must
involve us as persons. But this implies a categorical definition which may not be justified, and
raises questions about other assumed dichotomies.

Is engagement a continuum or a category? Is there any interaction which really does not
involve us emotionally? It could be argued that at least at some minimal level, all our
interactions must involve emotional connection; involvement of any kind is emotional. To be
aware of the attendant behind a counter looking at you expectantly, to perceive the guard’s
outstretched hand, to be aware that someone at the door is waiting till you get there requires
some degree of interest and awareness in you. And awareness implies some level of
affectivity (in Stern’s (1985) sense of vitality affects rather than categorical affects). Even if
one excludes the most basic interactions as being too minimally affective, it leaves the vast
range of interactions as a spread of varying degrees and kinds of emotional involvement.
This means of course that while the criterion of emotional involvement may be a good
representation of engagement, it can be misleading unless we see engagement as a continuum
rather than as an either/or category.
Is engagement singular or multiple? Once again, the answer to this opens inconvenient complexities. Like ‘being’, engagement not only occurs with different degrees of involvement, but also in different domains. We could be relishing the warmth of a hot shower while absorbed in solving a stubborn puzzle. We could be driving on a motorway but be reliving a painful event. We could be in the middle of a tragedy, weeping, while at the same time thinking through its practical implications as if it were a mundane logistical problem. Speaking of engagement as if it was one thing, even if at one moment, is inevitably limited. We only do so, in academic discourse, because to grasp and measure the full complexity of multiple engagements seems a Herculean task.

Does engagement occur not only with persons but also with objects? There are many reasons for answering yes and not venerating the person-object distinction. First, the distinction is by no means absolute: persons are objects too, and we run the risk of another kind of dualism by ignoring this. Second, there are many ways in which persons personalise the material world: artefacts are structured by the intelligence, the bodies and the desires of persons; and the object world is introduced to persons by a cultured and person-ed reality (Costall, 1997; Rodriguez & Moro, 2008, Rossmanith et al, 2014). Third, even new-born infants are interested in engaging with objects, with intense attention, whole body movements and rough swiping by the hands in their vicinity (von Hofsten, 1982), and are interested in the effects of their actions on the world, seeking to explore visual-kinaesthetic matching and the production of visual ‘sights’ contingent on their actions (Van der Meer et al, 1995). The emotionality involved in attraction to perceived objects and in wishing to ‘grasp’ them in some way is strong and moving evidence of an openness to engage even with the material world. No doubt engagements with objects are limited (Brazelton, et al, 1974; Legerstee, 1994), and do not involve the rich mutuality of engagement with animate beings. But they are engagements nonetheless.
Can there be engagement without action? One could regard looking at and ‘feeling’ a response to, what one sees, as action of sorts. And of course one can be engaged, intensely, with only attention and responsive feelings - as in a movie or with a puzzle (albeit with imagined actions). But doing something manifest in the world - reaching out to what one sees, or turning away or vocalising to call, can lead not only to a change of state inside oneself, but to a change in the world, allowing responses and responses to responses, expanding, curtailing or colouring the potential of the relation you are engaged in. The most impactful and consequential engagements, then, are those which involve action. Emotional involvement without action - as in intense feelings while watching a movie - is indeed engagement, but does not allow the kinds of mutuality, unscripted developments and consequences which are possible in active engagements. But action too, like personhood and emotional involvement, is inevitably a continuum.

So what is the difference between second-person and third-person relations?

It is clear that engagement can occur in many different kinds of interaction - thinking out an argument with yourself, frowning in a one-to-one encounter with another person, crying in sympathy with a character in a movie or on the news, joining in with the joyful mood of a group of friends, taking part in witty banter with a colleague while feeding off the amusement of the audience. And it is clear that the term engagement is loaded with complexities of level, of types of relation, of intensity, of activity, and of modality. Within this quagmire of overlapping terms and complex phenomena, is engagement not only still worth talking about, but worth talking about in relation to another distinction - that of second and third person relations? How would this help us understand the development of social cognition in human infants?
What could we mean by a second person and a third person relation (for the sake of simplicity talking only of relations between persons)? Although simply operationalisable in structural terms – as the difference between dyadic interactions and interactions in which one person is the third party observing another dyad interact – the substance of the distinction (and its relevance for cognition) comes from an emotional source. It derives from Martin Buber’s (1956) attempt to differentiate two different modes of knowing - the I-You and the I-It. At its heart, a second person relation involves the experience of being addressed by another, of being seen as a You by another person, and of the mutuality that is generated in seeing the other as a You in turn. A third person relation involves a more detached, observational, stance in which one sees the other as a He or She, rather than as a You.

There are different contexts in which a second or third person stance can occur: one can be in a dyadic interaction but be thinking about the interlocutor in detached analytic terms (“he has done it again” or “why does she act as if I am an idiot?” and so on); or one can be in a group or three person situation or even in a movie primarily as an observer, but still see the other(s) as if they were speaking to oneself and feel involved with responsive sympathy or hate or anger or adoration towards them. The key difference between relating in the second person and relating in the third person is not one of the structure of the situation, but one of the openness or closed-ness with which one faces (and is faced by) the other. Closed-ness to the other – even in dyadic engagements – can occur in many ways: through categorising or objectifying the other, through adopting an analytic stance, through having another agenda or concern. You can see the other through the filter of a label, a group category or a dismissive analysis (she is just a student, he is an immigrant, she is autistic); you can de-personalise the other, literally objectifying them as a means to another end (in pornography for instance, or in malicious teasing to entertain an audience); you can approach them with another agenda (I must convince him to buy this) which stops you from ‘hearing’
them or from your own genuine expressiveness; you can approach the crying of a baby with
an immediate search for expert advice (should I go to her or let her cry a little more?) which
could (momentarily) stop you hearing and relating to the crying as crying; you can be
discussing a student’s poor performance with him while worrying about what your head of
department wants you to say to him. The extent to which one is open to and present to the
other and to which the other is open to and present to you is what marks out a second person
relation. Of course this too consists of multiple levels and continua.

There are arguments against making this distinction at all (for understanding
developing social cognition). Some oppose any relevant distinction between second and
third-person interactions (what one might call, the ‘no real difference’ objection): saying that
in both cases the child or adult observes an other person (not the self) displaying actions and
expressions, making the 2nd person situation only another kind of 3rd person situation (Barresi
& Moore, 1996). This objection is premised on the assumption of a gap between interactants
which can only be bridged intellectually (hence the other is always a 3rd person, as it were)
even when in emotional engagement with you. A different objection (what one might call the
‘graded difference’ objection) comes from the observation that since the distinction between
active engagement and passive observation appears to be gradual rather than absolute, it
“undermines the claims about the developmental primacy and phenomenological
pervasiveness of 2P interactions.” (de Bruin et al 2012). But this only follows if the claims
are themselves categorical; and they need not be; the fact of their graded distinction does not
negate the fact of the different effects from different parts of the grading. Another argument
(what one could call the ‘can’t have two separate theories’ objection), primarily in the context
of understanding belief-desire reasoning (Schoenherr, 2016) questions the plausibility of the
distinction and its relevance for social cognition on two grounds. First, the boundaries
between 2nd and 3rd person contexts in the real world are seen as simply not clear-cut enough
to merit such a sharp theoretical distinction. Interactional and observational stances towards others are often so rapidly changing in a single social situation that “it would seem surprising if distinct theories were to apply to both contexts”. Second, it is argued, many of the inferences that can be drawn in second person interactive situations can also be drawn when observing others. The rapidity of shifting stances and the complexity of their intertwining is certainly true, and it is entirely sensible to assume that any theoretical and inferential grasp of minds will not differentiate between – or exist in two separate compartments for – the two types of interaction. Both these objections, however, relate to distinguishing 2nd person theories or inferences from 3rd person theories or inferences. My argument is not that 2nd and 3rd person relations build parallel theories of other minds, or indeed that they do not influence each other, but that they yield different experiences of others. And in the pre-theoretical infant the experience of mutuality and emotional connection in second-person relations is developmentally crucial (and is crucial throughout life, although probably in different ways).

Both types of experiences, second person involvements and third person observations, must influence each other and both may be necessary even for stable pre-inferential perceptions of other minds. Indeed, in phenomena such as teasing and humour, even infants seem to show a rapid alternation between stances (Reddy, 2008). But being addressed as a You and addressing the other as a You arouses emotional responses differently from watching someone else be addressed, and engenders – even if briefly – a mutuality and suspension of separateness. The other becomes a person to you, someone who knocks you off balance or enters your consciousness in a more fundamental way than when you are largely untouched by the other, or just watching them. Even though the boundaries between them may be thin and permeable and their occurrence fleeting and dynamic, the difference between the two relations – of being addressed or being ‘heard’ versus being observed or watched or analysed, of addressing or ‘hearing’ versus watching, observing or analysing – is undeniable.
We cannot talk of a single kind of ‘other’; how you experience an other depends on the type of engagement you have with them. And in early infancy the most salient access to others (and I mean this in the non-Cartesian sense of ‘others’ as persons or as beings with minds) is in engagements where the other addresses, is addressed by and responds to the infant, which involve the infant perceptually and emotionally.

A second answer to the question of whether it is worth valuing the distinction is this: it matters on the ground. However messy the definitional distinctions, being addressed and being heard makes a difference to one’s experience of being at any moment, and most crucially, it makes a difference to development. If we have evidence that such engagements matter, the phenomenon demands recognition, study and intervention where it is endangered. Many current developmental theories, however, don’t take it into account. One recent review (Schneidman & Woodward, 2015) challenges the importance of child-directedness in learning, finding little evidence for its superiority in learning words but some evidence for its importance in learning to act on cultural artefacts. The authors suggest that any advantage comes not from child-directedness per se, but from greater likelihood of attention to child-directed as compared to observed input, or from other properties accompanying it that support learning. Paradoxically, this suggestion supports a difference between 2nd person 3rd person situations: while they may not always aid learning, child-directed acts are salient to children – they obtain more attention and might be accompanied by other experiences not present in merely observational situations. When we are talking about engaging with other minds, salience, in terms of attention and emotional relevance, is central.

Evidence for the power and primacy of second person relations.
In adults, being directly addressed by another person leads to different neural processing and to enhanced sensory awareness. Hearing one’s name being called leads to activation of the same brain areas as does seeing someone look at one, independent of arousal (Kampe, Frith & Frith, 2003). Direct gaze, smiles and attracting the other’s attention show distinct and localised neural activation processes (Schilbach et al, 2006, 2008). There also appears to be a processing advantage of ‘being addressed’: even in fairly unnatural laboratory experiments where personal pronouns are presented through headphones, the second-person pronoun - “you” - is processed preferentially across different ERPs, showing up significantly earlier than the processing of first- or third-person pronouns, arousing enhanced self-related processing and reduced external processing (Herbert, Blume & Northoff, 2015). This suggests an attentional advantage - leading to greater internal sensory processing - by hearing an address. Being looked at (by eyes in a still photograph) arouses greater self-awareness: if preceded by a photograph of a face with direct gaze, participants show greater accuracy in self-reports of arousal to affectively arousing images than if they are preceded by a photograph of a face with averted gaze (Baltazar et al, 2014). If these patterns are also the case in infancy, early second person engagements might change and enhance the awareness of other and of self, allow an awareness of both self and other as persons and also allow recognition of the marking of a shared world by the other.

What is necessary for second-person relations in infancy? Three things: infants need to be open to engagement with others (i.e., have interest in and an ability to act towards them); infants need to have others who recognise them as persons (i.e., persons who address them as a You and respond to their addresses), and infants need to be able to recognise the recognition of the other and respond to it (i.e., they need to recognise the other’s actions as relevant responses connected with themselves and their actions and be able to pursue mutually responsive engagements).
Infant openness to engagement with others: There is considerable evidence of interest in others, and openness to their initiatives from birth and the first weeks and months of infancy. A specific predisposition for looking at human faces and face-like stimuli is present at birth and allows not just greater duration of looking but also tracking across space (Goren et al., 1975; Morton & Johnson, 1991). Direct gaze is not only preferred within at least a few days of birth (Farroni et al., 2002), but leads to greater accuracy of directional saccades if it precedes gaze directed elsewhere (Farroni et al., 2004). Recent findings from EEG and NIRS studies show that direct addresses result in different neurological effects even in infancy. Direct gaze results in different gamma band oscillation in 4-month-old infants (Grossman, Johnson, Farroni & Csibra, 2007) and both direct gaze and being addressed by their own name are correlated in 5-month-old infants (Grossman, Parise, & Friederici, 2010), suggesting a selective attention to direct communications. By 5 months mutual gaze leads to enhanced word learning (Parise et al., 2011) and being called by their own name to enhanced attention to objects (Parise, Friederici & Striano, 2010) and by 6 months mutual gaze leads to enhanced gaze following (Senju & Csibra, 2008). Similarly, a greater interest in sounds in the human voice range, in human voices, in female voices and in the mother’s voice, is evident at birth and may be a result of auditory experiences in utero (De Casper & Fife, 1984). Studies of neonatal imitation, regardless of the controversies surrounding their status as imitation, show powerfully that neonates within minutes of birth look with intense and focused interest at the facial and manual actions others direct towards them, and attempt to respond with actions themselves (Kougioumutzaki, 1985; Meltzoff & Moore, 1975). Such responsiveness suggests that the neonate’s openness to others may, at least in healthy and neurologically intact neonates, be free of distrust and distress (Brazelton, 1986) and is expressed in sustained and effortful attention.
Having others who recognise them as persons: The availability of ‘others’ who are keen to respond to the infant’s actions and interests may be a bit of a circumstantial lottery, and varies in form with culture, but is certainly common enough for this to form part of the normal social environment which infants meet at birth. The simplest operational measures we can think of to pin down the recognition of personhood involve direct dialogic addresses rather than treating the infant as an object (for instance by just performing necessary caretaking or medical procedures without interpersonal connection), affective attunement (by matching or complementing the other’s affect, showing recognition that the other is feeling something and trying to tune into that mood or that rhythm) and sensitivity to changing initiatives and interests of the infant (by listening to and responding to them even if they are unexpected). In early encounters, and increasingly over the next weeks, the infant is often directly addressed by others – in face to face attempts to engage the infant (Trevarthen, 1977) in responsive tactile engagements (Kaertner et al, 2010) and perhaps even in vocal responsiveness. Parents vary in their modes and degrees of affective attunement, sometimes described as variations in affective mirroring (Legerstee & Varghese, 2001), contingent smiling (McQuaid et al, 2010), and are inevitably affected by their own state of depressiveness leading to altered patterns of communication and different communicative expectancies in their infants (Field, 1984; Murray, Fiori-Cowley, Hooper & Cooper, 1996). Having one’s communicative initiatives and emotionality responded to by others not only gives the infant feedback about their relevance to the other, but potentially confirms the infant as an expressive emotional being. The influence of this feedback and confirmation, however, depends on what the infant recognises of it.

The ability to recognise others’ recognition: The recognition of the contingency of others’ responses to their own actions is evident from at least 2 months of age, but possibly also from shortly after birth. We know that infants can detect contingencies between their
own actions and their effects by 4 weeks of age (Van der Meer et al, 1985), and even in the neonatal period as revealed by rapid instrumental conditioning (Lipsitt & Siqueland, 1977), and in neonates’ ability to detect temporal mismatches between actions felt on their face and actions seen on a face on a monitor (Fillipetti et al, 2015). By two and three months of age, contingency detection is advanced and can be seen in different ways: in differentiating live versus video-taped ‘partners’ (i.e., live versus replayed versions of their own faces in a double video set up, Reddy et al, 2007) and live versus video-taped body movements (Bahrick & Watson, 1985). More than the detection of temporal contingency, however, infant distress at the lack of response to their initiatives is crucial evidence of the recognition of an absence of recognition at some level. Increased frowning, looking away, actual distress or signs of helplessness as indicated by reduced attempts to engage can result from different types of unresponsiveness: the cessation of responses within a good engagement with a familiar person (still face studies: Cohn & Tronick 1983; Markova & Legerstee, 2006; Nagy 2008), the mismatch between action and response (Murray & Trevarthen, 1985; Nadel et al, 1999), reduced contingent maternal responsiveness (Field, 1984; Murray & Cooper, 2006, Legerstee & Varghese, 2001; McQuaid et al, 2010) and so on. Anecdotal evidence from nursing care about the reduction in distress at invasive medical procedures if there is first an interpersonal address to the infant adds to the implications of these experimental studies, suggesting that the quality of contingent interactions and the extent to which they respect the communicative potential of the infant is important in the immediate and the longer term. The infant does seem to recognise when she is recognised as a conversational being from very early in life.

Part of the difficulty in arguing this case - that the infant is indeed recognising the other’s actions as recognition - comes from different meta-theoretical assumptions about the nature and availability of mind. Where the task for the infant is seen as being an inferential
construction of hidden mental states this claim becomes very difficult to sustain. But the considerable recent evidence for the embodiment and perceptual availability of many mental states (Robertson & Johnson, 2009; Becchio et al, 2010, 2013] have made the opacity of mind position rather difficult to sustain. Minds are embodied; mind is the way the body expresses attention, intention and emotion and while conceptualising mind is a complex and developmentally late achievement, perceiving it is a simple achievement within engagement (see discussions in Leudar & Costall, 2006; Reddy, 2008, Schilbach et al, 2013).

I take below two domains of social cognitive awareness - of others’ attention and others’ intentions - with two examples of second-person engagements in each domain: one showing infant responses to attention or intentions directed to the infant, and the other showing responses to attention or intentions directed to the infant’s actions. I use these to argue two points: i) that these early engagements are crucial steps along typical developmental paths in awareness of attention and intention. Not recognising their impact results, I argue, in inadequate theorising; and ii) that these engagements show a profound mutuality of emotional involvement. Both infant and other are responding to each other, in the moment, as persons and these responses open up new ways of being, and new possibilities for understanding. The engagements not only reveal infant awareness, but also create new things to be aware of, for both infant and adult.

Developments in Joint Attentional Engagements

Infants don’t just detect others’ attention towards them, they respond emotionally to it (for the moment, given patchy evidence about other modalities, I discuss only visual attention). And throughout early development the range of emotional responses to attention - positive, negative, indifferent and ambivalent - is broadly similar, even while these infant emotional responses are increasingly elicited with age by more diverse and complex
attentional acts by others. From birth, gaze to self is interesting, leading to longer durations of looks and more frequent looks (Farroni et al, 2002); and this very same interest and perception of relevance can lead to distress if the infant cannot disengage from it (Brazelton, 1986). And by the second month the onset of others’ gaze evokes smiling and positive affect (Wolff, 1986). But affective responses to attention can be more complex. Coy smiles in the two month-old suggest a recognition of others’ attention as attention (but only when directed to self), and clowning and showing-off in the 8 month-old suggest complex elicitation of attention, not just to self, but to specific actions by the self.

Positive shyness or coy smiles to gaze to self at 3 and 4 months: Coy smiles in response to others’ smiling gaze directed to them start to appear in the third month of life in typically developing infants, consisting of a pattern typically associated with embarrassed smiles in adults (Asendorpf, 1990): intense smile, with gaze, head aversion or arm raising occurring within the peak of the smile, and with frequent return of gaze. They occur more frequently after a break in interaction than later on in the interaction, and although initially occur with familiar adults, by four months of age can be seen prominently with strangers (Reddy, 2000; Colonnesi et al, 2011). Positive shyness or coy smiles do not, however, occur in children with autism despite developmental age-appropriate achievement of mirror self-recognition (Hobson et al 2006; Reddy et al, 2010). This phenomenon has two theoretical implications. First, it challenges the claim that emotional reactions akin to embarrassment cannot occur until the development of a concept of self in the second year (Lewis, 1995). The pattern of behaviour evident in these coy smiles is similar to adult expressions of embarrassed smiles (see Reddy 2005 for a discussion of similarities and differences over age).

Importantly, it is dissociated from mirror self-recognition in autism and precedes rather than follows it in typical development, suggesting that self-consciousness begins as an affective response to attention to self, and that self-conscious affectivity might help to constitute, rather
than solely be derivative of, a concept of self (Izard & Hyson, 1986; Hobson, 1995; Reddy, 2008).

Second, the emotional responses to others’ attention that begin from birth and shortly after, show continuities over the first year, long before the infant can engage in the traditionally conceived triadic joint attention (Bates et al, 1979; Tomasello, 1999). The range of emotions originally elicited only by attention directed to self, expand and start to be elicited by more complex situations and ‘stimuli’, increasing in subtlety of expression and control but without dramatic changes in categorical affect (Reddy, 2003, 2005, 2011). These continuities must be explained. They could, unconvincingly, be dismissed as mere pseudo-emotional responses (Reddy & Morris, 2004), awaiting a developmental watershed such as the so-called 9-month-revolution before being considered as responses to attention. Alternatively, they could be seen as emotional responses to attention to self which are crucial first steps in an expanding realisation of what attention is and can be.

Clowning and showing-off from 7 or 8 months of age: Soon after the middle of the first year infants start to pick up on others’ emotional responses to infant actions (Reddy, 1991, 2005, Mireault et al, 2011). Often an accidental discovery that a certain facial expression, sound or movement of the body leads to adult laughter or general positive attention can lead to the infant repeating that expression, sound or movement (sometimes for weeks) to re-elicit the response (Reddy, 1991, 2005). The occurrence of clowning is fundamentally mutual - it depends on the presence of an ‘other’ who finds an action or expression funny, and it depends on the infant’s pleasure in the other’s amusement, the infant’s awareness that the awareness is linked to the action or expression and the infant’s ability to re-elicit it by repeating the action or expression. Such emotional engagements show
deficits in pre-school children with autism (but not in developmental age-matched children with Down syndrome) with either very limited or formulaic evidence of clowning and showing-off (Reddy, Williams, & Vaughan, 2002). This phenomenon reveals an expanding grasp of attentionality (the infant now knows that others attend not only to her, but to her actions), before the infant can engage in triadic joint attention - such as pointing to external objects - towards the end of the first year. Joint engagement with the infant’s action or expression as a ‘shared object’ is evident much earlier than the joint engagement at 14 months (Moll et al, 2007); the simpler mutuality of these engagements, with adult emotional responses and infant attempts to re-elicit it interdependent on each other, may be crucial for allowing infants further access to the nature of others’ attentionality.

Both positive shyness and clowning/showing-off illustrate the powerful way in which people - both infants and adults - are emotionally moved by others’ attention and actively seek to move each other. The crucial point about these engagements is not their age of occurrence – earlier than hitherto thought – but their developmental sequencing. First, infants become emotionally moved by others’ attention to self. Some months after that they become emotionally moved by others’ attention to their actions. Only after that do they become able to grasp others’ attention to distal objects in space. This sequence is chronological (in typical development) as well as potentially causal (in children on the autistic spectrum). Given this continuity of emotional response to attention and this sequence, the argument that attention as a psychological phenomenon is ‘discovered’ at 9 months is untenable.

Developments in Joint Intentional Engagements

Infants don’t just observe others’ intentional actions towards distal objects, they also, and often powerfully from the moment of birth, experience intentional actions directed
towards them. It would be strange to imagine that infant awareness of others’ intentions was independent of infant experience of such infant-directed actions. And if there is evidence that infants not only recognise such actions, but anticipate them with appropriate responses, then it would be strange indeed to ignore these responses in theorising about action understanding.

*Anticipatory adjustments to being picked up at 2 and 3 months:* Kanner discussed at some length as long ago as 1943, reports from several parents of school-age children with autism, that their children did not show anticipatory adjustments to being picked up. Such a lack of motor preparedness or anticipation of actions towards the self has been subsequently reported in feeding and other kinds of actions in children with autism (Brisson et al., 2012). Typically developing infants, however, from at least around two months of age, show anticipatory adjustments of their bodies when familiar adults are approaching to pick them up, even before actual contact is made. Three types of body adjustment occur, differently in different infants: raising and extending or tucking up the legs, opening out or raising the arms, raising the chin or turning the head and neck. The adjustments are present at two months but become smoother by three months (Reddy, Markova & Wallot, 2013). By at least three months, these adjustments cease if the adult delays a few seconds in picking the infant up (Fantasia et al., 2016). This phenomenon does not easily fit current explanations of social cognition. Given that the infant at two months is not yet capable of reaching and grasping in this manner, their own action experience cannot be the explanation of the adjustments to the adult’s approach. The arms directed to the infant could, at a stretch, be interpreted as ostensive cues, but if so, they are communicative in themselves rather than cues to information about the world, and these adjustments appear to develop during the first weeks of life (Wallot et al., in preparation), suggesting that the arms approaching are not hard-wired signals. The key point here is that the infant at two months, before she grasps intentional actions directed to other objects, is appropriately responsive to intentional actions directed to
herself. Such engagements are often emotionally significant to the infant, and involve unthinking mutuality (sensitivity and ongoing adjustments by each person to actions as they unfold); they must form the basis of later, more complex awareness of intentions.

Compliant responses to others’ directives in the second half of first year: Although infant compliance with others’ verbal directives is clear by the end of the first year, adults issue commands and requests to infants from many months earlier. From the middle of the first year, but varying between family and culture, adults tend to increase the frequency of directives, embedding them in frequently repeated pragmatic formats (mostly involving display of positive actions and skills rather than prohibitions). And infants from around 7 or 8 months, also varying with their experience of the directives, start complying with the directives (Reddy, Liebal, Hicks, Jonnalgadda & Chintalapuri, 2013). What is very clear in these directive episodes is the infant’s interest in joining with the adult’s intentions and interests; the adult’s directive is interesting to the infant because it comes from the adult to the infant. And the infant’s (initial) willingness to engage - however tentatively - with their directives, encourages adults in constantly upping the ante and putting forward more and more complex invitations. The gradual increase in frequency of compliant responses from six and a half to twelve and a half months in both cultural groups in this study suggests that these engagements are part of a continual and expanding process of understanding intentions, challenging the claim that communicative intentions are complex things only understood in the second year of life (Tomasello, 1999). The infant-directedness of the directives – as an ‘ostensive cue’ to the purpose of the utterance - may indeed be evident before the content of the directive becomes clear as some argue (Csibra, & Gergely, 2009). But compliance with different directive contents does not occur all at once and it is likely that it is the repetitive pragmatic formats with partial and increasing fulfilment of the intent that clarify their meaning.
Intention awareness would seem to be inseparable from intentional engagements. The infant gets drawn in to respond to others’ intentions by having them be directed towards herself in the first place and then towards her actions. Infant appropriate responses to others’ intentional actions follow the same chronological scheme as with attention - going from grasping intentions directed to self, to intentions directed to actions and only later to intentions directed to objects in space. Adult sensitivity to infant capabilities and motivations is crucial in such engagements; both parties to the interactions change and adjust as they unfold, creating paths characteristic of each relationship.

Concluding points

Most theorists of social cognition would take as given, that in some sense, interaction or engagement is necessary for social cognition. But what is meant by this ‘necessity’ differs significantly among them. The work of Csibra and Gergely (2009) in studying the role of ostensive cues in the development of learning through communication has been an important step forward in the field and is sometimes identified as an example of a second-person theory. They have argued that the direct address - looking at the infant, calling their name or using infant-directed speech - is crucial in activating built-in attentional biases in the infant and in allowing the infant to see what follows as generalizable knowledge. However, their use of the term ‘ostension’, deriving from their central focus on understanding the occurrence of communicative learning about the world (not especially about the other), takes a different slant from the second-person position outlined here. Their use of a cue-based explanation is substantially different from the emotional-involvement based explanation I offer here. The evidence (of the primacy of infant awareness of attention and intentions directed to the infant, and of the expansion of this awareness to attention and intentions directed to the infant’s actions, and only then to awareness of attention and intentions directed to objects in space) extends their findings to the early emergence of other-awareness (not just about others as
cues leading to the world awareness). Similarly, the work of Michael Tomasello and his colleagues has been central to almost all debates about the emergence of social cognition and is also sometimes identified as involving a second person position. Their recent focus on joint engagement - e.g., showing that understanding what another person knows about an object depends initially on actual joint engagement with that adult and the object, and only subsequently relies on simple observation of the adult’s engagement with the object (Moll, Carpenter & Tomasello, 2007) - suggests an intriguingly crucial role for engagement in social cognition. However, even though the implications of the position have very early roots in dyadic interaction, the evidence used is largely limited to triadic interactions occurring after 9 months and the so-called 9 month socio-cognitive revolution in understanding of intentionality (Tomasello, 1999; Tomasello et al 2005). The recognition of early second person engagements allows a fuller recognition of what it means to be addressed, to being seen as a You by another person (and not only to seeing the other as a You with an object). The mutuality of seeing the other as a You and being seen as a You at the same time needs us to go a step further - or earlier - for understanding the second-personal roots of social cognition.

To sum up, to offer an adequate explanation of social cognition we need explanations which do not cast the infant in the role of an epistemic observer or analyst, which do not begin to explain social cognition until late in infancy, and which do not conceptualise engagement merely in terms of cues to the perception of the object world. To understand the origins of social cognition, we need to understand direct infant engagements with the social – their occurrence in early addresses by the world, the reasons for their occurrence and the manner in which they seem to be sustained in a delicate mutuality of emotional involvement. That which is mutual not only involves a more messy developmental experience, but also
involves a changing field; we seem to be poised at the beginning of an explanation but have not got very far.

Some might argue that the proper domain of study to explain social cognition is the individual brain - that is, that “the social neuroscientist does not also need to be studying the stock market” (Adolphs in De Jaegher et al, 2015). However, the presence of the stock market must change not only the content, but the form of the cognitions. The shape and potential of cognition is not fixed independent of the material it cognises. The material itself changes the shape and scope of cognition. Engagements are not just grist for a cognitive mill; they create the very material that we seek to understand.

Like all those other terms that we use - mind, culture, emotion, love - engagement is a vague and multi-faceted term. One could argue that this is an essential vagueness – an indeterminacy that encompasses the possibility of as yet unknown manifestations. But it captures something crucial about our social lives. The ability to connect with the emotions of others and to recognise their recognition of our own is at the heart of development and may involve a dynamic process of ‘identifying with’ (Hobson, 2002, 2007) and an openness to dialogue where the emergence cannot be pre-determined. Similarly, in attempting to study socio-cognitive development from birth, the scientist’s recognition of the subjective orientations of infant and adult in engagement are crucial for understanding what it is that prompts infants to try and understand others, and what it is that adults are doing to invite and indeed allow it. As much as mutual engagement between infants and adults is needed for developing social cognition, mutual engagement between developmental scientists and infants is needed to develop an adequate theory of social cognition.
References


Brinck, I. (2016). Personal communication


Herbert, C., Blume, C., & Northoff, G. (2015) Can we distinguish an “I” and “ME” during listening?—an event-related EEG study on the processing of first and second person personal and possessive pronouns, Self and Identity, DOI: 10.1080/15298868.2015.1085893


Johnson & Morton, 19??


