

# **The parent-infant dyad and the construction of the subjective self**

Peter Fonagy, PhD, FBA  
Freud Memorial Professor of Psychoanalysis, University College London  
Chief Executive, The Anna Freud Centre, London

George Gergely, PhD  
Institute for Psychological Research, Budapest, Hungarian Academy of Sciences &  
The Anna Freud Centre, London

Mary Target, PhD  
Reader in Psychoanalysis, University College London  
Professional Director, The Anna Freud Centre, London

Address for correspondence:  
Sub-Department of Clinical Health Psychology  
University College London  
Gower Street  
London WC1E 6BT  
E-mail: p.fonagy@ucl.ac.uk  
Tel: 44 20 7679 1791  
Fax: 44 20 7916 8502

Revised version submitted to Journal of Child Psychology and Psychiatry, November 2006.

## **Abstract**

Developmental psychology and psychopathology has in the past been more concerned with the quality of self-representation than with the development of the subjective agency which underpins our experience of feeling, thought and action, a key function of mentalization. This review begins by contrasting a Cartesian view of pre-wired introspective subjectivity with a constructionist model based on the assumption of an innate contingency detector which orients the infant towards aspects of the social world that react congruently and in a specifically cued informative manner that expresses and facilitates the assimilation of cultural knowledge. Research on the neural mechanisms associated with mentalization and social influences on its development are reviewed. It is suggested that the infant focuses on the attachment figure as a source of reliable information about the world. The construction of the sense of a subjective self is then an aspect of acquiring knowledge about the world through the caregiver's pedagogical communicative displays which in this context focuses on the child's thoughts and feelings. We argue that a number of possible mechanisms, including complementary activation of attachment and mentalization, the disruptive effect of maltreatment on parent-child communication, the biobehavioural overlap of cues for learning and cues for attachment, may have a role in ensuring that the quality of relationship with the caregiver influences the development of the child's experience of thoughts and feelings.

**Keywords:** attachment, mentalization, theory of mind, social development, social cognition, self, pedagogical stance

**Acknowledgements:** The authors wish to acknowledge the creative and constructive contribution of Dr. Elizabeth Allison and the exceptionally helpful comments of an anonymous reviewer and Dr. James Leckman, MD.



# The parent-infant dyad and the construction of the subjective self

## Introduction

The ability to give subjective meaning to psychological experiences becomes possible as a result of our developing ability for explicit and reflective understanding that others' (as well as our own) actions are driven by underlying mental states and the establishment of adaptive mentalizing strategies to reason about interactive experiences in terms of such mental states. This review aims to examine evidence and theory that pertains to the relevance of the parent-infant relationship for the emergence of mentalizing. We shall explore if the establishment of the representational and attentional preconditions for such a reflective mentalizing capacity develops optimally in a relatively safe and secure social context and if so, how we might understand this. We will commence our review with considering models that potentially entail a Cartesian view of the nature of subjectivity and overview evidence concerning brain structures known to be recruited by mentalisation. We will consider, on the basis of evolutionary speculation and recent neuroimaging data, why we might consider mentalization and the social context provided by parent-infant relations to be linked. We will also consider in some detail if the literature on the social influences on mentalisation might give us ground for assuming that the parent-child relationship contributes to the 'construction' of the psychological self. Finally, we will consider the implications of a recently advanced model for the intergenerational transfer of cultural knowledge, pedagogy theory, for the unfolding of social cognitive competences.

In the 1980s developmental psychology began to investigate when we become able to understand that people can have false beliefs about the world (Perner & Lang,

2000; Wellman, 1990; Wellman & Liu, 2004). A number of researchers consider the resulting construct of theory of mind and its false belief paradigm to be too narrow (Carpendale & Lewis, 2006) as it fails to encapsulate the relational and affect regulative aspects of interpreting behaviour in mental state terms. Developmentalists have also started to use the term ‘mentalizing’ as an alternative, because it is not limited either to specific tasks or particular age groups (Morton & Frith, 1995; O'Connor & Hirsch, 1999).

We define mentalization following a tradition in philosophy of mind established by Brentano (1973/1874), Dennett (1978) and others as a form of mostly preconscious imaginative mental activity, namely, perceiving and interpreting human behaviour in terms of intentional mental states (e.g. needs, desires, feelings, beliefs, goals, and reasons). It is imaginative because we have to imagine what other people might be thinking or feeling – an important indicator of high quality of mentalisation is the awareness that we cannot know what is in someone else’s mind (for a discussion of the definition of the concept see Allen, 2006). We would even suggest that a similar kind of imaginative leap is required to understand one’s own mental experience, particularly in relation to emotionally charged issues and certainly some neural networks subserving judgments of intentionality in self and other appear to overlap (den Ouden, Frith, Frith, & Blakemore, 2005; Frith & Frith, 2003). In order to be able to adopt this stance (consciously or unconsciously), to have and conceive of others as having a “mind”, the individual needs a symbolic representational system for mental states and also needs to be able to selectively activate states of mind in line with particular intentions (attentional control, Leslie, 2000).

Thus mentalisation entails at least three key overlapping functions: (1) an intuitive ‘theory’ of action that we might term ‘mentalism’ that compels us to

interpret (human) actions as caused by intentional mental states (beliefs, desires, wishes); (2) a representation of others' minds that enables humans to infer, attribute and represent the intentional mental states of others – a capacity that can clearly extend to generate representations of one's own mind; (3) a capacity to predict, explain, and justify the actions of others by inferring the intentional mental states that cause them. If we are to predict and justify each others' actions we have to understand that we have separate minds that (often) contain different mental models of reality but that it is this internal reality rather than the external one that causes our actions. To do this we have to be able to infer and represent both the mental models of the other's mind and the mental models of our own mind.

If they are to achieve this children need to acquire a complex set of cognitive capacities: (1) to represent causal mental states of others with counterfactual contents (false beliefs), (2) to represent causal mental states of others with fictional contents (pretense, imagination, fantasy), (3) to simultaneously represent and differentiate between the mental models of the self and of the other about reality, (4) to infer and attribute the mental states of others from visible behavioural cues as mind states are invisible, and we have to rely on cues such as gaze-direction, emotion expressions, gestures, verbal and non-verbal communicative signals, non-communicative behavioural correlates and signs, (5) to detect our own perceptible (behavioural, physiological, emotional, arousal, etc.) cues in order to infer, interpret, and attribute mental states to our self. This is in our view a substantive question as we shall try to show that the causal mental states of the Self - contrary to Cartesian doctrine - are also invisible to introspection. To put it simply: the mind of the self is not transparent to itself.

## **The Cartesian view of the nature of the subjective sense of self**

It is a commonly expressed reproach (e.g. Dennett, 1991) that the question of the developmental and social-environmental origins of our subjective sense of affective states has all too often been answered using the Cartesian assumption of a universal, shared subjectivity across individuals and through development. This Cartesian view assumes an innate, prewired organization of our mind that ensures 'primary introspective access' to our internal mental states providing us with 'first person authority' over the contents of our private subjective mental life (for a critical discussion of this general view, see Carpendale & Lewis, 2006; Fonagy, Gergely, Jurist, & Target, 2002; Gergely, 2002; Gopnik, 1993; Wegner & Wheatley, 1999).

### ***Simulation and the mirror neuron system***

The Cartesian approach is often coupled (in so-called 'simulationist' models of mind-reading, e.g. Gallese & Goldman, 1998; 2004; Goldman, 1993; Goldman & Sripada, 2005; Gordon, 1995; Harris, 1991; 1992) with the idea that the way we come to understand (or, in a sense, to internally directly 'perceive') *other* people's subjective mental states is by (automatically) 'putting ourselves in their shoes' using (in our imagination) our self as a mental model of the other (for a fuller exposition see Saxe, Carey, & Kanwisher, 2004). Through this process of internally 'simulating' the other person's goals and particular situation one comes to infer and represent the other's mental states as well as anticipating the actions these intentional mind states are likely to cause. This involves mentally inducing the internal subjective states of the other in ourselves by imitation, imagination, identification, or lately, through 'neuronal resonance' evoked by the automatic activation of our brain's 'mirror neuron system' during the observation of the other person's behavior (Gallese et al., 2004).

Recent work on the mirror neurone system (Gallese et al., 2004; Rizzolatti & Craighero, 2004) suggests that the fundamental mechanism that allows us to understand the actions and emotions of others involves the activation of the mirror neurone system for actions and the activation of visceromotor centres for the understanding of affect. The claim is made on the basis of the observation that the motor neurones, originally found in the ventral premotor cortex of the macaque monkey respond both when the monkey performs a particular goal-directed act and when it observes another individual performing a similar action (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996). Action observation causes the automatic activation of the same neural mechanism triggered by action execution or even by the sound produced by the same action (Kohler et al., 2002). There is evidence that the mirror neuron system, both in monkeys (Ferrari, Gallese, Rizzolatti, & Fogassi, 2003) and humans (Rizzolatti & Craighero, 2004), also encompasses communicative actions. In an fMRI study, participants observed communicative mouth actions in humans, monkeys, and dogs which led to the activation of different cortical foci with actions belonging to the motor repertoire of the observer's species (e.g. biting and speech reading) being mapped on the observer's motor system (Buccino et al., 2004). Since the discovery of mirror neurons, a number of similar experiments (Calmels et al., 2006; Gazzola, Aziz-Zadeh, & Keysers, 2006; Lotze et al., 2006; Molnar-Szakacs, Kaplan, Greenfield, & Iacoboni, 2006) as well as indirectly connected studies for example on facial mimicry (Sato & Yoshikawa, 2006), gender differences (Cheng, Tzeng, Decety, Imada, & Hsieh, 2006), and autism (Dapretto et al., 2006; Williams, Waiter et al., 2006) have been interpreted as implying that we understand the actions, emotions and sensations of others from the perspective of sharing their actions (Keysers & Gazzola, 2006; Rizzolatti, Ferrari, Rozzi, & Fogassi, 2006). It is

suggested that a single mechanism (shared circuits) applies to witnessing the actions, sensations and emotions of other individuals and to performing the same actions. Similarly, feeling the same sensations and emotions and translating the vision and sound of what other people do and feel into the language of the observer's own actions and feelings provides intuitive insights into their inner life. The thesis of embodied semantics holds that conceptual representations accessed during linguistic processing are, in part, equivalent to the sensory-motor representations required for the enactment of the concepts described (Aziz-Zadeh, Wilson, Rizzolatti, & Iacoboni, 2006).

This suggests a dichotomy between an immediate direct, motor-mediated type of action understanding, and a more cognitive type based on the interpretation of visual representations. This is thought to be also true for emotion understanding and we might conceive of a two-level system underpinning mentalisation with a (frontal) cortical system that invokes declarative representations and a mirror neurone system sub-serving a more immediate direct understanding of the other. In the anterior insula, visual information concerning the emotions of others is directly mapped onto the same visceromotor neural structures that determine the experience of that emotion in the observer (Wicker et al., 2003). This direct mapping can occur even when the emotion of others can only be imagined (Singer et al., 2004) or inferred from visual stimuli (Jackson, Meltzoff, & Decety, 2005). Gallese, Goldman and others hypothesize a shared sub-personal neural mapping between what is acted and what is perceived that can be used to predict the actions of others (Gallese, 2003, 2006; Goldman & Sripada, 2005; Saarela et al., 2006). This automatically established link between agent and observer may not be the only way to understand the emotions of others, but the simulation of actions by means of the activation of parietal and

premotor cortical networks may constitute a basic level of experiential understanding that does not entail the explicit use of any theory or declarative representation.

Once such a mental model has been set up, all one has to do is to introspectively access its contents and ‘read off’ from this ‘off-line self-simulation of the other’ what the other must be feeling, intending, or believing in the given situation. In other words, by accessing the thoughts and feelings that one would have in the other’s - internally represented - situation, one can attribute (by analogy) these simulated subjective states to the other person’s mind. The central assumption of this simulationist account of understanding other minds is that the basic set of subjective mental states of different individuals are identical and ‘interchangeable’ and that similar situations generate the same causal mental states and consequent action-tendencies in all of us. However, it has been pointed out that the models do not take full account of the computational burdens on the system that they clearly imply (Oztop, Kawato, & Arbib, 2006).

The direct matching account of understanding others’ actions in terms of goals and intentions by mapping them directly onto one’s corresponding motor actions through the mirror neuron system has been criticized on a number of other grounds as well. Csibra (in press) reviewed evidence showing that brain areas that are not part of the mirror neuron system (and have no motor properties, such as the superior temporal sulcus - STS) are routinely activated during action observation and seem to play a crucial role in assigning goals to actions. In this view, the premotor action representations of the mirror neuron system are activated in a top-down fashion by such previously assigned goal representations (rather than through ‘direct matching’) and play a predictive (rather than a recognitive) role by anticipating (and monitoring) the other’s action to achieve the goal through simulation. There is developmental

evidence from human infants showing that infants as young as 6 months of age can understand and anticipate goal-directed actions of others even when they are performed by unfamiliar, inanimate, or abstract (animated) agents (e.g Csibra, Gergely, Bíró, Koós, & Brockbank, 1999; Kamerawi, Kato, Kanda, Ishiguro, & Hiraki, 2005; Luo & Baillargeon, 2005; Wagner & Carey, 2005) or by computer-generated simulations of human hands performing biomechanically impossible actions (that, nevertheless, involve an efficient goal approach, see Gergely & Csibra, 2003).

These findings cannot be easily accommodated by the mirror neuron account as such observed actions cannot be directly mapped onto the self's own existing motor action representations (as there are no corresponding action schemes in the infant observer's motor repertoire). In a recent fMRI study in which adults were viewing a person performing (non-rational vs. rational) goal-directed actions (such as someone pushing an elevator button with her knee while her hands were either free or occupied, Brass et al., 2007) reported a specific increase in the case of non-rational goal-approach (the hands-free condition) in the activation of brain areas (such as the superior temporal sulcus (STS), the temporo-parietal junction (TPJ) and the anterior fronto-median cortex) that have no mirror properties and that are typically involved in mentalization and belief attribution tasks (Ferstl & von Cramon, 2002; Fletcher et al., 1995; Gallagher et al., 2000; Goel, Grafman, Sadato, & Hallett, 1995; Grezes, Frith, & Passingham, 2004; Saxe & Kanwisher, 2003; Saxe & Wexler, 2005; Vogeley et al., 2001). These findings support the view that action understanding in terms of reasons is primarily mediated by functional brain mechanisms other than those involved in motor simulation through direct neuronal 'resonance'. In short, according to these criticisms, while the mirror neuron system may provide an important

simulation-based predictive mechanism for the anticipation and monitoring of others' observed actions, the more radical claim of standard mirror neuron accounts that understanding the intentions or goals of others' actions is solely accomplished by the direct matching of observed actions onto one's own corresponding motor schemes seems untenable.

The mirror systems view also has strong implications for the self-other distinction. If understanding of others' actions and emotions is directly mediated by shared representations that are equally activated by the self's or the other's behaviours, then it becomes hard to explain why we do not confuse others with ourselves and how we manage to attribute actions to either ourselves or to other agents. Recently, Schütz-Bosbach, Mancini, Aglioti, & Haggard (Schutz-Bosbach, Mancini, Aglioti, & Haggard, 2006) investigated this problem by an ingenious method (the so-called "rubber hand illusion") through manipulating experimentally – by induced contingency experience – whether the brain attributed the same observed action to the self versus to another agent. The study demonstrates that while the same actions attributed to another person facilitated the observer's action system, when it was attributed to the self the observer's action system was suppressed rather than facilitated. The authors conclude that contrary to the radical "shared representation" model of self-other understanding, "the motor system...includes representations of other agents as qualitatively different from the self." (p. 1834).

### *Primary intersubjectivity*

The Cartesian approach to the self a) presupposes direct introspective access to subjective intentional and emotional mind states, and b) implies the existence of prewired, universal and subjectively equally accessible intentional and emotional self states in all human individuals. It could be argued that the intersubjectivist view

implies an innatist position which leaves little room for developmental changes to the subjective sense of self induced by social environmental factors, producing individual variability in the quality and content of subjective affective and mental states across different persons. Phenotypical variations in the range and kinds of internal mental and emotional states, in their relative degree of subjective accessibility, or in the ability to use them to simulate the contents of other minds, would then best be explained as a result of genetic differences, maturational dysfunctions or brain injury. By contrast, from our constructionist perspective we claim that as our understanding of the interface of brain development and early psychosocial experience increases, we can see that the evolutionary role of the attachment relationship goes far beyond giving physical protection to the human infant. Attachment facilitates the appropriate organization of the brain processes that come to subservise social cognition and helps prepare them to equip the individual for the collaborative and cooperative existence with others for which their brain was designed (Fonagy, 2003).

Alan Sroufe (1996) and Myron Hofer (2004) were the key instigators in extending attachment theory from a concern with the developmental emergence of a complex set of social expectancies to a far broader conception of attachment as facilitating the organisation of physiological and brain regulation (Burgess, Marshall, Rubin, & Fox, 2003). More recent work has begun to articulate the associated biological pathways at least in animal models (Champagne et al., 2004; Champagne, Weaver, Diorio, Sharma, & Meaney, 2003; Francis, Szegda, Campbell, Martin, & Insel, 2003; Jaworski, Francis, Brommer, Morgan, & Kuhar, 2005; Plotsky et al., 2005; Zhang, Chretien, Meaney, & Gratton, 2005). This body of work illustrates how processes as fundamental as gene expression or changes in receptor densities can be

influenced by the infant's environment. The brain is experience-expectant (Siegel, 1999).

The Cartesian view of the mind and the self has been criticized on a number of grounds in current philosophy of mind, cognitive neuroscience, social psychology, developmental psychology, and clinical theory (Csibra & Gergely, 2006; Damasio, 1995; Damasio, 2003; Dennett, 1991; Fonagy et al., 2002; Gergely, 2002; 2004; Gopnik, 1993; Saxe, 2005; Wegner & Wheatley, 1999). Nevertheless, it has continued to influence numerous recent theories of early socio-emotional development and attachment. These developmental theories all emphasise the centrality of what has come to be termed as primary 'intersubjectivity' that is assumed to characterize the mental experience of infants during infant-caregiver interactions from the earliest phases of life (Braten, 1988, 1992; Gianino & Tronick, 1988; Hobson, 2002; Hobson, 1993; Meltzoff & Gopnik, 1993; Meltzoff & Moore, 1977; Meltzoff & Moore, 1998; Stern, 1985; Trevarthen, 1979; 1993; Trevarthen & Aitken, 2001; Trevarthen, Vandekerckhove, Delafield-Butt, & Nagy, 2006). For example, Meltzoff & Moore's (1977; 1998) well-known discovery of neonatal imitation in humans showing an innate capacity and motivation by newborns to re-enact specific parental facial displays (such as tongue protrusion, frowning, raised-eye-brows, lip protrusion, mouth opening, and some basic emotion displays, see also Field, Woodson, Cohen, Garcia, & Greenberg, 1983) have been widely cited as strong evidence for primary intersubjectivity in humans.

The notion of primary intersubjectivity assumes (a) that human infants are born with innate perceptual and inferential mechanisms to identify and attribute a rich set of subjective mental states (such as intentions, desires and feelings) to the other's mind during early contingent social interactions, (b) that from the beginning of life

the infant is aware of a relatively rich set of differentiated mental states of the self, (c) that these states of mind can be recognized as being similar or identical to the corresponding mental states expressed by the caregiver during turn-taking interactions, and (d) as a result, the infant experiences his self states as ‘shared’ with the attachment figure (e.g. Braten, 1988; Braten, 1992; for a collection of papers on intersubjectivity, see Braten, 1998; Stern, 1995; Trevarthen, 1993; Trevarthen & Aitken, 2001; Trevarthen et al., 2006).

A central characteristic of these models of primary ‘intersubjectivity’ is a shared emphasis on the continuity from infancy to adulthood of subjective emotional experience, of the kinds of ‘intersubjective’ states of interpersonal relatedness, and of the identity of basic human motives that are supposed to drive the mutual affect-regulation and attunement assumed to characterize dyadic interactions from the beginning of life. Such theories often assume – either explicitly or implicitly - a basic human-specific drive to share psychological states with others (e.g., Tomasello, Carpenter, Call, Behne, & H., 2005). (For a critical appraisal of this assumption see Gergely & Csibra, 2005a). This sharing of psychological states is often seen as the ultimate and intersubjectively shared basic goal that is inherent in and determines the structure of human interactions from the beginning of life.

The concept of primary intersubjectivity involves a ‘rich’ mentalistic interpretation of the nature of the young baby’s subjective experience of her own as well as of the caregiver’s mind states during the organized patterns of mother-infant interactions from birth (Hobson, 2002; Rochat & Striano, 1999; Trevarthen, 1993; Trevarthen & Aitken, 1994, 2001). Trevarthen, for example, (Trevarthen, 2005) showed that the newborn baby is already interested in the attention of a nearby person. The sharing of minds established at this early stage is also (implicitly or

explicitly) considered critical by many philosophers of mind (Cavell, 1994; Davidson, 1987; Wittgenstein, 1969). Trevarthen has shown that the infant appears to be endowed with the tendency to make emotionally expressive movements with voice, face and hands which are ideally adapted for interpersonal emotional expression (Trevarthen, 2001; 2005). Similarly, fMRI evidence is accumulating that infant brain organization may be well-adapted to be an 'intersubjectivity system' (Aitken & Trevarthen, 1997; Tzourio-Mazoyer et al., 2002).

The basic evidence marshaled in favor of primary 'intersubjectivity' includes the intricate organization of the early bi-directional affective and imitative interaction sequences and their characteristic contingent 'protoconversational' turn-taking structure (e.g. Beebe, Lachmann, & Jaffe, 1997; Jaffe, Beebe, Feldstein, Crown, & Jasnow, 2001; Malloch, 1999; Tronick, 1989). Examples of other early social competences include innate attentiveness to and preference for the pattern of the human face, a prewired interest in eye contact and an innate propensity to follow gaze shift when this is subsequent to direct eye contact (Csibra & Gergely, 2006; Farroni, Csibra, Simion, & Johnson, 2002; Farroni, Massaccesi, Pividori, Simion, & Johnson, 2004; Trevarthen, 2005), the innate inclination to imitate certain human facial gestures (Meltzoff & Moore, 1989), the early sensitivity and motivation to explore and analyse the causal contingency structure of interactions (Gergely & Watson, 1999; Lewis, Allessandri, & Sullivan, 1990; Watson, 1972; 1994; 2001) or the implicit understanding of others' pretend actions at 15 months (Onishi, Baillargeon, & Leslie, in press). In fact, several recent studies (Onishi & Baillargeon, 2005; Southgate, Senju, & Csibra, in press-b; Surian, Caldi, & Sperber, in press) using non-verbal implicit measures of false belief attribution (such as the violation-of-expectation looking paradigm) seem now strongly to suggest that human infants as

young as 13 months of age have the mentalizing capacity to attribute beliefs to others based on automatic monitoring of the other's perceptual access to the situation. They can anticipate the other's actions on the basis of such belief contents even when those have become outdated (false) due to a change of reality that had not been witnessed by the person. Infants by 12 months of age do not just participate in joint attention, they also actively attempt to establish it, often apparently simply to share interest in something (Liszkowski, 2006). Increasingly elaborate games are developed (Watson, 1972), in which a shared focus is established apparently to strengthen 'dyadic states of consciousness' (Tronick, 2005) which incorporate familiar tasks, gestures and objects into jointly elaborated routines. An intriguing study from Tomasello's group (Liszkowski, Carpenter, Henning, Striano, & Tomasello, 2004), observed the impact of an adult reacting to the pointing behavior of 12-month-olds. Infants were not happy when the adult simply followed the infant's pointing and looked to the object, or looked to the infant with positive affect, or did nothing. But they were satisfied when she responded by looking back and forth from the object to the infant and commented positively. Liszkowski et al (Liszkowski et al., 2004) interpret this as implying that this sharing of attention and interest was indeed their goal. An alternative interpretation of this phenomenon in terms of the theory of human pedagogy (Csibra & Gergely, 2006; Gergely, Kiraly, & Egyed, in press) considers its primary function to be epistemic in nature and will be considered in detail below. Such interactions represent a 'proto-interrogative' request for relevant information about the object of joint attention (such as its name, function, or valence) for the infant to learn. In fact, infants of 12 months also happily point just to inform an adult of the location of a misplaced object they have no direct interest in (Liszkowski et al., 2004).

However, many researchers of early emotional development do not share the ‘intersubjectivist’ view that differentiation of discrete emotions, to which we have conscious access, is present during the first few months. They see the behaviours described above as consequences of either early self-organizing dynamic systems processes (Fogel et al., 1992; Lewis & Granic, 2000) or cognitive developmental processes leading to the early socialization of, and sensitization to, different feelings during affect-regulative caregiver-infant interactions (Gergely & Watson, 1996, 1999; Sroufe, 1979; 1996). Lewis and Michaelson (Lewis & Michaelson, 1983) also argue that during the earliest phases of infancy internal states and expressive behaviors are not yet coordinated. In their view, conscious feelings that are linked to discrete expressive displays emerge only later due to the influence of socialization and cognitive growth (see also Barrett & Campos, 1987; Kagan, 1992; Lewis & Brooks, 1978).

Critics of primary ‘intersubjectivity’ point to plausible alternatives to the central functionalist and motivational interpretations proposed by the ‘intersubjectivist’ school as the primary organizational factors behind the ‘protoconversational’ turn-taking structure of early affective caregiver-infant interactions. A number of well-documented innate cognitive and perceptual capacities of the human infant can account for the early turn-taking structural organization of affective interactions without invoking mentalization or attribution of ‘intersubjective’ emotional states to the other (Csibra & Gergely, 2006; Gergely et al., in press; Gergely & Watson, 1996; 1999; Watson, 1994). We do not believe that the primary function of human infants’ innate sensitivity to contingent turn taking is the fulfillment of any of these functions (Gergely, 2002). For example, filial attachment is established in many mammalian and avian species without extended

protoconversational routines. It also seems to be an overstatement that mothers and infants are both motivated by and subjectively aware of ‘sharing’ each other’s mental or emotional states in these interactions. No doubt, they both enjoy these situations, and one can say that they, in fact, ‘share’ this positive hedonic experience, at least in the sense of being simultaneously in a similar affective state. But apart from generating simultaneous enjoyment, what aspect of the evidence would indicate that any other, more differentiated discrete emotional states are shared during turn taking? Do mothers and babies share sadness, fear, anger, disgust, or distress just in order to be in the same internal state? The evidence suggests that during their first 6 months infants may not yet be able to recognize a number of basic categorical emotions of others (see Gergely, 2002; Nelson, 1987) even though mothers will certainly react to these emotions if their child expresses them. But this kind of adequate maternal reaction will only rarely involve an imitative emotion expression or even an initiation of a turn-taking interchange: she is much more likely to just pick the child up and establish close bodily contact with him. Engaging in protoconversational turn-taking is neither a typical nor an effective response when the baby is in need of soothing.

Csibra & Gergely (Csibra & Gergely, 2006; Gergely, 2002) have provided a detailed critical analysis of the flaws inherent in the primary ‘intersubjectivity’ theory. They argue that early social interactional phenomena are better and more coherently explained in an entirely different theoretical framework as manifestations of a human-specific adaptation for ‘pedagogy’ (Gergely & Csibra, 2006), a social communication learning system of mutual design that evolved to ensure fast and efficient transfer of relevant cultural knowledge between conspecifics (to be discussed in more detail later in this paper). We shall return to this more parsimonious account of the development

of social cognition after considering recent research on the neuroscientific basis of social cognition. The bulk of the evidence points to the functional independence of the subjective sense of self. The current evidence indicates that a range of structures are involved and there are excellent reviews of the relative merits of the intersubjectivist and constructionist positions (e.g., Reddy & Morris, 2004).

### **Brain structures directly relevant to mentalizing**

The diversity of brain structure recruited by mentalization points to the multicomponent character of this capacity. Changes in human brain structures particularly concerned with cognitive mediation were closely associated with evolutionary changes in social intelligence. Mentalization may not be a uniquely human capacity in so far as important components of mentalisation (e.g. that seeing leads to knowing) have independently evolved in a number of non-human species in competitive niches (e.g. birds such as ravens and scrub-jays as well as primates Bugnyar & Heinrich, 2005; Emery & Clayton, 2004; Hare, Call, & Tomasello, 2001; Tomasello, Call, & Hare, 2003). However, comprehensive social understanding that makes use of a wide range of intentional states and encompasses both other and self seems to be a solely human capacity. A broad brush approach to its localisation might be to identify brain aspects unique to humans. There is a class of large and clustered spindle cells unique to humans in the anterior cingulate cortex (Allman, Hakeem, & Watson, 2002; Allman, Hakeem, Erwin, Nimchinsky, & Hof, 2001; Nimchinsky et al., 1999), and there are other unique features of human neuroanatomy such as increased lateralization that underpin social interpretation. There was a disproportionate expansion in humans of the right prefrontal cortex and the frontal pole (Holloway, 1996; Zilles et al., 1996), areas of the brain that have been shown by imaging studies to be involved in self-awareness, the ability to remember personal

experiences, and to project oneself into the future (Tulving, 2002). There was a modest proportional expansion in parts of the prefrontal cortex (about 10%, Semendeferi & Damasio, 2000), with an increased richness of interconnections between neurones in these areas (Miller & Cohen, 2001). The evolution of the human prefrontal cortex has been suggested to be closely related to the emergence of human morality (Allman et al., 2002; Grafman, 1995; Moll, de Oliveira-Souza, & Eslinger, 2003; Moll, Zahn, de Oliveira-Souza, Krueger, & Grafman, 2005; Wood & Grafman, 2003). Motivational functions are integrated with functions predicting the outcome of actions in others. This has been argued to be a major advantage permitting the cultural explosion of the Upper Paleolithic period (Mithen, 2005). These structures underpin awareness of social dynamics and the capacity to imagine responses to changes in these dynamics (Geary, 2005). Developmental and neuropsychological observations help us in identifying some of the likely components of the complex function of mentalisation.

The capacity to engage in *shared or joint attention* is seen as a ‘precursor’ to mentalisation by most of those interested in ‘theory of mind’ (i.e. preschoolers’ understanding that people act on their beliefs even when these are false). Tomasello (1995) argued that the capacity to engage in joint attention is the fundamental human ability that permits the infant to experience interaction and acquire language that in turn opens the door for more complex understandings of the social world. The representation of triadic relationships between self, other and object (“mother sees that I see the cup”) mediate behaviours such as protodeclarative pointing and gaze monitoring which indicate shared attention (Scaife & Bruner, 1975). A complex recent investigation succeeded in developing video stimuli that induced an experience of joint attention in the observer, allowing a comparison of brain activation in joint

attention and non-joint attention conditions (Williams, Waiter, Perra, Perrett, & Whiten, 2005). In this study the experience of shared attention was accompanied by activity in the ventromedial frontal cortex, the left superior frontal gyrus (BA10), cingulate cortex, and caudate nuclei. It is important to note that the ventromedial frontal cortex has been consistently shown to be activated during mental state attribution tasks (see below). It could be that the left superior frontal gyrus serves the cognitive integration function, which in this case seems to utilize a perception-action matching process and overlaps with a location of increased grey matter density that was found by the same group to be associated with autistic spectrum disorder. This would be consistent with the suggestion here that the neural substrate of joint attention also serves a mentalizing function.

Mentalization entails *inhibitory controls* necessary for the child to suppress the pre-potent assumption that everyone else shares the same knowledge and beliefs. These controls may enable the child to recognize the existence of separate minds (Leslie, 2000; Perner & Lang, 2000). Studies of response conflict unequivocally indicate that inhibitory controls require the activation of the anterior cingulate cortex (ACC) along with the dorsolateral prefrontal cortex (DLPFC) and superior parietal lobe (Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999; Braver, Barch, Gray, Molfese, & Snyder, 2001; Milham et al., 2001; Miller & Cohen, 2001; Sylvester et al., 2003). Imaging studies confirm the activation of ACC during tasks calling for a theory of mind (Calarge, Andreasen, & O'Leary, 2003; Kobayashi, Glover, & Temple, 2006; Vogelely et al., 2001) Animal research has shown the key role of the ACC and hippocampus in emotional regulation of activity and social relating, and the relevance of this to child psychopathology is becoming clearer (see Allman et al., 2001; Amaral, 2003). An fMRI study that contrasted subjects playing a game involving interpersonal

interactions with either a computer or a putative human partner (both conditions involved interaction with a computer) found DLPFC involvement in both conditions (although knowing the partner was a computer led to greater activation from this cognitive control area (Rilling, Sanfey, Aronson, Nystrom, & Cohen, 2004). A limited number of studies have implicated the same area particularly in the right hemisphere with changes in moral behavior (Eslinger, 2001; Miller, Chang, Mena, Boone, & Lesser, 1993; Perry & al, 2001; Tranel, Bechara, & Denburg, 2002) and empathy (Vollm et al., 2006).

Accumulating evidence indicates that some structures responsible for *understanding affect* in others are independent from systems that mediate belief attribution. Neural systems associated with the perception and experience of emotions include the extrastriate cortex, right parietal cortex, right fusiform gyrus, orbitofrontal cortices, amygdala, insula, and basal ganglia (Adolphs, 2002; Bachevalier & Loveland, 2006; Blair, 2003; Canli & Amin, 2002; Hamann, 2003; LeDoux, 2000; Posner, Russell, & Peterson, 2005; Wildgruber et al., 2005). The activation of areas associated with emotionally salient stimuli, such as the posterior cingulate, may also be involved in interpersonal interaction tasks when the response of the partner is likely to generate feelings that affect interpersonal judgements (Posner et al., 2005; Rilling et al., 2004).

Of specific concern to researchers has been the location of systems involved in emotional concern with others (often referred to as empathy, Decety & Jackson, 2004). The ventromedial prefrontal cortex, in association with the insula and limbic cortices, seems to be a regulator of emotion, and of motivation for monitoring social behaviour, including social self-awareness and moral behaviour (e.g. Damasio et al.,

2000). It seems that theory of mind (belief attribution) and empathy may be associated with overlapping but distinct neuronal networks (Vollm et al., 2006). Both involve the medial prefrontal cortex, temporoparietal junction and temporal poles. However, empathy is associated with enhanced activations of paracingulate, anterior and posterior cingulate and amygdala, while belief attribution entails increased activations in lateral orbitofrontal cortex, middle frontal gyrus, cuneus and superior temporal gyrus (see below). While mentalizing others' cognitions and empathy both rely on networks associated with making inferences about mental states of others, empathic responding also requires the additional recruitment of networks involved in emotional processing (Vollm et al., 2006). Patients with developmental damage to the ventromedial prefrontal cortex have been observed to be remarkably impaired in terms of moral behavior while apparently unimpaired in specific moral reasoning tasks (Eslinger & Damasio, 1985). Later studies found that the development of both moral reasoning and moral behavior can be undermined by damage to the ventromedial prefrontal cortex (Anderson, Bechara, Damasio, Tranel, & Damasio, 1999; Eslinger, Grattan, Damasio, & Damasio, 1992) yielding to a presentation somewhat like childhood psychopathy.

*Belief attribution*—reasoning about false beliefs or making judgements about someone's knowledge or ignorance about a topic—increases brain activity in the medial prefrontal cortex (anterior to the ACC), temporal poles bilaterally, anterior superior temporal sulcus, and bilateral temporo-parietal junction extending into posterior temporal sulcus (Ferstl & von Cramon, 2002; Fletcher et al., 1995; Gallagher et al., 2000; Goel et al., 1995; Grezes et al., 2004; Saxe & Kanwisher, 2003; Saxe & Wexler, 2005; Vogeley et al., 2001). These areas appear to be specific to belief attribution rather than general reasoning about people or reasoning about

non-mental false representations or hidden cues in general. The medial prefrontal cortex showed significant increases in activation during false belief stories but not stories about true beliefs that could be action based and that require no representational component (Fletcher et al., 1995; Gallagher et al., 2000). In another study (den Ouden et al., 2005), subjects who were asked to keep in mind an intention whilst carrying out the ongoing Causality task activated a network of regions including the medial prefrontal cortex (mPFC), the superior temporal sulcus and the temporal poles bilaterally. From such observations some reviews limit the uniquely theory-of-mind area to the medial prefrontal region of the cortex (Gallagher & Frith, 2003). A task that involved real-time interpersonal interaction (the prisoner's dilemma task) recruited greater activation from the anterior paracingulate cortex, and the posterior superior temporal sulcus for trials when participants believed they were playing with a person as opposed to a computer (Rilling et al., 2004). Distinct regions of the medial prefrontal cortex contribute differentially to social cognition: the ventral medial prefrontal cortex is activated during the analysis of social content and a more dorsal part of the medial prefrontal cortex subserves the detection of self-relevance and may thus establish an intersubjective context in which communicative signals are evaluated (Schilbach et al., 2006).

The cortical systems associated with *attribution of desires and goals* have been investigated using vignettes, cartoons, and animations that depict or suggest a character's intentions. Looking at these stimuli tends to be associated with moderately enhanced activity in the brain regions linked to belief attributions, including the medial prefrontal cortex and posterior superior temporal sulcus (Brunet, Sarfati, Hardy-Bayle, & Decety, 2000; Buccino et al., 2001; Castelli, Happe, Frith, & Frith, 2000; Gallagher et al., 2000; Saxe & Kanwisher, 2003; Schultz et al., 2003). When

subjects engage in simple games with an unseen agent as contrasted with playing a computer, the activation of the medial prefrontal cortex is increased (Gallagher, Jack, Roepstorff, & Frith, 2002; McCabe, Houser, Ryan, Smith, & Trouard, 2001). This part of the brain appears to respond more to any story that contains a person than to stories that do not involve humans (Gallagher et al., 2002; Saxe & Kanwisher, 2003)

There have been several attempts to draw up comprehensive models of the neural basis of mentalisation which incorporates all the components listed above. The brain structures involved appear to be manifold yet reasonably specific to aspects and components of the task of understanding mind states. Interestingly, the majority of brain structures subserving social cognition appear to be also implicated in the processing of emotions (Grady & Keightley, 2002). This demonstrates a putatively critical set of relations between feeling and thought and, perhaps not surprisingly, implicates the basis of disorders of emotion as occurring in the same neural systems.

Some authors recommend separating social cognition into an implicit and explicit system. For example, Satpute and colleagues (Satpute & Lieberman, 2006) differentiate structures underpinning a reflexive and a reflective system of social cognition. Reflexive systems correspond to automatic processes and include the amygdala, basal ganglia, ventromedial prefrontal cortex, dorsal anterior cingulate cortex, and lateral temporal cortex. Reflective systems correspond to controlled processes and include lateral prefrontal cortex, posterior parietal cortex, medial prefrontal cortex, rostral anterior cingulate cortex, and the hippocampus and surrounding medial temporal lobe region.

The ‘two-component’ model of mentalising implied by the implicit – explicit dichotomy is likely to have a developmental dimension. Early implicit intuitive

mentalising (Frith, 1989) or a socio-perceptual awareness of mind (Tager-Flusberg, 2001) and its failure may have greater social impact than the acquisition in late preschool years of an explicitly representational concept of mind that would be revealed by performance on a false-belief task. Recent findings have revealed that children have clear expectations about the behaviour of a person with false belief even at 15 months (Onishi & Baillargeon, 2005; Southgate, Senju, & Csibra, in press-a) and that they attempt to make sense of others' intentional actions from the second half of the first year (Csibra et al., 1999; Guajardo & Woodward, 2004; Király, Jovanovic, Prinz, Aschersleben, & Gergely, 2003; Sommerville & Woodward, 2005). Infants by the age of 15 months do not automatically imitate an actor's observed behaviour; rather, they imitate selectively those aspects that the demonstrator's communicative cues indicate are relevant (Gergely, Bekkering, & Kiraly, 2002). One-year-old infants become angry only when an actor is unwilling to give them a toy, not when she is unable (Behne, Carpenter, Call, & Tomasello, 2005). Even earlier, when an 5 or 6 months old infant watches an agent act on objects he/she may attribute to the agent goals and dispositions that help explain and predict the agent's actions (e.g. Bíró & M., in press; Kamerawi et al., 2005; Luo & Baillargeon, 2005). Thus, it seems fairly well established that infants in the first year of life already possess implicit understanding of intentional action although the exact nature of these abilities and the implications they have for our understanding of the development of early psychological reasoning remain controversial (e.g. Csibra & Gergely, in press; Meltzoff, 2005; Tomasello et al., 2005; Woodward, 2005).

The best integrative developmental summary model to our mind was provided by Simon Baron-Cohen (Baron-Cohen, 2005; Baron-Cohen & Belmonte, 2005). Labelling his model of mentalisation the 'empathising system' he specifies an

emotion detector (ED), intention detector (ID) and an eye direction detector (EDD) as developmental and neuropsychological precursors to the functioning triadic interaction system SAM (shared attention mechanism). SAM enables the emergence of the theory of mind mechanism (TOMM) charged with mediating M-representations consisting of an Agent-attitude-proposition (e.g. “Mother- believes Johnny - took the cookies”). The empathising system (TESS) uses E-representations of the form: Self-Affective state-proposition (e.g. “I am sorry-you feel hurt-by what I said.”). An important constraint that Baron-Cohen hypothesizes for TESS is that it will always create representations where emotion in the other is consistent with the Self-Affective state (e.g. it will not create the representation that “I am pleased that you are in pain”). It has to be a state that the self can generate in relation to the presumed state in the other (this constraint is assumed not to be present in psychopaths Blair, 2003). While emotion understanding and belief-desire reasoning or theory of mind are essential aspects of children’s sociocognitive understanding, it would probably be an error to overemphasise the separation between the two (de Rosnay & Hughes, 2006). They interact and only in combination generate mentalization or social understanding. A schematic model attempting to summarise information from this section within Baron-Cohen’s (2005) model is shown in Figure 1.

[INSERT FIGURE 1 ABOUT HERE]

### ***Speculation about possible evolutionary links between attachment and mentalization***

Mentalization is arguably the evolutionary pinnacle of human intellectual achievement. But what has driven the selection processes of the 2 million or so years of human evolution towards a recognition of mental states in others? Was it to meet the periodic challenges the physical environment presented to our ancestors, who

were presumably only somewhat more agile and strong than we currently are?

Surprisingly, leaps forward in human brain size in the course of evolution do not correspond to what we know about ecological demands on our hominin ancestors (e.g. climate change, predators, availability of prey). The evolutionary biologist Richard Alexander, (1989) proposed that our exceptional intelligence evolved not to deal with the hostile forces of nature but rather to deal with competition from other people. This occurred only after our species had already achieved relative dominance over their environment. At that point it seems we became our “own principal hostile forces of nature” (Alexander, 1989, p.469). To meet this challenge to the survival of our genes those with common genetic material had to collaborate. This ‘Machiavellian’ (Byrne & Whiten, 1988) or competitive aspect may not be specific to humans and evidently other species also evolved aspects of mentalisation (such as ‘seeing leads to knowing’) when this was supported by competitive pressure (Bugnyar & Heinrich, 2005; Emery & Clayton, 2004).

As our understanding of the interface of brain development and early psychosocial experience increases, we can see the evolutionary role of the attachment relationship as going far beyond giving physical protection to the human infant. From the social constructionist perspective adopted here we might speculate that attachment may have a role in facilitating the development of brain processes that come to subserve social cognition and help organise and prepare the individual for the collaborative and cooperative existence with others for which his or her brain was designed (Fonagy, 2003). Mentalizing refers both to reflecting on the contents of others’ minds, and to having knowledge of one’s own intentions, desires and thoughts. Representing the contents of one’s own mind is seen as tapping into the same meta-representational capacity required for representing the contents of

another's mind (den Ouden et al., 2005; Frith & Frith, 2003). Self awareness and awareness of the mental states of others are closely linked, certainly in terms of the brain areas involved across a number of domains including, for example, the experience of pain (Jackson et al., 2005; Singer et al., 2004). Aspects of mentalization are there not just to facilitate human collaboration and positive relationships but also to facilitate individual social survival. Mentalization may serve competition: as the antlers of a reindeer are there to fight other reindeer rather than to catch prey or ward off predators. Self-awareness enables us to modify the way we wish to present ourselves, and to mislead (Barrett & Henzi, 2005; Brune, 2001; Striedter, 2006). While other species also appear to have the capacity to deceive (Bugnyar & Heinrich, 2005; Emery & Clayton, 2004; Hare et al., 2001), in humans the right prefrontal cortex may have the special function of enabling us 'to see ourselves as others see us so that we may cause competitive others to see us as we wish them to' (Alexander, 1990, p. 7). The original evolutionary function of experiences such as daydreams and fantasies might have been to allow individuals to form goals, carry out plans and integrate these within a seamless knowledge of their life history (Levine, 1999).

Mentalization permits superior adaptation to the physical environment in part through facilitating social collaboration and well-functioning kinship groups. It also, of course, supports competition for survival when social groups are in competition. As in other species, the competitive niches we occupy drive the social brain to reach higher and higher levels of sophistication. Since the mind needs to adapt to ever more challenging competitive conditions and since these conditions are socially determined and therefore highly variable across geographical locations and time, the capacity cannot be fixed by genetics or constitution; it is left to be optimized for the infant

through a prolonged childhood by a group of trusted kin (attachment figures). We are suggesting that evolution has left it to the intimate relationships of early childhood to elaborate the capacity for social cognition fully. The capacity for mentalization, along with many other social-cognitive capacities, is designed to evolve out of the experience of social interaction with the early social environment, including that provided by the child's caregivers. From this standpoint it is interesting to note that increased sophistication in social cognition evolved over the same evolutionary period as apparently unrelated aspects of development, such as increased helplessness in infancy, prolonged childhood, and the emergence of intensive parenting (Geary & Huffman, 2002; Hrdy, 2000; Siegal & Varley, 2002).

### *Activation of attachment & deactivation of mentalization: Reciprocity*

Recent neuroimaging studies have further linked some attachment phenomena to the *deactivation* of mentalizing. Rodent research on the neurobiology of attachment has linked this to the mesocorticolimbic dopaminergic reward circuit, which also plays a key role in mediating the process of addiction (Insel, 1997; MacLean, 1990; Panksepp, 1998). It is unlikely that nature created a brain system specifically to subserve cocaine and alcohol abuse; addictions are the accidental byproduct of the activation of a biological system that underpins the crucial evolutionary function of attachment (Insel, 1997; MacLean, 1990; Panksepp, 1998). Ironically, attachment can be construed as an addictive disorder (Insel, 2003) in the sense that falling in love, which is stimulated by social/sexual activity, entails the activation of an oxytocin- and vasopressin-sensitive circuit within the anterior hypothalamus (MPOA) linked to the VTA and the nucleus accumbens (Insel, 2003). Unfortunately, most evidence concerning the involvement of mesocorticolimbic dopaminergic pathways in the neurobiology of attachment comes from rodent research (e.g. Lim, Murphy, & Young,

2004; Lim, Wang et al., 2004; Lim & Young, 2004). Human fMRI studies also tend to indicate specific activation of these reward-sensitive pathways in the brain of somebody seeing their own baby or partner as compared to another familiar baby or other people's partners (Nitschke et al., 2004). Early deprivation affects the vasopressin and oxytocin systems that are critical for the establishment of social bonds and the regulation of emotional behaviour (Fries, Ziegler, Kurian, Jacoris, & Pollak, 2005). The neural bases of attachment, including the limitations of neuroscience research in this area, have recently been comprehensively reviewed for both human and non-human data for this volume (Swain, Lorberbaum, Kose, & Strathearn, in press).

In two separate imaging studies, Bartels and Zeki (Bartels & Zeki, 2000; 2004) reported that the activation of areas mediating maternal and/or romantic attachments appeared simultaneously to suppress brain activity in several regions mediating different aspects of cognitive control and including those associated with making social judgements and mentalizing. Bartels and Zeki (Bartels & Zeki, 2004) suggest grouping these reciprocally active areas into two functional regions. The first of these systems includes the medial prefrontal, inferior parietal and medial temporal cortices mainly in the right hemisphere, as well as the posterior cingulate cortex. These areas are part of the circuitry specialised for attention and long-term memory (Cabeza & Nyberg, 2000), and they also have variable involvement in both positive (Maddock, 1999) and negative (Mayberg et al., 1999) emotions. It is argued that these areas may be specifically responsible for integrating emotion and cognition (e.g. emotional encoding of episodic memories Maddock, 1999). In addition, lesion studies suggest a role in judgements involving negative emotions (Adolphs, Damasio, Tranel, Cooper, & Damasio, 2000). It is possible that, as projections from the affect-oriented

limbic/paralimbic regions modulate the activity of these areas, they could subserve mood-mediated inhibition or enhancement of cognitive processing (Mayberg et al., 1999). These areas also may play a role in recalling emotion-related material and generating emotion-related imagery (Maddock, 1999) that may be relevant to understanding the typology of attachment.

The second set of brain areas observed to be deactivated by the activation of the attachment concerns, included the temporal poles, parietotemporal junction, amygdala, and mesial prefrontal cortex. The authors argued that activation of these areas is consistently linked to negative affect, judgements of social trustworthiness, moral judgements, theory-of-mind tasks, and attention to one's own emotions. As we have considered above, this system probably constitutes part of the primary neural network underlying the ability to identify and interpret mental states (both thoughts and feelings) in other people (Frith & Frith, 2003; Gallagher & Frith, 2003) as well as in the self (Gusnard, Akbudak, Shulman, & Raichle, 2001). The structures are also thought to be associated with intuitive judgements of moral appropriateness (Greene & Haidt, 2002) and of social trustworthiness based on facial expressions (Winston, Strange, O'Doherty, & Dolan, 2002).

The pattern of activation of the attachment system and the two overlapping cognitive information processing control systems may have implications for our understanding of the nature of individual differences in attachment behaviour, the relationship of attachment and mentalization and, consequently, our understanding of dysfunctions associated with mentalizing deficits. Broadly, three conditions in attachment relationships may be assumed to inhibit or suppress aspects of social cognition associated with mentalizing the attachment figure. First, the love-related

activation of the attachment system, mediated by dopaminergic structures of the reward system in the presence of oxytocin and vasopressin, probably inhibits neural systems that underpin the generation of negative affect that sometimes may prompt problem-solving social cognitions entailing mentalizing. This is to be expected: a key function of the attachment system is to moderate negative emotions in infancy and, indeed, throughout life (Sroufe, 1996). Second, threat-related activation of the attachment system (e.g., triggered by perceived threat, loss or harm) deactivates mentalizing by virtue of evoking intense arousal and overwhelming negative affect (Arnsten, 1998; Arnsten, Mathew, Ubriani, Taylor, & Li, 1999; Mayes, 2000). Third, a stable, secure, predictable attachment relationship may be most effective in pre-empting threat and probably obviates the need for the frequent activation of the attachment system. It may be helpful at this stage to remind ourselves that the attachment system is assumed to have been designed to be activated by fear, often associated with the loss of protection of the attachment figure (Bowlby, 1959; 1969; 1973). An unpredictable, insecure caregiver-infant relationship is likely to call more frequently for the activation of the attachment system than a predictable one and thus more frequently bring about the deactivation of neural structures underpinning aspects of social cognition. There is also evidence that the level of attachment anxiety is positively correlated with activation in emotion-related areas of the brain (e.g., the anterior temporal pole, implicated in sadness) and inversely correlated with activation in a region associated with emotion regulation (orbitofrontal cortex) (Gillath, Bunge, Shaver, Wendelken, & Mikulincer, 2005). The findings suggested that anxiously attached people might under-recruit brain regions normally used to down-regulate negative emotions. Those high on avoidance failed to show as much deactivation in two brain regions (subcallosal cingulate cortex; lateral prefrontal cortex) as less

avoidant participants suggesting that their suppression was less complete or less efficient, in line with results from previous behavioral experiments.

At this point we might summarise the material reviewed so far as follows: there are substantial biological (mirror neurons) and psychological (intersubjectivity) accounts of the emergence of the subjective self that do not necessarily call for a social constructionist approach to the development of mentalization. However we have also seen that a range of brain processes that are likely to be involved in the mediation of this complex function and that social cognition is closely linked, at least anatomically to the regulation of emotional experiences. Functional links at the level of brain processes also appear to couple mentalization and attachment. Whilst an important evolutionary function of mentalization may be to provide advantage in competition, its full development at least in part must mostly take place in the context of a protective (attachment) relationship. It is suggested that the parent-child relationship is likely to facilitate the unfolding of the subjective self precisely because in this biological context competitive pressures are likely to be minimised. As it is insecure unpredictable parent-child relationships that are most likely to activate the attachment system, we may predict, on the basis of these recent neuro-imaging data alone that a secure parent-child bond is most likely to facilitate the development of mentalization as it is likely to be associated with limited inhibitory effects on the brain networks subserving mentalization. This is a somewhat mechanical model that nevertheless may form part of the social constructionist approach to the emergence of subjective selfhood proposed here. In the next section we shall consider the literature on social influences on the development of mentalization which give further support to the social constructionist perspective.

## **Social influences on the development of mentalisation**

Why should children's social experiences be associated with the developmental unfolding of mentalizing? There is a line of thinking that considers the variability across individuals in the stage at which ToMM or TESS functions emerge to be of little relevance as in all but extreme cases these capacities emerge in any case (e.g. Jenkins & Oatley, 2004). From our social constructionist perspective we consider that such variability offers vital clues about the mechanisms that underpin the development of mentalization (Carpendale & Lewis, 2006). Social experiences that are correlated with early acquisition of mentalising may be argued to be more endowed with the elements that promote the emergence of this vital capacity.

The nativistic position entails the assumption that children's social environments can trigger but cannot determine the development of theory of mind (Baron-Cohen, 2005; Leslie, Friedman, & German, 2004). There is some evidence that the timetable of theory of mind development is universal (Avis & Harris, 1991; Wellman, Cross, & Watson, 2001), the bulk of the evidence, however, is not consistent with the assumption of a universal timetable and suggests substantial cultural differences in the rate and order of emergence of theory of mind skills (see review by Carpendale & Lewis, 2006).

Behaviour genetic data has been traditionally considered key to assessment of social influences. At the extreme of low mentalizing, there has for some time been compelling evidence of genetic influences (e.g. Dorris, Espie, Knott, & Salt, 2004). Similarly, a relatively small scale (n=120) study suggested that theory of mind scores of 40 months old twins were powerfully influenced by genetic factors and had little if any shared environmental variance (Hughes & Cutting, 1999). The study estimated

heritability at 67% with the remaining 33% accounted for by non-shared environmental factors. By contrast, in a larger study, (Hughes et al., 2005) with a longitudinal twin sample of 1,116 sixty-month-old twin pairs who completed a comprehensive battery of ToM tasks, behavioral genetic models of the data showed that environmental factors explained the largest part (48%) of the variance in ToM performance. Individual differences in ToM were striking and strongly associated with verbal ability. Non-shared influences on ToM were also marked (44%). Bivariate genetic analysis revealed that to the extent that genetic factors can be said to influence ToM these were the same as those that determine verbal ability and account for a relatively small proportion of the variance in this ability (15%). Environmental influences that were shared by the twins on verbal ability also had impact on ToM (21%). Further, non-shared environmental influences (parental and other non-genetic influences that are specific to each child) were not common to verbal ability and ToM. Forty-four percent of this variance in ToM was non-shared and specific to ToM. The possible underlying proximal mechanisms considered by the authors included maternal speech and mind-mindedness, sibling interactions, and peer influences. Interestingly, attachment classification is rare among behaviours in showing little heritability in twin studies (Bokhorst et al., 2003; O'Connor, Croft, & Steele, 2000; O'Connor & Croft, 2001) and the shared environmental influence accounts for a similar proportion of the variance (53%). This once again hints at the possible cooperative advantage of leaving early development of emotionally invested ties and related social cognition maximally open to environmental influence and social heredity.

### *Family structure and family size*

Evidence for the influence of social life on mentalisation began with findings concerning family size. Studies in the early 90s reported that children with older siblings passed ToMM tasks earlier (Perner, Ruffman, & Leekman, 1994; Ruffman, Perner, Naito, Parkin, & Clements, 1998). Children who had infant or adolescent siblings benefited little (Peterson, 2000b) but those with relatively low verbal abilities benefited most (Jenkins & Astington, 1996). The effect seems to be related not just to siblings but to the number of older family members who spend time with the child (Lewis, Freeman, Kyriakidou, Maridaki-Kassotaki, & Berridge, 1996). As not all studies show the effect (Cole & Mitchell, 2000; Cutting & Dunn, 1999; Peterson & Slaughter, 2003) it is likely that family size is a 'proxy' for particular activities that take place with older siblings and others that do not take place without them, for example listening to talk about mental states (Jenkins, Turrell, Kogushi, Lollis, & Ross, 2003). This is supported by both the outstanding observational studies of family interaction by Dunn and her colleagues (Dunn, Brown, & Beardsall, 1991; Dunn, Brown, Somkowski, Telsa, & Youngblade, 1991) and associations of mentalization with other demographic variables such as social class (Cutting & Dunn, 1999).

### *Play activities with peers*

The family size literature, the advantage to children who have siblings of an age to be partners in fantasy and pretend play, suggests that play, or more specifically pretend play with its characteristic suspension of external considerations, may play a crucial role in the facilitation of mentalization. Children who score high on false belief tests also frequently engage in pretence (Dunn & Cutting, 1999; Taylor & Carlson, 1997) and have discussions of pretend roles and play proposals in preschool

(Jenkins & Astington, 2000). Preschoolers with siblings of an age to be partners in pretend or fantasy play appear to be superior in mentalization as they are more likely to pass the false belief test (Jenkins & Astington, 1996; Perner et al., 1994; Peterson, 2000). Blind children who have a dearth of pretend play and lack private or social imaginative activities (Fraiberg, 1977; Tröster & Bambring, 1994) and have trouble comprehending others' pretending (Hughes, Dote-Kwan, & Dolendo, 1998; Lewis, Norgate, Collis, & Reynolds, 2000) tend to manifest a delay in acquiring ToM (Green, Pring, & Swettenham, 2004; Hobson & Bishop, 2003; McAlpine & Moore, 1985; Peterson, Peterson, & Webb, 2000). Pretend play involves mental representations as well as relatively complex syntactic structures to establish alternative realities (de Villiers, 2005). Engaging in pretending involves the child sharing others' mental perspectives, conceptualising variations on his/her own reality, and setting against each other events that are real and symbolic representations of these. Pretence may be a precursor of the ability to pass the false belief test but it is equally likely to be a consequence (Macguire & Dunn, 1997; Slomkowski & Dunn, 1996) and while pretend may be argued to create a social zone of proximal development around the child to facilitate the practice of mentalization, pretend play is far from an essential aspect of all peer relationships (Dunn & Brophy, 2005).

### *Quality of parenting - Secure attachment associated with parental mind-mindedness*

The quality of parenting appears to have a complex relationship with mentalization. In an early study more reflective parenting practices were associated with precocious understanding of false beliefs (Ruffman, Perner, & Parkin, 1999). Consistent with this approach, many findings suggest that the nature of family interactions, the quality of parental control (Astington, 1996; Cutting & Dunn, 1999; Dunn, Brown, Somkowski et al., 1991; Ruffman et al., 1999; Vinden, 2001), parental

discourse about emotions (Denham, Zoller, & Couchoud, 1994; Meins et al., 2002), the depth of parental discussion involving affect (Dunn, Brown, & Beardsall, 1991) and parents' beliefs about parenting (Baumrind, 1991; Ruffman et al., 1999; Vinden, 2001) are all strongly associated with the child's acquisition of a coherent conceptual apparatus for understanding behavior in mentalistic terms. Similarly, parents whose disciplinary strategies focus on mental states (e.g. a victim's feelings, or the non-intentional nature of transgressions) have children who succeed in ToM tasks earlier (Charman, Ruffman, & Clements, 2002; Sabbagh & Callanan, 1998).

However, in another early study boys whose parents exercised greater strictness and severity of discipline were found to be more advanced on ToM tests (Hughes, Deater-Deckard, & Cutting, 1999). A more recent study, (Pears & Moses, 2003) reported power-assertive parenting (including spanking and yelling) to be apparently retarding the understanding of false beliefs. Before accepting the obvious conclusion that less power-assertive parenting facilitates mentalisation we should also consider the possibility of a child-to-parent effect, namely that less mentalizing children are more likely to elicit controlling parenting behavior.

A relation between attachment in infancy and early social understanding was reported by Bretherton, Bates, Benigni, Camaioni, & Volterra (Bretherton, Bates, Benigni, Camaioni, & Volterra, 1979) who found that children who were securely attached at age 12 months used more protodeclarative pointing at age 11 months than other infants. Laible and Thompson (Laible & Thompson, 1998) also found that securely attached children have higher competence in understanding negative emotion. There is general agreement in the classical literature that, as well as increasing attachment security in the child, the harmoniousness of the mother-child relationship contributes to the emergence of symbolic thought (Bretherton et al., 1979

p.224; see also Mahler, Pine, & Bergman, 1975; Vygotsky, 1978; Werner & Kaplan, 1963). Bowlby (1969) recognized the significance of the developmental step entailed in the emergence of “the child’s capacity both to conceive of his mother as having her own goals and interests separate from his own and to take them into account” (Bowlby, 1969 p.368). Peter Hobson (2002) in a powerful and persuasive monograph described ‘the triangle of relatedness’ as the source of alternative perspectives upon the world, which he considers to be lacking in autism and to be the foundation of symbolic thought.

A significant body of observations offers some support for the suggestion that the quality of children’s primary attachment relationship may facilitate the development of mentalization, leading to passing standard theory of mind tasks somewhat earlier (e.g., de Rosnay & Harris, 2002; Fonagy, Redfern, & Charman, 1997; Fonagy & Target, 1997; Harris, 1999; Meins, Fernyhough, Russel, & Clark-Carter, 1998; Ontai & Thompson, 2002; Raikes & Thompson, 2006; Steele, Steele, Croft, & Fonagy, 1999; Symons, 2004; Thompson, 2000). For example, the Separation Anxiety Test, a projective test of attachment security, was found to predict belief-desire reasoning capacity in 3½ to 6 year old children when age, verbal ability and social maturity were all controlled for (Fonagy, Redfern et al., 1997). In this task the child is asked what a character would feel, based on his or her knowledge of the character’s belief. Quality of belief-desire reasoning was predicted from attachment security in infancy: 82% of babies classified as secure at 12 months with mother passed the belief-desire reasoning task at 5½ years (Fonagy, Steele, Steele, & Holder, 1997). 46% of those who had been classified as insecure failed. Infant-father attachment (at 18 months) also predicted the child’s performance.

It should be noted that not all studies find this relationship and it is more likely to be observed for emotion understanding than ToM (Meins et al., 2002; Oppenheim, Koren-Karie, Etzion-Carasso, & Sagi-Schwartz, 2005, April; Raikes & Thompson, 2006). The lack of consistency in findings between observational measures of attachment and mentalization suggest that the causal pathway is unlikely to be direct; rather, secure attachment and mentalization may have shared facilitating influence in aspects of parenting. The strongest evidence for this comes from observations that the inclination of mothers to take a psychological perspective on their child, including maternal mind-mindedness and reflective function in interacting with or describing their infants, is associated with both secure attachment and mentalization (Fonagy & Target, 1997; Meins et al., 2003; Meins et al., 2002; Peterson & Slaughter, 2003; Sharp, Fonagy, & Goodyer, 2006; Slade, 2005).

Tolerating negative affect may be a shared characteristic of secure attachment and a family environment facilitating mentalizing. For example, family-wide talk about negative emotions, often precipitated by the child's own emotions, predicts later success on tests of emotion understanding (Dunn & Brown, 2001). The capacity to reflect on intense emotion is a marker of secure attachment (Sroufe, 1996). Similar considerations may explain the finding that the number of references to thoughts and beliefs and the relationship specificity of children's real-life accounts of negative emotions correlate with early ToM acquisition (false belief performance) (Hughes & Dunn, 2002).

Three programs of work, by Elizabeth Meins (Meins, Ferryhough, Fradley, & Tuckey, 2001), David Oppenheim (Koren-Karie, Oppenheim, Dolev, Sher, & Etzion-Carasso, 2002; Oppenheim & Koren-Karie, 2002) and Arietta Slade and their

respective groups (Grienenberger, Kelly, & Slade, 2005; Schechter et al., 2005; Slade, 2005; Slade, Grienenberger, Bernbach, Levy, & Locker, 2005) have sought to link parental mentalization to the development of affect regulation and secure attachment by examining interactional narratives between parents and children (for a more comprehensive account of these and other investigations of the impact of the parent's capacity to treat the child as a psychological agent on emotional development see review by Sharp & Fonagy, submitted). Meins and colleagues assessed mentalizing from mothers' verbalisations to a 6-months old infant. For example, mothers were asked, "Can you describe [child] for me?" and their responses were categorized as being mental, behavioural, physical, or general (Meins & Fernyhough, 1999). This is an 'off-line' (non-interactive) measure of mentalizing, but Meins and colleagues also developed a more on-line measure based on twenty minutes of free play coded for appropriate mind-related comments. Mind-related comments were shown to be predictive of attachment security at 6 months (Meins et al., 2001), mentalizing capacity at 45 and 48 months (Meins et al., 2002), and Stream of Consciousness performance at 55 months (Meins et al., 2003). In the Oppenheim et al. studies, the mothers commented on their own previously recorded playful interactions with their child. Both studies found that high levels of mentalization related to the child in the mothers' narratives were associated with secure infant-mother attachment.

Mentalizing of the child in the context of the mother-child relationship, rather than global sensitivity, predicted security of attachment. Yet the studies assessed mothers' mentalization differently: Meins assessed the quality of the parents' thinking about the child in real time in the course of an interaction, whereas the Oppenheim studies employed a more reflective, off-line measure.

Slade and colleagues (Slade et al., 2005) extended previous observations by using an autobiographical memory measure, the Parent Development Interview (PDI), rather than an episode of observed interaction. These researchers found strong relationships between attachment in the infant and the quality of the parent's mentalizing (reflective function – RF) about the child. The PDI aggregates mentalizing across many episodes of interaction, yielding a prototype from the mother's autobiographical memory (Conway, 1996). In a structural model of autobiographical memory, Conway (1992) proposed that two types of autobiographical memories exist within a hierarchical autobiographical memory system: unique, specific events and repeated, general memories. The PDI accesses general memories assumed to have a preferred level of entry to the autobiographical memory system (Addis, McIntosh, Moscovitch, Crawley, & McAndrews, 2004). The PDI permits scrutiny of the mother's off-line reflective mentalizing capacity and her predominant stance towards the child as an intentional being, perhaps reflecting many hundreds of interactions. High scorers on the PDI mentalizing scale are aware of the characteristics of the mental functioning in their infants and grasp the complex interplay between their own mental states and the child's putative inner experience.

Slade and colleagues' (Slade et al., 2005) study included ten infants with disorganized attachment whose mothers' mentalizing scores are a standard deviation below those who are secure. What might low mentalizing parents do to disorganise the infant's attachment classification? Grienberger et al. (Grienberger et al., 2005) rated the Strange Situations collected as part of the study on Karlen Lyons-Ruth and colleagues' AMBIANCE (Atypical Maternal Behavior Instrument for Assessment and Classification, Bronfman, Parsons, & Lyons-Ruth, 1999) coding system. AMBIANCE arguably measures the mother's responsiveness to the intentions

conveyed in the infant's communication and the frequency of atypical behaviours suggesting gross failures to grasp the intentionality of the infant. Their narratives for example, showed little appreciation that the infant's mind cannot be directly read, or depict her as having no feelings, thoughts or wishes. The study demonstrated that mothers with low RF scores were higher on codings for behaviors associated with attachment disorganisation (demanding a show of affection from the infant, fearful behaviour or intrusive or negative behaviours such as mocking or criticizing) than parents with high scores with secure children. The strong correlation suggests that the same control mechanism may be responsible for the inhibitory regulation of certain aspects of the mother's behavior with the infant, and her organisation of narratives about her. A common brain mechanism might subserve both tasks. For example the paracingulate area might provide input for the organisation of both social interaction and person-centered autobiographical narrative (den Ouden et al., 2005; Farrant et al., 2005; Gallagher et al., 2000).

Through this research, Slade and her colleagues might have partially closed the 'so-called' transmission gap between parent and infant attachment that Marinus van IJzendoorn (1995) identified over a decade ago. As we currently formulate it, the mother's secure attachment history permits and enhances her capacity to explore her own mind and promotes a similar enquiring stance towards the mental state of the new human being who has just joined her social world. This stance of open, respectful enquiry makes use of her awareness of her own mental state to understand her infant, but not to a point where her understanding would obscure a genuine awareness of her child as an independent being. The awareness of the infant in turn reduces the frequency of behaviours that would undermine the infant's natural progression towards evolving its own sense of mental self through the dialectic of her interactions

with the mother. The work of Goldberg and colleagues (Goldberg, Benoit, Blokland, & Madigan, 2003) indeed shows that atypical maternal behaviour, as coded on the AMBIANCE system, did relate not only to infant disorganization of attachment but also to unresolved (disorganized) attachment status in the mother's AAI. Thus, while secure mother-infant attachment may not directly facilitate the development of mentalisation, it is an indicator of a parental stance to the child's state of mind that may have a direct facilitative effect. Perhaps more crucially, secure infant attachment evidences that aspects of parental behaviour that might have undermined mentalisation are unlikely to be present. Attachment is a generally non-competitive relationship in which the aim is not to outsmart others, thus learning about minds can be safely practiced. Conversely, severe neglect—the absence of mentalizing on the part of attachment figures—may undermine the development of mentalizing (see below). Preliminary evidence that the capacity for change in attachment organisation decreases over time, underlines the danger that persistent trauma will lead to long-term disorganisation of attachment, with attendant poor development of social cognition and of substantially raised risks of psychopathology (Kobak, Cassidy, Lyons-Ruth, & Ziv, 2006). However, we are not suggesting that parental mind-mindedness is inevitably helpful for the children's emotional development. Mind-mindedness is likely to be one of those parental attributes that is most adaptive in moderation. While evidence on this issue is still lacking, on the basis of our clinical observations we have proposed that maladaptive aspects of parental mentalizing of a child can be either deficient (concrete & stimulus-bound) or excessive or hypermentalizing (necessarily going beyond the data, often quite distorted and sometimes paranoid) (Fearon et al., 2006; Williams, Fonagy et al., 2006). In research considered above the measure of mind-mindedness was confounded with the accuracy

in the scoring; low scorers could be either deficient or excessive mentalizers because both would be rated as failing to reflect the child's mental state with what we may refer to as 'grounded imagination' (Allen, 2006).

### *Deprivation of verbal input*

The association of attachment and mentalisation has directed our attention to the possibility that the key to understanding the impact of social conditions on the development of mentalisation may be through deprivation of normally available catalysts to its development rather than facilitation by unusually sensitive or caring parenting. This kind of formulation would follow from considering mentalisation as developing along lines analogous to grammar acquisition, following a predetermined sequence and chronology as we can observe in the spontaneous language production of sign language in deaf children unfolding at a predetermined rate regardless of the language environment (Petitto, Holowka, Sergio, & Ostry, 2001; Petitto & Marentette, 1991). However even in such sequences there are critical periods during which exposure to language appears to be vital (Curtiss, 1977; Grimshaw, Adelstein, Bryden, & MacKinnon, 1998). Deaf infants of hearing parents are delayed relative to hearing infants (de Villiers & de Villiers, 2000) and the infants of natively signing parents who appear to follow the same early timetable as hearing infants in their acquisition of spoken language (Bonvillian, 1999).

The acquisition of mentalizing could run along lines parallel to the acquisition of the grammar of language independent of variations of social experience, with the exception of critical inputs from conversations about mental states (Leslie et al., 2004). Mentalisation may be the manifestation of a mental state reasoning module which emerges through the development of a 'selection processor' mechanism that enables the child to inhibit the assumption that beliefs correspond to reality in favour

of the understanding that beliefs may be either accurate or inaccurate. However, mentalisation appears to need considerable exposure to conversational and social opportunities if the child is to display ToM reasoning skills on normative timetable (Siegal & Patterson, in press). Nicaraguan deaf adults who grew up devoid of referents to cognitive mental states appear to be incapable of passing false belief tests (Pyers, 2003 cited in Siegal & Patterson, in press). Late signing deaf children of hearing parents are delayed not only in language but also in ToM developments (de Villiers & de Villiers, 2000; de Villiers & Pyers, 2002; Peterson, 2004). Exposure to mental state language may then be critical. Where the actual verb used in the ToM test question implies that the actor has a false belief (this can be the case in Mandarin Chinese, Greek, Turkish, Puerto Rican Spanish) the acquisition of ToM appears to be facilitated (Lee, Olson, & Torrance, 1999; Maridaki-Kassotaki, Lewis, & Freeman, 2003; Shatz & al, 2003).

Language is not just facilitative of ToM; it has been seen as essential to its development. Mentalizing and language could be linked in several ways (see Lohmann & Tomasello, 2003). First, language could be just one form of evidence helping the child to construct and support his “theory” of beliefs (e.g. Bartsch, 2002). Second, the language of mental states specifically may teach the child about the characteristics and processes of mentalizing; through learning the meaning of mental state terms they acquire knowledge of its workings (e.g. Olson, 1988). Third, conversation with others constantly reminds us that they have desires, beliefs and intentions that are the same or different from ours, and this learning will take place in conversation whether mental state terms are used explicitly or not (e.g. Harris, 1996; Harris, 2005). Fourth, the specific syntactical structure of complementation (a complement of the sentence can be false yet the whole sentence is true – e.g. “Ian

believes psychosurgery is effective”) may enable children to hold in mind an internal state that does not correspond to reality (de Villiers, 2005; de Villiers & de Villiers, 2000). Finally, it is possible that language and mentalization are inseparable as they are different aspects of the same thing (Carpendale & Lewis, 2006). Each of these models are consistent with the empirically solid observation that language development is predictive of mentalisation (at least as measured by the false belief task) (e.g. Astington & Jenkins, 1999; Ruffman, Slade, Rowlandson, Rumsey, & Garnham, 2003).

Each of these formulations implies that deprivation of appropriate language input will undermine mentalisation competence. Mentalizing in general, and representing false beliefs in particular, requires mental state vocabulary and grammar of embedded tensed complements to construct propositions about mental states. The vocabulary necessary to talk about the beliefs of others and the syntactic features that enable the child to conceptualise propositional attitudes may be necessary to enable children to entertain propositions that involve the simultaneous representation of alternative states of affairs (false propositions in true statements). Better language may help children converse better with others and through conversation gain insights into other people’s minds which in turn will help refine their linguistic competence, leading to a better conceptual understanding of vocabulary and syntax (de Villiers, 2005; Nelson, 2004) in a virtuous cycle of social improvement. However, emotion understanding and reasoning about beliefs may not follow identical paths in this regard. Late signing deaf children from hearing families spontaneously include terms for inner states of desire and, affect and perception ahead of terms for cognition (Peterson & Slaughter, 2006). However, they are delayed in false belief tests, even into adolescence (e.g. Morgan & Kegl, 2006).

Consistent with the studies on early theory of mind reasoning reviewed above, it may be argued that information related to people as repositories of beliefs is available implicitly from language based interactions between parent and child for both normally developing children and deaf children with deaf parents. Children who are deaf and who live with hearing parents may be deprived of social interaction with an adult that they can fully engage with. Deaf children with hearing parents may experience frequent mismatch, with parents attempting to communicate to their deaf infants through speech (Woolfe, Want, & Siegal, 2002). Thus what may be most important is reflective discourse in which there is frequent accurate elaboration of psychological themes (Peterson & Slaughter, 2003), in which adults and children engage in and respond to the intentional states implied by each others' reasonable verbal comments linked to each others' appropriately interpreted actions. Conversation that was non-reflective, that did not use language to link accurately internal states to actions of the self, the other, or explicate reactions to others' actions, would be of little value in facilitating the development of metacognition, at least for young children. By and large, under normal circumstances it is fair to say that sensitive, emotionally normally modulated conversational exchange may be the "royal road" to understanding minds (Dunn, 1996; Harris, 2005; Nelson, 2005). Mentalizing verbal rationalizations of actions teach the child to acquire the intentional stance – rather than teleological externalist – explanatory schemes to rationalize and predict the behaviour of others. This may contribute to the facilitation of performance on explicit verbal mindreading test. The appropriate mutually attuned engagement of adult and child are crucial to this.

### *Deprivation of expectable social input: blindness and maltreatment*

There are data from two populations of children that are consistent with the above assumption: blind children and maltreated children. We shall consider them briefly in turn. Selma Fraiberg (1977) noted that social and language development of blind children was delayed and distorted. Blind infants do not initiate much contact with parents by voice, touch or posture, and are limited in their expression of affect (Tröster & Brambring, 1992). Peter Hobson has explored in a most illuminating way the social and developmental implications of this sensory disability (Hobson, 2002; Hobson & Bishop, 2003; Hobson, Lee, & Brown, 1999). Blind infants miss out on access to parental non-verbal information on inner states (affect, attention, perception, intention). They refer to themselves as 'YOU' while in the speaker's role (Andersen, Dunlea, & Kekelis, 1984; Dunlea, 1989; Fraiberg, 1977) and Hobson has linked this characteristic deficit with perspective taking problems as well as visual deprivation of social cues, especially facial expression. The significant retardation of language skills in blind children (Andersen et al., 1984; Dunlea, 1989; Fraiberg, 1977) may be linked to problems in the mastery of syntactic rules, deprivation of some of the pragmatics of conversation (e.g. turn-taking, the more limited presence of a conversational partner).

These sensory, social and language limitations create a situation where the kind of mutually attuned, shared reflective discourse we suggest is facilitative of mentalization occurs with reduced frequency. Perhaps associated with this, blind children's pretend play is quite limited until early school age (Fraiberg, 1977; Tröster & Bambring, 1994) and they understand pretend play poorly (Hughes et al., 1998; Lewis et al., 2000). On false belief tests blind children appear to be delayed (Green et al., 2004; Minter, Hobson, & Bishop, 1998) and to require a substantially higher verbal mental age to pass (11 vs 5) (McAlpine & Moore, 1985). A relatively large

carefully conducted study of 23 blind children aged 5-13 found that less than a quarter of the sample could show competence in this task by passing 4 false belief tests. The performance of younger blind children was around chance (Peterson et al., 2000).

The second group of children for whom deprivation of expectable attuned social input can be argued to cause a distortion and deficit in mentalisation is abused or maltreated children. The following findings are relatively well established in relation to young maltreated children and could be linked with problems of mentalisation: (1) they engage in less symbolic and dyadic play (Alessandri, 1991). (2) They sometimes fail to show typical empathic responses to distress in other children (Howes & Espinosa, 1985; Klimes-Dougan, 1990; Main & George, 1985). (3) They more often manifest emotionally dysregulated behaviour (e.g., Maughan, 2002). (4) They make proportionately fewer references to internal states and maltreating mother-child dyads discuss emotions less frequently than non-maltreating dyads (Beeghly & Cicchetti, 1994; Shipman & Zeman, 1999). (5) They manifest a range of problems indicative of a difficulty in understanding emotional expressions. (a) Between three and seven years of age they appear to have poorer understanding of universal child facial expressions of emotion (Camras, Grow, & Ribordy, 1983), masked negative emotional facial expressions (Camras et al., 1988a), adult facial expression (During & McMahon, 1991), even when controlled for verbal IQ (Camras et al., 1990); (b) Findings on the whole suggest a delay in development of emotion understanding as there is no evidence for affect specific deviation (Camras et al., 1988b; Smith & Walden, 1999) except for the tendency by maltreated children to misattribute anger (Camras, Sachs-Alter, & Ribordy, 1996) and show elevated event-related potential (ERP) to angry faces by maltreated 6-12 years olds (Pollak, Klorman, Thatcher, & Cicchetti, 2001; Pollak & Sinha, 2003). It should be noted that there is

no evidence yet linking a difficulty in understanding affect to the abnormal social behavior of maltreated children.

While the evidence for significant developmental delay in the emotion understanding of maltreated young children is consistent (Frodi & Smetana, 1984; Pears & Fisher, 2005 ; Rogosch, Cicchetti, & Aber, 1995; Smith & Walden, 1999) it should be pointed out that the observed difference tends to be reduced and sometimes eliminated if the groups are carefully matched intellectually and socioeconomically (Frodi & Smetana, 1984; Smith & Walden, 1999). However, there is meaningful evidence indicating that the extent of the delay in emotion focused mentalization is of developmental significance for maltreated children. For example, the quality of understanding of the possible situational determinants of sad and angry emotions at approximately six years of age was found to predict social competence at eight years of age (Rogosch et al., 1995). Consistent with this finding, the experience of physical abuse was found to predict social isolation at eight years of age to the extent that it had impacted on emotion understanding (controlling for verbal ability). There have also been reports of delayed theory-of-mind understanding in maltreated children (Cicchetti, Rogosch, Maughan, Toth, & Bruce, 2003; Pears & Fisher, 2005 ) but the studies leave open the question, whether the deficits experienced by maltreated children result from maltreatment per se or whether they were a function of the broader intellectual delays experienced by many maltreated children. A persuasive study from the Rochester Mount Hope Family Center (Cicchetti et al., 2003) tested 203 maltreated low SES children and 104 non-maltreated controls using the unexpected content false-belief task. Children's language was assessed, and verbal mental age (VMA) was estimated based on Peabody Picture Vocabulary Test. There was a highly significant effect of maltreatment on ToM ( $p < .001$ ). Controlling for CA,

SES reduced the effect but it remained statistically significant when potential confounding variables were controlled for (Cicchetti et al., 2003). In a study of 80 maltreated pre-schoolers, Macfie, Toth and Rogosch et al. (Macfie et al., 1999) showed clear limitations in the representation of social cognition in a story stem completion task where the story stem called for the relief of distress.

In a further study (Macfie, Cicchetti, & Toth, 2001), maltreated children, especially physically or sexually abused children, were shown to manifest more dissociation, disruptions of identity and incoherence of parental representations which may be seen as indicators of a failure of mentalizing capacities (Fonagy, 1997). In maltreated children the capacity for social cognition, particularly the complexity of the representation of the parent in conflict imbued settings, decreases with development while the children's representations of themselves become increasingly simplified and exaggerated (Toth, Cicchetti, Macfie, Maughan, & Vanmeenen, 2000).

Perhaps the most compelling evidence for maltreatment-related social-cognitive deficit rooted in the child-caregiver relationship is provided by the elegant randomised controlled intervention study reported by Toth et al. (Toth, Maughan, Manly, Spagnola, & Cicchetti, 2002). Pre-schooler parent psychotherapy was offered to 23 maltreating families. This 12-month intervention aimed at elaborating and modifying the relationship between parent and child by linking current maternal conceptualisation of relationships to the mother's childhood caregiving response. Social cognitive measures of outcome favoured this group in contrast to psychoeducational home visitation or treatment as usual in a range of domains including degree of maladaptiveness of maternal representation and the quality of self and mother-child relationship representation. In all, recent work from the Mount

Hope Family Center is consistent with the assumption of serious impairment of social cognition associated with maltreatment and the potential for reducing this impairment through a relationship focused intervention.

In a recent, as yet unpublished study we used Baron-Cohen's Reading the Mind in the Eyes test to test mentalization in 147 adults (Fonagy et al., 2006). Half the participants were individuals with a history of childhood maltreatment, many of whom showed present Axis I (69.7%) or Axis II (45.3%) psychopathology. The other half were community controls. All participants were administered the Childhood Experience of Care and Abuse (CECA) and SCID-I interviews. Results showed that participants with a history of maltreatment achieved lower scores on the Eyes Test but this varied with the severity and type of maltreatment. A history of moderate sexual or physical abuse was most strongly associated with low scores on the test. It should be noted, however, that the strength of association with Eyes Test scores was far stronger with adolescent than with childhood maltreatment about which records were likely to be less accurate.

These data are consistent with the assumption that child-adult interaction characterised by maltreatment deprives children of the attuned mutual engagement focused around internal states that may be critical for the development of mentalization leading to developmentally critical delay. It should be noted that this is neither the only possible or even possibly the most likely account of the findings. Mentalization deficit associated with maltreatment may not necessarily reflect incapacity but rather a form of decoupling, inhibition or even a phobic reaction to mentalizing in maltreated individuals. There are multiple possibilities: (1) We have suggested elsewhere that the reluctance to conceive of mental states on the part of

maltreated individuals might be understandable given the frankly hostile and malevolent thoughts and feelings which the abuser must realistically hold to explain his or her actions against a vulnerable young person (e.g. Fonagy, 1991). Consistent with this assumption, forms of maltreatment that are most clearly malevolent and clearly target the child have greatest impact on mentalization (viz. physical, sexual and psychological abuse). (2) It could be argued that adversity undermines cognitive development in general (Cicchetti & Lynch, 1995; Cicchetti & Toth, 1995). Certainly, there is strong evidence to suggest that addressing issues of maltreatment in parent-child relationships can facilitate the children's cognitive development (Cicchetti, Rogosch, & Toth, 2000). In a number of studies reviewed here, controlling for verbal IQ reduced the apparent impact of maltreatment. However, this is not the case for all studies and in our study we observed no significant contribution of verbal IQ to performance on the Eyes Test. (3) Another alternative may be that mentalization problems reflect an anomaly related to arousal. Maltreatment entails exposure to chronic stress (see Cicchetti & Walker, 2001). This can lead to the ready activation of the arousal system that underpins the posterior cortical and sub-cortical functions (Arnsten, 1998; Arnsten et al., 1999). It is possible that experimental tasks testing affect and emotion understanding, and theory of mind, increase the arousal of maltreated individuals, resulting in the less efficient functioning of the prefrontal regions of the brain of maltreated individuals with maltreatment histories.

Here we favour the developmental account because it encompasses findings on other types of social influences. There is ample evidence that maltreatment puts children at risk of profound deficits in the skills required to negotiate social interactions with peers and friends. These are broad ranging and include verbal ability, the comprehension of emotional stimuli and situations, and possibly also

theory of mind. We have seen that the level of mental state understanding (particularly emotion understanding) is closely linked to the extent that emotions are openly discussed in the mother-child dyad or can be discussed given the child's disabilities and the parents' ability to overcome these. We may then argue, that maltreatment acts on mentalisation in many ways like sensory deficits; it compromises the unconstrained, open reflective communication between parent and child or indeed between child and child. Maltreatment undermines the parent's credibility in linking internal states and actions. This limitation in communication is not hard to comprehend and could hardly be otherwise if the maltreatment is perpetrated by a family member. But even in cases where it is not, the centrality of the maltreatment experience for the child coupled with the oversight on the part of the parent of an experience of maltreatment which the child encounters outside the home could serve to invalidate the child's communications with the parent concerning the child's subjective state. Thus apparently reflective discourse will not correspond to the core of the child's subjective experiences, and this moderates or reduces the facilitative effect of mentalizing verbal rationalizations of actions in generating an intentional as opposed to a teleological orientation. The formulations advanced here imply that therapeutic interventions should aim to engage maltreated children in causally coherent psychological discourse within appropriate contexts.

### *Limitations of relationship influences on mentalization*

Relationship influences on the development of mentalization are probably limited and specific rather than broad and unqualified. Three key limitations to simplistic linking of mentalization and positive relationship quality should be kept in mind (Hughes & Leekham, 2004): (1) The application of mentalization and the acquisition of theory of mind skills should be treated as separate functions. The

acquisition of the capacity to mentalize may, for example, open the door to more malicious teasing (e.g. Dunn, 1988), increase the individual's sensitivity to relational aggression (Cutting & Dunn, 2002), or even mean that they take a lead in bullying others (Sutton, Smith, & Swettenham, 1999a, 1999b). The possession of the capacity to mentalize is neither a guarantee that it will be used to serve pro-social ends, nor a guarantee of protection from malign interpersonal influence. (2) While, as we have seen, broadly, positive emotion promotes the emergence of mentalization (Dunn, 1999), negative emotion could be an equally powerful facilitator. For example, children engage in deception that is indicative of mentalizing in emotionally charged conflict situations (Newton, Reddy, & Bull, 2000). (3) The impact of relationships on the development of mentalization is probably highly complex involving numerous aspects of relational influences (e.g. quality of language of mental states, quality of emotional interaction, themes of discourse, amount of shared pretend play, negotiations of conflict, humour in the family, discourse with peers, etc) probably affecting several components of the mentalizing function (joint attention, understanding of affect states, capacity for emotion regulation, language competence, competence with specific grammatical structures such as sentential complements, etc.) (Hughes & Leekham, 2004).

### **The development of an agentive self: a model of social influences on social cognition**

There is reasonable evidence to suggest that children's caregiving environments play a key role in the development of some of their social cognitive capacities. The suggestions beg major questions of mechanisms. We have considered two models so far: (1) the interaction of attachment and social cognition systems at

the level of brain function and (2) facilitative attuned conversational engagement between parent and child focused on linking actions to internal states. While consistent with correlational data on the environmental predictors of delay in the acquisition of mentalization, the suggestions fall short of identifying a specific set of biobehavioural mechanisms that could mediate this developmental process. A further currently speculative model of social influence links this to the acquisition of cultural knowledge.

### *The pedagogical stance*

Recently, a conceptual framework has been advanced that can provide a more parsimonious account of much of the behavioural data reviewed including findings that have often been used to support the intersubjectivist position (see above) while also assigning a role to relational influences in the development of social cognition. We shall briefly describe this conceptual framework and then apply it to the emergence of emotion understanding and affect regulation. It is almost axiomatic that the evolutionary underpinnings of human culture require that the infant turns to others for essential information about the world (Csibra & Gergely, 2006; Gergely & Csibra, 2005b). The child naturally turns to the caregiver to provide him with information about the nature of the world, internal and external. We assume that attuned caregivers behave towards and speak to children in such ways that they gradually conclude that their behavior may be best understood if they assume that they have feelings, wishes, ideas and beliefs, which determine their actions, and the reactions of others can be generalized to other similar beings. The caregiver is biologically prepared to act in the role of the ‘teacher’, the pedagogue. Pedagogy is hypothesized to be a primary cognitive system with a collaborative design that has evolved to facilitate the efficient transmission of relevant cultural information from

knowledgeable people to ignorant, but specifically receptive human babies (Csibra & Gergely, 2006; Gergely & Csibra, 2006).

Several theorists have pointed to the importance of teaching in the ontogenesis of human cognition (e.g. Barnett, 1973; Caro & Hauser, 1992; Kruger & Tomasello, 1996; Premack, 1984; Premack & Premack, 2003; Tomasello, 1999; Tomasello, Kruger, & Ratner, 1993). However, in these models teaching is usually described as a secondary derivative of some more fundamental human-specific adaptation, such as language (Dunbar, 1996), theory of mind (Tomasello, 1999), aesthetics (Premack & Premack, 2003) or culture itself (Bruner, 1993; Vygotsky, 1978). In contrast, Csibra and Gergely believe that the ability to teach and to learn from teaching is a primary, independent, and possibly phylogenetically an even earlier adaptation than either language or the ability to attribute mental states.

The social world and human culture represent two somewhat incompatible challenges for human infants. As this review attempts to detail, children have to develop an understanding of minds to be able to predict and interpret people's actions in terms of causal mental states attributed to them and as part of this process acquire an agentic sense of self. On the other hand, they must acquire an immense amount of cultural knowledge much of which is not obvious as it involves arbitrary features of social belief systems and complex cultural artifacts that each member of the community has to learn to use. Theories of self and social cognitive development have tended to focus on the first challenge with many elaborate models to explain the ontogenetic development of young children's mind-reading skills and the unfolding of explicit representational understanding of minds. The theory of human pedagogy provides a new perspective for approaching both of the major challenges of social-

cultural development outlined above as well as offering an alternative, additional explanation of the social influences on the development of mentalising capacities.

The theory of human pedagogy (Csibra & Gergely, 2005; Gergely & Csibra, 2005b; Gergely & Csibra, 2006) proposes that early emerging triadic communications about referent objects are often best conceived of as serving a primarily epistemic function. Triadic communications enable the child to obtain reliable, new and relevant information from knowledgeable adults about the generalizable properties of referent objects that constitute universally shared cultural knowledge to be fast-learned by infants (such as the object's name, proper function, manner of use, whether it is good or bad).

Adults produce two types of pedagogical communicative cues for which infants show specific receptivity: cues of 'ostensive communication' and cues of 'referential knowledge manifestation'. The teacher must not only transmit her knowledge to the learner, but also alert him to the fact that she is teaching. This requirement is analogous to the Gricean view of ostensive communication, which holds that normal human communication makes manifest not just the intended message content but also the communicative intent of the speaker. Gergely & Csibra call this aspect of pedagogy ostension, after Sperber and Wilson (Sperber & Wilson, 1986). 'Ostensive cues' of communication have two major functions: a) they tell the infant that the adult has an overt 'communicative intent' (cf. Sperber & Wilson, 1986), and b) they function as 'addressing cues' telling the infant that the communication is specifically addressed to her. Ostensive cues involve the establishment of eye-contact typically marked by further ostensive gestures such as 'knowingly' raising one's eyebrows, momentarily widening (or shrinking) one's eyes,

and tilting one's head slightly forward towards the infant. These are often accompanied by calling the infant by name using the salient and specific type of 'marked' speech intonation pattern of 'motherese'. Further ostensive cues include turn-taking and contingent reactivity (see Csibra & Gergely, 2005 for a review of evidence of very early sensitivity and preference for such cues by human infants). We assume that ostensive cues constrain and direct infants' interpretation of adults' object-directed actions (such as their object-referential emotion expressions, verbal labelling, demonstrations of the functional properties of objects) as conveying to them new and relevant knowledge about the referent that they need to extract and bind to its representation as its essential property.

In this theoretical framework the phenomenon of early turn-taking 'protoconversational' interactions is interpreted as manifesting the infant's innate sensitivity to and preference for stimuli exhibiting 'contingent reactivity'. This innate propensity to engage in turn-taking contingencies (Floccia, Christophe, & Bertoncini, 1997) together with the infant's innate preference for eye-contact (Farroni, Csibra, Simion, & Johnson, 2002; Farroni et al., 2004), orientation towards face-like configuration, (Cassia, Turati, & Simion, 2004; Turati, Simion, Milini, & Umiltà, 2002) and preference for the characteristic intonation pattern of infant-directed speech or 'motherese' (Cooper & Aslin, 1990; Fernald, 1985) belong to the set of cues of 'ostensive communication' that function to identify a potential teacher for the infant. Ostensive cues are interpreted as signaling the presence of a communicative intention in the other that is 'addressed' to the infant. Ostensive cues are assumed to trigger a specific receptive attentional and interpretive attitude, the 'pedagogical stance', in the infant. Thus, early turn-taking - together with other early social communicative phenomena such as joint attention, 'proto-declarative' pointing, social referencing

(Egyed, Király, & Gergely, 2004), or imitative learning (Gergely, Bekkering, & Kiraly, 2002; Gergely & Csibra, 2006; Király, Csibra, & Gergely, 2004) are interpreted as examples of pedagogical communication whose primary function is epistemic in nature; its aim is to facilitate fast and efficient transfer of knowledge about the world, rather than that of intersubjective 'sharing' of internal psychological states.

*Teaching about the self's emotions via the pedagogical stance: the origins of affective self-awareness and self-regulation*

Let us take the development of an understanding of affects as an example. In line with our social constructionist stance, we assume that at first infants are not introspectively aware of their differential emotion states. Babies learn to differentiate the internal patterns of physiological and visceral stimulation that accompany different feelings through observing their caregivers' facial or vocal mirroring responses to these (Gergely & Watson, 1996, 1999; Legerstee & Varghese, 2001; Meltzoff, 1990; Mitchell, 1993; Schneider-Rosen & Cicchetti, 1991). Furthermore, infants establish introspectively accessible second order representations of their primary and procedural emotional states on the basis of the caregiver's consistent and marked mirroring reactions to their automatic emotion-expressive displays (Fonagy et al., 2002; Gergely & Watson, 1996). The baby also comes to associate the control he has over the parents' mirroring displays with the resulting improvement in his emotional state, leading, eventually, to an experience of the self as a regulating agent. The establishment of a second order representation of affect states creates the basis for affect regulation and impulse control: affects can be manipulated and discharged internally as well as through action, they can also be experienced as something recognizable and hence shared. Affect expressions by the parent that are not

contingent on the infant's affect will undermine the appropriate 'labelling' of internal states (i.e. the establishment of introspectively accessible second order representations for them) which may, in turn, remain confusing, experienced as unsymbolized and hard to regulate.

Two conditions need to be met if the capacity to understand and regulate emotion is to develop: (a) reasonable congruency of mirroring whereby the caregiver accurately matches the infant's mental state and (b) 'markedness' of the mirroring, whereby the caregiver is able to express an affect while indicating that she is not expressing her own feelings (Gergely & Watson, 1996, 1999). For affect mirroring to serve as the basis of the development of a representational framework, the parent must indicate that her display is not an indication of how she herself feels. 'Marked' affect-mirroring interactions can be interpreted as a special case of pedagogical communication that functions to teach infants about their primary emotions through establishing cognitively accessible second-order representations for their – initially non-conscious – procedural (automatic) emotion states. This proposal is based on the realization that 'marked' affect-mirroring displays involve the same infant-directed cues of 'ostensive communication' and 'referential knowledge manifestation' that play a key role in the hypothesized species-specific cognitive adaptation for human pedagogy (Csibra & Gergely, 2006; Gergely & Csibra, 2006).

By activating the infant's pedagogical stance, repeated experience with ostensive 'marked' affect-mirroring feedback reactions from infant-attuned caregivers in the baby's early attachment environment can a) 'teach' the infant about the existence of her internal subjective emotion states, b) lead to the internalization of the caregiver's 'marked' mirroring displays as second-order representations associated with the infant's (inferred) primary self states, and c) introspectively sensitize the

infant's attentional system to the presence of internal referents in the self (through the process of 'social biofeedback'). This process, we argue, extends mentalisation to include in its domain internal states of the self.

The characteristic formal features of empathic 'marked' affect-mirroring displays share the characteristic features of other types of ostensively communicated referential knowledge manifestations in that "marked" affect displays are themselves salient and schematically executed transformations of the corresponding normative, realistic emotion expressions. Also, 'marked' emotion displays are typically accompanied by ostensive cues of communicative intent such as eye-contact, raised eyebrows, slightly tilted head or gestural widening (or shrinking) of the eyes. The 'marked' form of the caregiver's emotion display and the other ostensive cues accompanying it tell the infant that it is not performed in its primary function as expressing the caregiver's actual emotion state. We hypothesise that ostensive cues accompanying the caregiver's affect-mirroring induce the referential interpretive attitude of the 'pedagogical stance' in the infant and activate a search for the intended referent. In trying to work out what the 'marked' emotion display refers to (since as a result of its 'markedness' the emotion expressed is 'decoupled' from the caregiver as not expressing her own emotion state), the infant will rely on the cues of referent identification (such as eye-gaze direction) of the caregiver that accompany her communicative emotion display. Since the caregiver is looking at and being oriented towards the infant while producing these infant-directed 'marked' emotion mirroring displays, the infant's attention will be directed towards her own face and body, i.e., her own physical self as the spatial locus of the referent entity that the caregiver's attention orienting referent identification cues indicate and to which the "marked" (and 'decoupled') affect display should be referentially 'anchored'.

In summary, a ‘mind-minded’ reflective mirroring environment extends mentalization to include the internal states of the self in its domain (thereby making self-prediction and emotional-self-control possible) by applying pedagogical referential communication to the domain of the emotional and dispositional/intentional states of the self. Thus awareness, cognitive access and subjective internal self-states become part of this now extended domain of mentalization. Two critical changes in processing are achieved: (a) second-order representations of internal self-states are created and, (b) the attention system may be socialised towards an introspective monitoring direction. Parenthetically, in relation to the latter effect, we should note that ‘effortful control’ (the ability to inhibit a dominant response to perform a subdominant response Posner & Rothbart, 2000) has been linked to the quality of parent-infant relationship both theoretically (Fonagy, 2001) and empirically (Fearon & Belsky, 2004; Mundy & Neal, 2001). Both these changes are specific to humans and do not follow from the evolutionary need to predict others in competitive niches (that many other species have), but from teaching the child about (otherwise) un-learnable culturally relevant knowledge.)

*Evidence linking contingent interactions to the development of well regulated affect and mentalisation*

What of parents who are incongruent in their mirroring of internal states and are unable to mark for the infant that their mirroring is of the infant’s and not the caregiver’s state of mind? An expression congruent with the baby’s state, but lacking markedness, may overwhelm the infant. It is felt to be the parent’s own real emotion, perhaps making his experience seem contagious, or universal, and thus more dangerous. Many of the parental behaviours noted in the AMBIANCE codes may be seen as examples of the caregiver being unable to ‘mark’ her mirroring sufficiently to

direct the infant's attention towards her own face and body. In the short term, the baby's perception of a corresponding but realistic negative emotion is likely to escalate rather than regulate his state, leading to cumulative disorganisation rather than containment. If turntaking contingency and markedness are ostensive cues drawing the child's attention to internal processes, then higher levels of this type of experience should facilitate the acquisition of emotion regulation and more generally that of mentalizing.

As part of a larger longitudinal study, we recently examined the developmental relation between contingent maternal mirroring and the "markedness" of contingent maternal reactions in a group of 12-month-old infants, on the one hand, and different aspects of pretence competence of the same children at 2.5 years of age, on the other (Futó, Bátki, Koós, Fonagy, & Gergely, 2004). Maternal mirroring and contingent "markedness" was measured at 12 months in the so-called three-phase Mirror Interaction Situation (MIS) (see Koós & Gergely, 2001), a modified version of the standard Still-face paradigm (Tronick, Als, Adamson, Wise, & Brazelton, 1978). We have designed this procedure to induce contingent mother-infant interactions in an affect-regulative situation that involved the induction of mild stress in the infant. Mother and infant were seated next to each other in front of a (one-way) mirror. They were separated by an occlusion screen that prevented them from physically contacting each other: however, they were free to interact facially and vocally with each others' mirror image. The interactions were coded for a number of interactive and state-expressive behavioral variables. The category of relevance here was 'contingent maternal reactivity' ("mirroring"). This was defined as facial and/or vocal reflections by the mother of the infant's behaviours or as temporally contingent verbal 'acknowledgements' by the mother if these made reference to the infant's behavioral,

attentional, intentional, or emotional state. We have separately coded contingent verbal reactions when these also involved a saliently “marked”, exaggerated form of expression. The ratings for both contingent responsiveness and markedness could be achieved with high degrees of agreement.

We followed up a sub-sample of these infants selected on the bases of maternal codes for contingency and markedness when the infants were between 2-3 years of age. At this time we administered, in the child’s home, a modified and enriched version of the battery of pretence tasks originally developed by Harris & Kavanaugh (Harris & Kavanaugh, 1993) to measure representational aspects of pretence competence. This follows along the lines of Leslie’s (Leslie, 1987) analysis of the metarepresentational structure and representational operations implied by understanding and producing pretend play. For “markedness”, our preliminary findings (Futó et al., 2004) indicate that high degree of “markedness” of contingent maternal references to infant state (during the phases of the interaction before and after the still-face episode at 12 months) predicted both higher overall representational pretence competence scores and higher scores on spontaneous, adequate and creative extensions in the use of pretence at 2.5 years of age. For “contingency”, we found that high contingent maternal reactivity (“mirroring”) at 12 months predicted high scores on spontaneous, adequate and creative extensions of pretence performance at 2.5 years. (Interestingly, we also found that low tolerance to loss of maternal contingency during the still-face episode of the MIS predicted low pretence performance in open-ended pretence situations involving separation or physical injury at 2.5 years of age). Taken together, these findings support the claim that high levels of contingency between mother and child is associated with at least one key aspect of mentalisation: to represent (and manipulate) mental states with fictional contents.

Stanley, Murray, & Stein (Stanley, Murray, & Stein, 2004) measured contingent maternal reactivity to infant behaviours during face-to-face interactions at 2 months (in a sample of infants with postnatally depressed mothers and non-depressed controls). Months later these infants participated in an Instrumental Learning Task in which they had to learn that their spontaneous responses exerted contingent causal control over an external event (inducing the contingent movements of a mobile). Stanley et al. reported that high contingent maternal reactivity at 2 months predicted faster instrumental learning in infants (irrespective of maternal status with regard to depression). This finding suggested that maternal contingent reactivity to the infant's state expressions increases the accessibility to introspection of internal proprioceptive cues that accompany expressions of the infant's state. This increased sensitivity to and accessibility of proprioceptive cues may have allowed for the faster discovery and learning of the instrumental contingent control that the infant's spontaneous responses exerted over the mobile's movements, leading to more efficient learning among infants with highly contingently reactive mothers. The finding also indicates that the experience of causal agency and self-efficacy gained by those infants whose responses evoked high contingent maternal reactivity may have resulted in a generalized interest in attending to (and active testing of) the potency of their actions in controlling different aspects of the world around them.

In this account, affect regulation is closely related to the developing capacity to mentalize about affects. Being able to think about socially constructed and internalised second order representations of one's own emotional states transforms one's ability for affect regulation. Not only does it allow adjustment of affect states, but more fundamentally it is used to regulate the self. This may be an instance of the general principle that the child's capacity to create a coherent image of mind depends

on an experience of being perceived as a mind by the contingently responsive attachment figure. Social understanding of the subjective self can then be seen as an emergent property of the child's experience of referential interactions with the caregiver, which will inevitably generate the discovery that others have different perceptions, beliefs and feelings about the world from one's own.

### *Understanding the relationship influences on the acquisition of mentalization in the context of the pedagogy theory*

Throughout this review we have argued that the enthusiastic search for early forms of intersubjective understanding of minds had an undesirable additional effect of sometimes too hastily embracing mentalistic interpretations for early social cognitive phenomena (including social referencing, imitative learning, facial and vocal interactions that have a turn-taking 'proto-conversational' structural organization, proto-declarative pointing, or predicting others' object-directed actions (e.g. Moses, Baldwin, Rosicky, & Tidball, 2001; Mumme & Fernald, 2003; Phillips, Wellman, & Spelke, 2002; Sodian & Thoermer, 2004)) at the expense of exploring alternative functional explanations that do not necessarily involve or rely on infants' capacity to attribute mental states (Csibra & Gergely, 2006; Gergely, 2002). We have suggested that phenomena such as the findings concerning proto-conversations with the infant can be accounted for without the need to assume an innate capacity for self-awareness. We need assume only an innate contingency detection mechanism and a biological predisposition to teaching and learning on the part of both caregiver and infant. To paraphrase this, the evolutionary underpinnings of human culture require that the infant turns to others for essential information about the world (Csibra & Gergely, 2006; Gergely & Csibra, 2005b).

The universality assumption of pedagogy suggests that children consider knowledge that they are taught to be shared cultural knowledge available to all others. It follows then that the small child assumes that his knowledge is knowledge held by all. What he knows is known by others and what is taught by others is accessible to all others. When I am taught a new word for a new referent I do not have to learn who else was exposed to this knowledge. I assume that others who did not teach or witness the teaching of the new word will also know the meaning. That is, that the world is shared between all of us and only slowly does the uniqueness of our own perspective differentiate so that a sense of individual mental self can develop.

Thus in relation to what we know and understand about the world we start with the assumption that knowledge is common and there is nothing unique about our own thoughts or feelings. The assumption of universality implies that whatever the child knows (especially if it was taught to him) will be known by everyone. Though this will be a valid inference most of the time, children eventually have to learn the conditions under which this assumption should be suspended to overcome the erroneous conclusions that have recently been dubbed the 'curse of knowledge bias' by Susan Birch and Paul Bloom (Birch & Bloom, 2004). This bias was originally formally described by three economists (Camerer, Lowenstein, & Weber, 1989), and refers to the common observation that if one knows something about the world one tends to assume that everyone else knows it too. So, young children report that other children will know facts that they themselves have just learned (Taylor, Esbensen, & Bennett, 1994). It seems clear and unsurprising that three-year-olds are more likely than older children to assume this (Birch & Bloom, 2003). The curse of knowledge phenomenon accounts for the so-called 'egocentrism' of young children. They cannot appreciate another person's perspective, not because they assume that everyone's

perspective is the same as theirs, but rather because everyone knows the same things. Piaget's concept of egocentrism has exactly the opposite emotional valence to what we suggest is actually taking place. It is not the overvaluing of private knowledge, it is the undifferentiated experience of shared knowledge that hinders perspective taking. Many diverse observations show this (Birch & Bloom, 2003; Fischhoff, 1975; Kelley & Jacoby, 1996; Keysar, Lin, & Barr, 2003; Taylor et al., 1994). We assume that everyone has the same knowledge that we do, because most of the beliefs that we have about the world were someone else's beliefs before we made them our own.

Children do not know fully that they are separate, that their internal world is something private and individual, of which they will eventually take ownership or at least claim privileged access. From this perspective intersubjectivity may be an accurate, if superficial description. They do not know that they can choose whether – for example – to share their thoughts and feelings with their parents, their teacher or their therapist. Perhaps one reason that toddlers are so prone to outbursts of rage and frustration is that as the world and individual minds are not yet clearly demarcated, they expect other people to know what they are thinking and feeling, and to see situations in the same way they do. Thus crossing their intentions seems malign or wilfully obtuse, rather than the result of a different point of view, alternative priorities, etc. That makes it not just hurtful but intolerable and maddening, a denial of what they believe to be a shared reality.

Much of early social-cognitive development establishing a realistic understanding of other minds involves learning about the specific conditions under which the built-in default assumption of universal knowledge and omniscient other minds must be inhibited. This gradual and experience-driven learning process eventually leads to the ability to draw correct inferences and to attribute and represent

the differing knowledge contents of separate and individual minds of others. Previous research on social-cognitive development considered as its central task the need to account for how children come to understand that other people *have* minds. The new theoretical perspective offered by pedagogy theory turns this question upside-down identifying as the central task for children's early development the need to come to understand that others have *separate minds with different knowledge contents*.

Clinically, psychotherapists are daily confronted with the recognition that other people (our patients) really do think in different ways. Some of the apocryphal (and not so apocryphal) stories we hear about conflicts, for example between women and men, may rest on the difficulty that the one cannot conceive of the other not understanding situations as they do, they believe that the other is really only pretending to disagree, not admitting to a shared reality. More commonly, the overwhelming expectation on the part of some of our patients of being "totally" understood must have a developmental root. Finally, the devastation of having not been accurately perceived, the so-called rupture in the therapeutic alliance (Safran & Muran, 1996), has therapeutic potential precisely because it forces therapist and patient beyond the illusion of shared consciousness and creates an opportunity for each to have a "mind of their own" at least in the patient's experience.

Mentalization, we argue, evolves out of this biological predisposition to a shared orientation to the representation of external reality. The pedagogical stance ensures that the child naturally turns to the caregiver to provide him with information about the nature of the world, internal and external. Marked mirroring (ostensive cues) ensure that awareness is inwardly as well as outwardly directed. By building second-order representations on the one hand, and exemplifying mental reasoning schemes to give sense to action on the other, the relationship with the mind-minded

reflective caregiver transforms the implicit and automatic mentalizing competence into an explicit, potentially verbally expressible, and systematized ‘theory of mind’. This formulation can accommodate the evidence on environmental factors found to influence the development of mentalisation (from play activities to parenting effects). Verbal practice and direction towards perspective taking plus mentalizing explanations provided by the attachment figure are again conducive to the development of explicit mentalizing. This eventually leads to the ability to draw correct inferences about the knowledge contents of others’ minds. The caregiver behaves towards the child in such a way that the child’s assumption of universal shared knowledge is mildly challenged and his knowledge of internal states expanded. The mentalizing caregiver can bridge the alternating focus on physical reality and internal state, sufficiently for the child to identify contingencies between them.

Ultimately, the child arrives at the conclusion that the caregiver’s reaction to him makes sense given internal states of belief or desire within himself, which, in the first instance he assumes are known and available to all. With repeated experiences that this assumption is wrong, that his experiences are not shared, the subjective self begins to evolve. Through learning aspects of the caregiver’s knowledge of the world, intentionally taught by a trusted other, the child develops a sense of overlapping knowledge and yet of the uniqueness and separateness of his self-experience.

In brief then, we have seen that the ability to monitor others’ perceptual access to reality, represent their mental representations of perceived reality and predict their behaviour on the basis of such perceptually induced beliefs develops quite early (Luo & Baillargeon, 2005; Onishi & Baillargeon, 2005; Southgate et al., in press-a). It may even be innate and – in a probably more restricted domain-specific sense of types of

contents monitored for and represented (e.g., food, territory, sex) – it is probably not even human-specific (Bugnyar & Heinrich, 2005; Emery & Clayton, 2004). The biological basis of implicit and automatic mentalization is probably active by 1 year but possibly even earlier (Csibra & Southgate, 2006). What happens with development then is far more to do with the differentiation of mentalisation from a stance of universality or shared (semantic) subjectivity to a stance which Sperber (2006) aptly termed ‘epistemic vigilance’. This change, which is likely to be associated with the 3-4 year old watershed in the acquisition of explicit mentalising, is in the requirement to be ‘sophisticated’ concerning the source of knowledge. Assuming the possibility of false belief, is associated with greater concern about where information originates. In Sperber’s study, nursery school children were shown to respond differentially to information supplied by people they saw positively and negatively (a good guy vs. bad guy). Children who selectively acted on positively connoted information were also likely to pass the false belief task, suggesting that monitoring mental states of others is intrinsically tied to establishing the possible motivations behind any communication. Note that most false belief tasks actually entail an act of deception. Thus, the narrowing of assumptions about shared subjectivity is fully nested in the context of ‘trust’ and ‘mistrust’. We may speculate that a child who experiences more confusion about the possibility of trust in the context of his primary attachment relationship would be at a disadvantage when it comes to this aspect of self-other differentiation.

We assume that teaching and learning about states of mind is mostly a mundane process within the attachment relationship, and that it is preconscious to both infant and parent - inaccessible to reflection or modification. Parents, however, execute this natural human function in different ways. Some are alert to the earliest

indications of intentionality, while others may need stronger clues before they can perceive the child's mental state and modify their behaviour accordingly. Yet other parents consistently misread the infant's internal state; their expectations, based on past experience or reactions to these dominate their mentalization of their infants and preclude accurate identification of intention. These biases preclude the possibility of contingent mirroring, and an emotional experience is mirrored which is incongruent with the child's constitutional experience, and is likely to lead to absent or distorted secondary representation of these experiences within the child's developing self. Yet other parents, as we have seen, fail to mark their mirroring.

### *The role of the quality of parent-child relationship*

The sociobiological roots of social cognition considered above may help us understand why the attachment system is intricately involved in the development of human subjectivity. Our constructionist model suggests that teaching about minds is part and parcel of the evolutionarily central cognitive adaptation of human pedagogy, the biological preparedness for both adult and infant to provide and receive cultural information that is new and relevant to the infant. If competition with conspecifics were the primary driver of the evolution of human cognition then the teaching function which the work of Csibra and Gergely points to would have to be protected from deliberate distortion by individuals who did not share genetic material with the infant. As has been frequently suggested, attachment may well be a helpful behavioural marker of shared genetic makeup, perhaps in both animal and human species (Belsky & Jaffee, 2006; Erickson, 1993; Fonagy, 2003). It is at least plausible, therefore, that attachment serves as a guarantee of authenticity of knowledge. Thus when overt pedagogical cues alert the infant to the adult's overt communicative intent, the infant may be more prepared to respond to these if they are

from an adult to whom he/she has an attachment bond. The prediction from this model is that there should be generic cognitive benefits to secure attachment in the sense that the child is more likely to attend fully for example to new objects identified by the known and trusted adult who is pointing them out and naming them, or who is indicating whether she or he sees the object as good or bad, as in social referencing (Baldwin & Moses, 1996; Tomasello, 1999). The more reliable processing of pedagogical information in the context of secure attachment would account for the broad and generic intellectual benefits that appear to accrue from secure attachment in infancy (Cicchetti et al., 2000; Crandell & Hobson, 1999; Jacobsen & Hofmann, 1997; van Ijzendoorn & van Vliet-Visser, 1988). Some preliminary findings (Gergely, Fonagy, & Watson, in preparation) from a study designed to test infants' relative degree of sensitivity to internal proprioceptive cues generated by their facial expression versus sensitivity to external visible equivalent expressions demonstrated that secure infants were able to switch back and forth between monitoring either external or internal expressive cues whilst insecure infants were more rigidly focused on external cues. Monitoring both the internal and external world is the most adaptive and desirable strategy for optimal coping and social reality testing in the interpersonal domain.

What we are suggesting is that the advantage of secure attachment for the precocious development of mentalization and the stronger establishment of an agentic sense of self arises out of a far more general predisposition for infants to be more ready to learn from adults with whom they have a secure bond. The attachment bond is established through the adult's attentiveness and contingent responsiveness to the infant. There is a clear overlap of biological markers. Ostensive cues of the caregiver not only bias the infant to interpret the adult's action as indicating

communicative intention to transfer relevant knowledge, but also engender attachment security through sensitive (contingent) responding (Ainsworth, Blehar, Waters, & Wall, 1978; Braungart-Rieker, Garwood, Powers, & Wang, 2001 464; De Wolff & van IJzendoorn, 1997; Fearon et al., in press; Stams, Juffer, & van IJzendoorn, 2002). For the infant these signs mark not only the possibility of physical security but also the likely veracity of information communicated by that individual. From an evolutionary standpoint, we may consider such ostensive cues (at least in infancy) to trigger a 'basic epistemic trust' in the caregiver as a benevolent, cooperative, and reliable source of cultural information. The same caregiver behaviour facilitates the creation of secure attachment while also activating a 'presumption of relevance' about the contents of manifested knowledge as a result of which the young apprentice can fast-learn it without the need to test or critically scrutinize its validity or relevance any further. As Gergely & Csibra point out (Gergely & Csibra, 2005b), the adult invests caregiving behaviour and communication of knowledge selectively, mainly focusing it on infants for whom he/she is caring and therefore usually has genetic material in common. The baby has a corresponding selectiveness, relying as far as possible on familiar, attentive and responsive adults to teach him what is safe and trustworthy in the world, and furthermore what his thoughts and feelings are and how knowledge of such internal states can eventually make a bridge to understanding and prediction in the wider social world. A world in which he and his group will have to compete and survive.

## **Conclusion**

In this review we attempted to show how insecure and unpredictable attachment relationships between parent and infant may create an adverse social environment for the acquisition of mentalization or 'mind-reading' in the child. This

may to a limited extent be adaptive in that within extreme social contexts mentalisation is a less useful strategy. If parent-child interaction is in crucial respects not genuine the child might well be de-conditioned from using this as her or his predictive strategy. Severely insecure, abusive, inconsistent and disorganized attachment relations may well be detrimental for mentalisation to survive as a dominant, predictive interpersonal strategy. However, within the same contexts of deprivation and risk, mentalization could hold the key to breaking the cycle of abuse and deprivation for that child growing up, and for the children he or she produces. In this paper, we have focused particularly on how the development of mentalization, and the building of a sense of oneself and others as thinking and feeling, may be part of a much more general process, again dependent on trustworthy attachment bonds, which supports physical and social survival. Certain aspects of the interactive background to secure attachment (e.g. attunement sensitivity) appear to have an evolutionary function that is to do with pedagogy, the teaching of what cannot be learned about the world by simple observation, or which would be too risky or time-consuming to learn by trial and error, or from strangers. One part of the world that must be learned about as quickly as possible is how people are likely to treat the child, how to predict their behaviour, will they be protective, punishing, interested and so on. We know from early attachment research that babies learn early on to adapt to the customary attitudes and behaviour of their caregivers. We now know too that secure attachment and a mind-minded reflective mirroring environment extend mentalisation to include the internal states of the self in their domain, thereby making self-prediction and emotional self control possible. We argue here that this is achieved by applying pedagogical referential communication to the domain of the internal emotional and dispositional/intentional states of the child. We tried to show that

pedagogy creates a context for the caregiver to teach the child about the subjective self and make available and construct second order representations for internal states. This ontogenetically and perhaps evolutionarily extends mentalisation to allow reading and anticipating one's own mental contents as well as those of others, thereby making emotional self-control possible, and adding a sense of coherence and predictability to inner experience as well as to the social world.

## References

- Addis, D. R., McIntosh, A. R., Moscovitch, M., Crawley, A. P., & McAndrews, M. P. (2004). Characterizing spatial and temporal features of autobiographical memory retrieval networks: a partial least squares approach. *Neuroimage*, 23(4), 1460-1471.
- Adolphs, R. (2002). Neural systems for recognizing emotion. *Curr Opin Neurobiol*, 12(2), 169-177.
- Adolphs, R., Damasio, H., Tranel, D., Cooper, G., & Damasio, A. R. (2000). A role for somatosensory cortices in the visual recognition of emotion as revealed by three-dimensional lesion mapping. *J Neurosci*, 20(7), 2683-2690.
- Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the Strange Situation*. Hillsdale, NJ: Erlbaum.
- Aitken, K. J., & Trevarthen, C. (1997). Self-other organization in human psychological development. *Development and Psychopathology*, 9, 651-675.
- Alessandri, S. M. (1991). Play and social behaviours in maltreated preschoolers. *Development and Psychopathology*, 3, 191-206.
- Alexander, R. D. (1989). Evolution of the human psyche. In P. Mellars & C. Stringer (Eds.), *The human revolution: behavioural and biological perspectives on the origins of modern humans* (pp. 455-513). Princeton: Princeton University Press.
- Alexander, R. D. (1990). *How did humans evolve? Reflections on the uniquely unique species. Museum of Zoology (Special publication no. 1)*. Ann Arbor, MI: University of Michigan.
- Allen, J. G. (2006). Mentalizing in practice. In J. G. Allen & P. Fonagy (Eds.), *Handbook of Mentalization-Based Treatment*. New York: Wiley.
- Allman, J., Hakeem, A., & Watson, K. (2002). Two phylogenetic specializations in the human brain. *Neuroscientist*, 8(4), 335-346.
- Allman, J. M., Hakeem, A., Erwin, J. M., Nimchinsky, E., & Hof, P. (2001). The anterior cingulate cortex. The evolution of an interface between emotion and cognition. *Ann N Y Acad Sci*, 935, 107-117.
- Amaral, D. G. (2003). The amygdala, social behavior, and danger detection. *Annals of the New York Academy of Sciences*, 1000, 337-347.
- Andersen, E. S., Dunlea, A., & Kekelis, L. S. (1984). Blind children's language: Resolving some differences. *Journal of Child Language*, 11, 645-664.
- Anderson, S. W., Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1999). Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nature Neurosci.*, 2, 1032-1037.
- Arnsten, A. F. T. (1998). The biology of being frazzled. *Science*, 280, 1711-1712.
- Arnsten, A. F. T., Mathew, R., Ubriani, R., Taylor, J. R., & Li, B.-M. (1999). alpha-1 noradrenergic receptor stimulation impairs prefrontal cortical cognitive function. *Biological Psychiatry*, 45, 26-31.
- Astington, J. (1996). What is theoretical about the child's theory of mind?: A Vygotskian view of its development. In P. Carruthers & P. K. Smith (Eds.), *Theories of theories of mind* (pp. 184-199). Cambridge: Cambridge University Press.
- Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology*, 35, 1311-1320.

- Avis, J., & Harris, P. (1991). Belief-desire reasoning among Baka children: evidence for a universal conception of mind. *Child Development*, *62*, 460–467.
- Aziz-Zadeh, L., Wilson, S. M., Rizzolatti, G., & Iacoboni, M. (2006). Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Curr Biol*, *16*(18), 1818-1823.
- Bachevalier, J., & Loveland, K. A. (2006). The orbitofrontal-amygdala circuit and self-regulation of social-emotional behavior in autism. *Neurosci Biobehav Rev*, *30*(1), 97-117.
- Baldwin, D. A., & Moses, L. J. (1996). The ontogeny of social information gathering. *Child Development*, *67*, 1915-1939.
- Barnett, S. A. (1973). Homo docens. *Journal of Biosocial Science*, *5*, 393–403.
- Baron-Cohen, S. (2005). Autism. In *Cambridge encyclopedia of child development* (Vol. 28, pp. 109-126). Cambridge: Cambridge University Press.
- Baron-Cohen, S., & Belmonte, M. K. (2005). Autism: a window onto the development of the social and the analytic brain. *Annu Rev Neurosci*, *28*, 109-126.
- Barrett, K., & Campos, J. (1987). Perspectives on emotional development: II. A functionalist approach to emotions. In J. D. Osofsky (Ed.), *Handbook of infant development (2nd ed)* (pp. 555-578). New York: Wiley.
- Barrett, L., & Henzi, P. (2005). The social nature of primate cognition. *Proc Biol Sci*, *272*(1575), 1865-1875.
- Bartels, A., & Zeki, S. (2000). The neural basis of romantic love. *Neuroreport*, *11*(17), 3829-3834.
- Bartels, A., & Zeki, S. (2004). The neural correlates of maternal and romantic love. *Neuroimage*, *21*(3), 1155-1166.
- Bartsch, K. (2002). The role of experience in children's developing folk epistemology: Review and analysis from the theory-theory perspective. *New Ideas in Psychology*, *20*, 145-161.
- Baumrind, D. (1991). Parenting styles and adolescent development. In J. Brooks-Gunn, R. Lerner & A. C. Petersen (Eds.), *The encyclopedia on adolescence* (pp. 746-758). New York: Garland.
- Beebe, B., Lachmann, F., & Jaffe, J. (1997). Mother - infant interaction structures and presymbolic self and object representations. *Psychoanalytic Dialogues*, *7*, 113-182.
- Beeghly, M., & Cicchetti, D. (1994). Child maltreatment, attachment, and the self system: Emergence of an internal state lexicon in toddlers at high social risk. *Development and Psychopathology*, *6*, 5-30.
- Behne, T., Carpenter, M., Call, J., & Tomasello, M. (2005). Unwilling versus unable: infants' understanding of intentional action. *Dev Psychol*, *41*(2), 328-337.
- Belsky, J., & Jaffee, S. R. (2006). The multiple determinants of parenting. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental Psychopathology (second edition)*. Vol. 3: *Risk, Disorder and Adaptation* (pp. 38-85). New York: Wiley.
- Birch, S. A., & Bloom, P. (2004). Understanding children's and adults' limitations in mental state reasoning. *Trends Cogn Sci*, *8*(6), 255-260.
- Birch, S. A. J., & Bloom, P. (2003). Children are cursed: An asymmetric bias in mental state attributions. *Psychological Science*, *14*, 283-286.
- Biró, S., & M., L. A. (in press). Infants' perception of goal-directed actions: Development through cue-based bootstrapping. *Developmental Science*.
- Blair, R. J. (2003). Facial expressions, their communicatory functions and neuro-cognitive substrates. *Philos Trans R Soc Lond B Biol Sci*, *358*(1431), 561-572.

- Bokhorst, C. L., Bakermans-Kranenburg, M. J., Fearon, R. M., van IJzendoorn, M. H., Fonagy, P., & Schuengel, C. (2003). The importance of shared environment in mother-infant attachment security: a behavioral genetic study. *Child Dev, 74*(6), 1769-1782.
- Bonvillian, J. D. (1999). Sign language development. In M. Barrett (Ed.), *The Development of Language* (pp. 277-309). Hove: Psychology Press.
- Botvinick, M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature, 402*(6758), 179-181.
- Bowlby, J. (1959). Separation anxiety. *International Journal of Psycho-Analysis, 41*, 1-25.
- Bowlby, J. (1969). *Attachment and Loss, Vol. 1: Attachment*. London: Hogarth Press and the Institute of Psycho-Analysis.
- Bowlby, J. (1973). *Attachment and Loss, Vol. 2: Separation: Anxiety and Anger*. London: Hogarth Press and Institute of Psycho-Analysis.
- Braten, S. (1988). Dialogic mind: the infant and the adult in protoconversation. In M. Carvallo (Ed.), *Nature, Cognition and System* (Vol. I, pp. 187-205). Dordrecht: Kluwer Academic Publishers.
- Braten, S. (1992). The virtual other in infants' minds and social feelings. In H. Wold (Ed.), *The Dialogical Alternative* (pp. 77-97). Oslo: Scandinavian University Press.
- Braten, S. (1998). *Intersubjective Communication and Emotion in Early Ontogeny*. Paris: Cambridge University Press.
- Braungart-Rieker, J. M., Garwood, M. M., Powers, B. P., & Wang, X. (2001). Parental sensitivity, infant affect, and affect regulation: predictors of later attachment. *Child Dev, 72*(1), 252-270.
- Braver, T. S., Barch, D. M., Gray, J. R., Molfese, D. L., & Snyder, A. (2001). Anterior cingulate cortex and response conflict: effects of frequency, inhibition and errors. *Cereb Cortex, 11*(9), 825-836.
- Brentano, F. (1973/1874). *Psychology from an Empirical Standpoint*. London: Routledge.
- Bretherton, I., Bates, E., Benigni, L., Camaioni, L., & Volterra, V. (1979). Relationships between cognition, communication, and quality of attachment. In E. Bates, L. Benigni, I. Bretherton, L. Camaioni & V. Volterra (Eds.), *The emergence of symbols: Cognition and communication in infancy* (pp. 223-269). New York: Academic Press.
- Bronfman, E., Parsons, E., & Lyons-Ruth, K. (1999). *Atypical Maternal Behavior Instrument for Assessment and Classification (AMBIANCE): Manual for coding disrupted affective communication, version 2. Unpublished manuscript*. Cambridge, MA: Harvard Medical School.
- Brune, M. (2001). Social cognition and psychopathology in an evolutionary perspective. Current status and proposals for research. *Psychopathology, 34*(2), 85-94.
- Bruner, J. S. (1993). Do we "acquire" culture or vice versa? *Behavioral and Brain Sciences, 16*, 515-516.
- Brunet, E., Sarfati, Y., Hardy-Bayle, M. C., & Decety, J. (2000). A PET investigation of the attribution of intentions with a nonverbal task. *Neuroimage, 11*(2), 157-166.

- Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur J Neurosci*, *13*(2), 400-404.
- Buccino, G., Lui, F., Canessa, N., Patteri, I., Lagravinese, G., Benuzzi, F., et al. (2004). Neural circuits involved in the recognition of actions performed by nonconspecifics: an FMRI study. *J Cogn Neurosci*, *16*(1), 114-126.
- Bugnyar, T., & Heinrich, B. (2005). Ravens, *Corvus corax*, differentiate between knowledgeable and ignorant competitors. *Proceedings of the Royal Society*, *272*, 1641-1646.
- Burgess, K. B., Marshall, P. J., Rubin, K. H., & Fox, N. A. (2003). Infant attachment and temperament as predictors of subsequent externalizing problems and cardiac physiology. *J Child Psychol Psychiatry*, *44*(6), 819-831.
- Byrne, R., & Whiten, A. (1988). *Machiavellian intelligence*. Oxford: Oxford University Press.
- Cabeza, R., & Nyberg, L. (2000). Neural bases of learning and memory: functional neuroimaging evidence. *Curr Opin Neurol*, *13*(4), 415-421.
- Calarge, C., Andreasen, N. C., & O'Leary, D. S. (2003). Visualizing how one brain understands another: a PET study of theory of mind. *Am J Psychiatry*, *160*(11), 1954-1964.
- Calmels, C., Holmes, P., Jarry, G., Hars, M., Lopez, E., Paillard, A., et al. (2006). Variability of EEG synchronization prior to and during observation and execution of a sequential finger movement. *Hum Brain Mapp*, *27*(3), 251-266.
- Camerer, C., Lowenstein, A., & Weber, B. (1989). The curse of knowledge in economic settings: an experimental analysis. *Journal of Political Economy*, *97*, 1232-1254.
- Camras, L. A., Grow, G., & Ribordy, S. (1983). Recognition of emotional expressions by abused children. *Journal of Clinical and Consulting Psychology*, *12*(3), 325-328.
- Camras, L. A., Ribordy, S., Hill, J., Martino, S., Sachs, V., Spaccarelli, S., et al. (1990). Maternal facial behavior and the recognition and production of emotional expression by maltreated and nonmaltreated children. *Developmental Psychology*, *26*(2), 304-312.
- Camras, L. A., Ribordy, S., Hill, J., Martino, S., Spaccarelli, S., & Stefani, R. (1988a). Recognition and posing of emotional expressions by abused children and their mothers. *Developmental Psychology*, *24*(6), 776-781.
- Camras, L. A., Ribordy, S., Hill, J., Martino, S., Spaccarelli, S., & Stefani, R. (1988b). Recognition and posing of emotional expressions by abused children and their mothers. *Developmental Psychology*, *24*(6), 776-781.
- Camras, L. A., Sachs-Alter, E., & Ribordy, S. C. (1996). Emotion understanding in maltreated children: Recognition of facial expressions and integration with other emotion cues. In M. D. Lewis & M. Sullivan (Eds.), *Emotional development in atypical children* (pp. 203-225). Mahwah, NJ: Erlbaum.
- Canli, T., & Amin, Z. (2002). Neuroimaging of emotion and personality: scientific evidence and ethical considerations. *Brain Cogn*, *50*(3), 414-431.
- Caro, T. M., & Hauser, M. D. (1992). Is there teaching in nonhuman animals? *Quarterly Journal of Biology*, *67*, 151-174.
- Carpendale, J. I. M., & Lewis, C. (2006). *How Children Develop Social Understanding (Understanding Children's Worlds)* London: Blackwell.

- Cassia, V. M., Turati, C., & Simion, F. (2004). Can a nonspecific bias toward top-heavy patterns explain newborns' face preference? *Psychological Science, 15*, 379–383.
- Castelli, F., Happe, F., Frith, U., & Frith, C. (2000). Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *Neuroimage, 12*(3), 314-325.
- Cavell, M. (1994). *The Psychoanalytic Mind*. Cambridge, MA: Harvard University Press.
- Champagne, F. A., Chretien, P., Stevenson, C. W., Zhang, T. Y., Gratton, A., & Meaney, M. J. (2004). Variations in nucleus accumbens dopamine associated with individual differences in maternal behavior in the rat. *J Neurosci, 24*(17), 4113-4123.
- Champagne, F. A., Weaver, I. C., Diorio, J., Sharma, S., & Meaney, M. J. (2003). Natural variations in maternal care are associated with estrogen receptor alpha expression and estrogen sensitivity in the medial preoptic area. *Endocrinology, 144*(11), 4720-4724.
- Charman, T., Ruffman, T., & Clements, W. (2002). Is there a gender difference in false belief development? . *Social Development, 11*, 1–10.
- Cheng, Y. W., Tzeng, O. J., Decety, J., Imada, T., & Hsieh, J. C. (2006). Gender differences in the human mirror system: a magnetoencephalography study. *Neuroreport, 17*(11), 1115-1119.
- Cicchetti, D., & Lynch, M. (1995). Failures in the expectable environment and their impact on individual development: The case of child maltreatment. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental Psychopathology (vol. 2)* (pp. 32-71). New York: Wiley.
- Cicchetti, D., Rogosch, F. A., Maughan, A., Toth, S. L., & Bruce, J. (2003). False belief understanding in maltreated children. *Dev Psychopathol, 15*(4), 1067-1091.
- Cicchetti, D., Rogosch, F. A., & Toth, S. L. (2000). The efficacy of toddler-parent psychotherapy for fostering cognitive development in offspring of depressed mothers. *J Abnorm Child Psychol, 28*(2), 135-148.
- Cicchetti, D., & Toth, S. L. (1995). A developmental psychopathology perspective on child abuse and neglect. *Journal of the American Academy of Child and Adolescent Psychiatry, 34*, 541-565.
- Cicchetti, D., & Walker, E. F. (2001). Editorial: Stress and development: Biological and psychological consequences. *Development and Psychopathology, 13*, 413-418.
- Cole, K., & Mitchell, P. (2000). Siblings in the development of executive control and a theory of mind. *British Journal of Developmental Psychology, 18*, 279-295.
- Conway, M. A. (1992). A structural model of autobiographical memory. In M. A. Conway, H. Spinnler & W. A. Wagenaar (Eds.), *Theoretical Perspectives on Autobiographical Memory*. (pp. 167-194). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Conway, M. A. (1996). Autobiographical knowledge and autobiographical memories. In D.C.Rubin (Ed.), *Remembering our past: Studies in autobiographical memory* (pp. 67-93). New York: Cambridge University Press.
- Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development, 61*, 1587-1595.

- Crandell, L. E., & Hobson, R. P. (1999). Individual differences in young children's IQ: a social-developmental perspective. *J Child Psychol Psychiatry*, 40(3), 455-464.
- Csibra, G. (in press). Action mirroring and action understanding: An alternative account. In P. Haggard, Y. Rosetti & M. Kawato (Eds.), *Sensorimotor Foundations of Higher Cognition. Attention and Performance?*
- Csibra, G., & Gergely, G. (2005). Social learning and social cognition: The case for pedagogy. In M. H. Johnson & Y. Munakata (Eds.), *Processes of change in brain and cognitive development. Attention and Performance XXI*. Oxford: Oxford University Press.
- Csibra, G., & Gergely, G. (2006). Social learning and social cognition: The case for pedagogy. In M. H. Johnson & Y. M. Munakata (Eds.), *Processes of change in brain and cognitive development. Attention and Performance* (Vol. XXI, pp. 249-274). Oxford: Oxford University Press.
- Csibra, G., & Gergely, G. (in press). 'Obsessed with goals': Functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica (special issue)*.
- Csibra, G., Gergely, G., Bíró, S., Koós, O., & Brockbank, M. (1999). Goal attribution without agency cues: The perception of 'pure reason' in infancy. *Cognition*, 72, 237-267.
- Csibra, G., & Southgate, V. (2006). Evidence for infants understanding false beliefs should not be dismissed. *Trends in Cognitive Sciences*, 10, 4-5.
- Curtiss, S. (1977). *Genie: A psycholinguistic study of a modern-day "wild child."* New York: Academic Press.
- Cutting, A. L., & Dunn, J. (1999). Theory of mind, emotion understanding, language, and family background: Individual differences and interrelations. *Child Development*, 70, 853-865.
- Cutting, A. L., & Dunn, J. (2002). The cost of understanding other people: social cognition predicts young children's sensitivity to criticism. *J Child Psychol Psychiatry*, 43(7), 849-860.
- Damasio, A. (1995). *Descartes' Error: Emotion, reason and the human brain*. London: Macmillan.
- Damasio, A. R. (2003). *Looking for Spinoza: Joy, sorrow, and the feeling brain*. New York: Harvest Books.
- Damasio, A. R., Grabowski, T. J., Bechara, A., Damasio, H., Ponto, L. L., Parvizi, J., et al. (2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nat Neurosci*, 3(10), 1049-1056.
- Dapretto, M., Davies, M. S., Pfeifer, J. H., Scott, A. A., Sigman, M., Bookheimer, S. Y., et al. (2006). Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. *Nat Neurosci*, 9(1), 28-30.
- Davidson, D. (1987). Knowing one's own mind. *Proceedings and Addresses of the American Philosophical Association*, 60, 441-457.
- de Rosnay, M., & Harris, P. L. (2002). Individual differences in children's understanding of emotion: the roles of attachment and language. *Attach Hum Dev*, 4(1), 39-54.
- de Rosnay, M., & Hughes, C. (2006). Conversation and theory of mind: Do children talk their way to socio-cognitive understanding? *British Journal of Developmental Psychology*, 24(1), 7-37.

- de Villiers, J. (2005). Can language acquisition give children a point of view? In J. W. Astington & J. A. Baird (Eds.), *Why language matters for theory of mind* (pp. 186-219). Oxford: Oxford University Press.
- de Villiers, J. G., & de Villiers, P. (2000). Linguistic determinism and the understanding of false belief. In P. Mitchell & K. Riggs (Eds.), *Children's reasoning and the mind* (pp. 191-228). Hove, UK: Psychology Press.
- de Villiers, J. G., & Pyers, J. E. (2002). Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief-understanding. *Cognitive Development, 17*, 1037-1060.
- De Wolff, M. S., & van IJzendoorn, M. H. (1997). Sensitivity and attachment: A meta-analysis on parental antecedents of infant attachment. *Child Development, 68*, 571-591.
- Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behav Cogn Neurosci Rev, 3*(2), 71-100.
- den Ouden, H. E., Frith, U., Frith, C., & Blakemore, S. J. (2005). Thinking about intentions. *Neuroimage, 28*(4), 787-796.
- Denham, S. A., Zoller, D., & Couchoud, E. A. (1994). Socialization of preschoolers emotion understanding. *Developmental Psychology, 30*, 928-936.
- Dennett, D. C. (1978). Beliefs about beliefs. *Behaviour and Brain Sciences, 4*, 568-570.
- Dennett, D. C. (1991). *Consciousness Explained*. Boston: Little Brown.
- Dorris, L., Espie, C. A., Knott, F., & Salt, J. (2004). Mind-reading difficulties in the siblings of people with Asperger's syndrome: evidence for a genetic influence in the abnormal development of a specific cognitive domain. *J Child Psychol Psychiatry, 45*(2), 412-418.
- Dunbar, R. (1996). *Grooming, gossip and the evolution of language*. London: Faber.
- Dunlea, A. (1989). *Vision and the emergence of meaning*. Cambridge: Cambridge University Press.
- Dunn, J. (1988). *The Beginnings of Social Understanding*. Oxford: Basil Blackwell Ltd and Cambridge, MA: Harvard University Press.
- Dunn, J. (1996). The Emanuel Miller Memorial Lecture 1995. Children's relationships: Bridging the divide between cognitive and social development. *Journal of Child Psychology and Psychiatry, 37*, 507-518.
- Dunn, J. (1999). Making sense of the social world: Mindreading, emotion and relationships. In P. D. Zelazo, J. W. Astington & D. R. Olson (Eds.), *Developing theories of intention: Social understanding and self control* (pp. 229-242). Mahwah, NJ: Lawrence Erlbaum Associates.
- Dunn, J., & Brophy, M. (2005). Communication relationships and individual differences in children's understanding of mind. In J. W. Astington & J. A. Baird (Eds.), *Why languages matter for theory of mind* (pp. 50-69). New York: Oxford University Press.
- Dunn, J., & Brown, J. (2001). Relationships, talk about feelings, and the development of affect regulation in early childhood. In J. Garber & K. Dodge (Eds.), *Affect regulation and dysregulation in childhood* (pp. 89-108). Cambridge: Cambridge University Press.
- Dunn, J., Brown, J., & Beardsall, L. (1991). Family talk about feeling states and children's later understanding of others' emotions. *Developmental Psychology, 27*, 448-455.

- Dunn, J., Brown, J., Somkowski, C., Telsa, C., & Youngblade, L. (1991). Young children's understanding of other people's feelings and beliefs: Individual differences and their antecedents. *Child Development*, 62, 1352-1366.
- Dunn, J., & Cutting, A. L. (1999). Understanding others, and individual differences in friendship interactions in young children. *Social Development*, 8(2), 202-219.
- During, S., & McMahon, R. (1991). Recognition of emotional facial expressions by abusive mothers and their children. . *Journal of Clinical and Consulting Psychology*, , 20(2), 132-139.
- Egyed, K., Király, & Gergely, G. (2004). *Object-centered versus agent-centered Interpretations of referential attitude expressions in 14-month-olds (poster)*. Paper presented at the 14th Biennial International Conference on Infant Studies, Chicago, May.
- Emery, N. J., & Clayton, N. S. (2004). The mentality of crows: Convergent evolution of intelligence in Corvids and Apes. *Science*, 306, 1903-1907.
- Erickson, M. T. (1993). Rethinking Oedipus: an evolutionary perspective of incest avoidance. *Am J Psychiatry*, 150(3), 411-416.
- Eslinger, P. J. (2001). Adolescent neuropsychological development after early right prefrontal cortex damage. *Dev. Neuropsychol.* , 18, 297-329.
- Eslinger, P. J., & Damasio, A. R. (1985). Severe disturbance of higher cognition after bilateral frontal lobe ablation: patient EVR. *Neurology* 35, 1731-1741.
- Eslinger, P. J., Grattan, L. M., Damasio, H., & Damasio, A. R. (1992). Developmental consequences of childhood frontal lobe damage. *Arch. Neurol.*, 49, 764-769.
- Farrant, A., Morris, R. G., Russell, T., Elwes, R., Akanuma, N., Alarcon, G., et al. (2005). Social cognition in frontal lobe epilepsy. *Epilepsy Behav*, 7(3), 506-516.
- Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the USA*, 99, 9602-9605.
- Farroni, T., Massaccesi, S., Pividori, D., Simion, F., & Johnson, M. H. (2004). Gaze following in newborns. *Infancy*, 5, 39-60.
- Fearon, P., Target, M., Fonagy, P., Williams, L., McGregor, J., Sargent, J., et al. (2006). Short-Term Mentalization and Relational Therapy (SMART): An integrative family therapy for children and adolescents. In J. Allen & P. Fonagy (Eds.), *Handbook of mentalisation based treatments*. London: John Wiley.
- Fearon, P., van IJzendoorn, M. H., Fonagy, P., Bokhorst, C. L., Bakermans-Kranenburg, M. J., & Schuengel, C. (in press). In search of shared and non-shared environmental factors in security of attachment: A behavior-genetic study of the association between sensitivity and attachment security. *Developmental Psychology*.
- Fearon, R. M., & Belsky, J. (2004). Attachment and attention: protection in relation to gender and cumulative social-contextual adversity. *Child Dev*, 75(6), 1677-1693.
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Developmental Psychology*, 8, 181-195.
- Ferrari, P. F., Gallese, V., Rizzolatti, G., & Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *Eur J Neurosci*, 17(8), 1703-1714.

- Ferstl, E. C., & von Cramon, D. Y. (2002). What does the frontomedian cortex contribute to language processing: coherence or theory of mind? *Neuroimage*, *17*(3), 1599-1612.
- Field, T., Woodson, R., Cohen, D., Garcia, R., & Greenberg, R. (1983). Discrimination and imitation of facial expressions by term and preterm neonates. *Infant Behavior and Development*, *6*, 485-490.
- Fischhoff, B. (1975). Hindsight does not equal foresight: the effect of extreme knowledge on judgement under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, *1*, 349-358.
- Fletcher, P. C., Happe, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S., et al. (1995). Other minds in the brain: a functional imaging study of "theory of mind" in story comprehension. *Cognition*, *57*(2), 109-128.
- Floccia, C., Christophe, A., & Bertoncini, J. (1997). High-amplitude sucking and newborns: The quest for underlying mechanisms. *Journal of Experimental Child Psychology*, *64*, 175-189.
- Fogel, A., Nwokah, E., Dedo, J. Y., Messinger, D., Dickson, K. L., & Holt, S. A. (1992). Social process theory of emotion: A dynamic systems approach. *Social Development*, *2*, 122-142.
- Fonagy, P. (1991). Thinking about thinking: Some clinical and theoretical considerations in the treatment of a borderline patient. *International Journal of Psycho-Analysis*, *72*, 1-18.
- Fonagy, P. (2001). The human genome and the representational world: The role of early mother-infant interaction in creating an interpersonal interpretive mechanism. *Bulletin of the Menninger Clinic*, *65*, 427-448.
- Fonagy, P. (2003). The development of psychopathology from infancy to adulthood: the mysterious unfolding of disturbance in time. *Infant Mental Health Journal*, *24*(3), 212-239.
- Fonagy, P., Gergely, G., Jurist, E., & Target, M. (2002). *Affect Regulation, Mentalization and the Development of the Self*. New York: Other Press.
- Fonagy, P., Redfern, S., & Charman, T. (1997). The relationship between belief-desire reasoning and a projective measure of attachment security (SAT). *British Journal of Developmental Psychology*, *15*, 51-61.
- Fonagy, P., Steele, H., Steele, M., & Holder, J. (1997). Attachment and theory of mind: Overlapping constructs? *Association for Child Psychology and Psychiatry Occasional Papers*, *14*, 31-40.
- Fonagy, P., Stein, H., Allen, D., Chen, C.-F., Allen, J. G., & Vrouva, I. (2006). The relationship of childhood and adolescent adversity to impairment of mentalizing capacity and psychological disorder. Unpublished paper, Department of Psychology, University College London.
- Fonagy, P., & Target, M. (1997). Attachment and reflective function: Their role in self-organization. *Development and Psychopathology*, *9*, 679-700.
- Fraiberg, S. (1977). *Insights from the blind*. London: Souvenir Press.
- Francis, D. D., Szegda, K., Campbell, G., Martin, W. D., & Insel, T. R. (2003). Epigenetic sources of behavioral differences in mice. *Nat Neurosci*, *6*(5), 445-446.
- Fries, A. B., Ziegler, T. E., Kurian, J. R., Jacoris, S., & Pollak, S. D. (2005). Early experience in humans is associated with changes in neuropeptides critical for regulating social behavior. *Proc Natl Acad Sci U S A*, *102*(47), 17237-17240.
- Frith, U. (1989). *Autism: Explaining the Enigma*. Oxford: Blackwell.

- Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London B, Biological Sciences*, 358, 459-473.
- Frodi, A., & Smetana, J. (1984). Abused, neglected, and nonmaltreated preschoolers' ability to discriminate emotions in others: The effects of IQ. *Child Abuse & Neglect*, 8(4), 459-465.
- Futó, J., Bártki, A., Koós, O., Fonagy, P., & Gergely, G. (2004). *Early social-interactive determinants of later representational and affect-regulative competence in pretend play (Poster)*. Paper presented at the 14th Biennial International Conference on Infant Studies, Chicago, May.
- Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of 'theory of mind'. *Trends Cogn Sci*, 7(2), 77-83.
- Gallagher, H. L., Happe, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: an fMRI study of 'theory of mind' in verbal and nonverbal tasks. *Neuropsychologia*, 38(1), 11-21.
- Gallagher, H. L., Jack, A. I., Roepstorff, A., & Frith, C. D. (2002). Imaging the intentional stance in a competitive game. *Neuroimage*, 16(3 Pt 1), 814-821.
- Gallese, V. (2003). The roots of empathy: the shared manifold hypothesis and the neural basis of intersubjectivity. *Psychopathology*, 36(4), 171-180.
- Gallese, V. (2006). Intentional attunement: a neurophysiological perspective on social cognition and its disruption in autism. *Brain Res*, 1079(1), 15-24.
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119 ( Pt 2), 593-609.
- Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 2(12), 493-501.
- Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends Cogn Sci*, 8(9), 396-403.
- Gazzola, V., Aziz-Zadeh, L., & Keysers, C. (2006). Empathy and the somatotopic auditory mirror system in humans. *Curr Biol*, 16(18), 1824-1829.
- Geary, D. C. (2005). *The origin of mind: Evolution of brain, cognition, and general intelligence*. Washington, DC: American Psychological Association.
- Geary, D. C., & Huffman, K. J. (2002). Brain and cognitive evolution: forms of modularity and functions of mind. *Psychol Bull*, 128(5), 667-698.
- Gergely, G. (2002). The development of understanding of self and agency. In U. Goswami (Ed.), *Handbook of Childhood Cognitive Development* (pp. 26-46). Oxford: Blackwell.
- Gergely, G. (2004). The role of contingency detection in early affect-regulative interactions and in the development of different types of infant attachment. *Social Behavior*, 13, 468-478.
- Gergely, G., Bekkering, H., & Kiraly, I. (2002). Rational imitation in preverbal infants. *Nature*, 415, 755.
- Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naive theory of rational action. *Trends in Cognitive Sciences*, 7, 287-292.
- Gergely, G., & Csibra, G. (2005a). A few reasons why we don't share Tomasello et al.'s Intuitions about sharing. Commentary on Tomasello et al.: Understanding and sharing intentions. *Behavioral and Brain Sciences*, 28(5), 731-732.
- Gergely, G., & Csibra, G. (2005b). The social construction of the cultural mind: Imitative learning as a mechanism of human pedagogy. *Interaction Studies*, 6, 463-481.

- Gergely, G., & Csibra, G. (2006). Sylvia's recipe: Human culture, imitation, and pedagogy. In N. J. Enfield & S. C. Levinson (Eds.), *Roots of Human Sociality: Culture, Cognition, and Human Interaction*. Berg Press: London. (pp. 229-255).
- Gergely, G., Fonagy, P., & Watson, J. S. (in preparation). Attachment security and early social sensitization to internal versus external cues of state-expressive behaviors in the self and the other.
- Gergely, G., Kiraly, I., & Egyed, K. (in press). On pedagogy. *Developmental Science*, *10*, 755.
- Gergely, G., & Watson, J. (1996). The social biofeedback model of parental affect-mirroring. *International Journal of Psycho-Analysis*, *77*, 1181-1212.
- Gergely, G., & Watson, J. (1999). Early social-emotional development: Contingency perception and the social biofeedback model. In P. Rochat (Ed.), *Early social cognition: Understanding others in the first months of life* (pp. 101-137). Hillsdale, NJ: Erlbaum.
- Gianino, A. F., & Tronick, E. Z. (1988). The mutual regulation model: The infant's self and interactive regulation and coping and defensive capacities. In T. M. Field, P. M. McCabe & N. Schneiderman (Eds.), *Stress and Coping Across Development* (pp. 47-68). Hillsdale, NJ: Erlbaum.
- Gillath, O., Bunge, S. A., Shaver, P. R., Wendelken, C., & Mikulincer, M. (2005). Attachment-style differences in the ability to suppress negative thoughts: exploring the neural correlates. *Neuroimage*, *28*(4), 835-847.
- Goel, V., Grafman, J., Sadato, N., & Hallett, M. (1995). Modeling other minds. *Neuroreport*, *6*(13), 1741-1746.
- Goldberg, S., Benoit, D., Blokland, K., & Madigan, S. (2003). Atypical maternal behavior, maternal representations, and infant disorganized attachment. *Dev Psychopathol*, *15*(2), 239-257.
- Goldman, A. I. (1993). The psychology of folk psychology. *Behavioral & Brain Sciences*, *16*, 29-113.
- Goldman, A. I., & Sripada, C. S. (2005). Simulationist models of face-based emotion recognition. *Cognition*, *94*(3), 193-213.
- Gopnik, A. (1993). How we know our minds: The illusion of first-person knowledge of intentionality. *Behavioral and Brain Sciences*, *16*, 1-14, 29-113.
- Gordon, R. M. (1995). Simulation without introspection or inference from me to you. In T. Stone & M. Davies (Eds.), *Mental simulation: Evaluations and applications* (pp. 101-119). Oxford: Blackwell.
- Grady, C. L., & Keightley, M. L. (2002). Studies of altered social cognition in neuropsychiatric disorders using functional neuroimaging. *Can J Psychiatry*, *47*(4), 327-336.
- Grafman, J. (1995). Similarities and distinctions among current models of prefrontal cortical functions. *Annals of the New York Academy of Sciences*, *769*, 337-368.
- Green, S., Pring, L., & Swettenham, J. (2004). An investigation of first-order false belief understanding of children with congenital profound visual impairment. *British Journal of Developmental Psychology*, *22*(1), 1-17.
- Greene, J., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends Cogn Sci*, *6*(12), 517-523.
- Grezes, J., Frith, C. D., & Passingham, R. E. (2004). Inferring false beliefs from the actions of oneself and others: an fMRI study. *Neuroimage*, *21*(2), 744-750.
- Grienemberger, J., Kelly, K., & Slade, A. (2005). Maternal reflective functioning, mother-infant affective communication, and infant attachment: Exploring the

- link between mental states and observed caregiving behaviour in the intergenerational transmission of attachment. *Attachment and Human Development*, 7(3), 299-311.
- Grimshaw, G. M., Adelstein, A., Bryden, M. P., & MacKinnon, G. E. (1998). First-language acquisition in adolescence: Evidence for a critical period for verbal language development. *Brain and Language*, 63, 237-255.
- Guajardo, J. J., & Woodward, A. L. (2004). Is agency skin-deep? Surface attributes influence infants' sensitivity to goal-directed action. *Infancy*, 6, 361-384.
- Gusnard, D. A., Akbudak, E., Shulman, G. L., & Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proc Natl Acad Sci U S A*, 98(7), 4259-4264.
- Hamann, S. (2003). Nosing in on the emotional brain. *Nat Neurosci*, 6(2), 106-108.
- Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behavior*, 61, 139-151.
- Harris, P. (1996). Desires, beliefs, and language. In P. Carruthers & P. K. Smith (Eds.), *Theories of theories of mind* (pp. 200-221). Cambridge: Cambridge University Press.
- Harris, P. (2005). Conversation, pretence and theory of mind. In J. Astington & J. Baird (Eds.), *Why language matters for theory of mind* (pp. 70-83). New York: Oxford University Press.
- Harris, P. L. (1991). The work of the imagination. In A. Whiten (Ed.), *Natural Theories of Mind* (pp. 283-304). Oxford: Blackwell.
- Harris, P. L. (1992). From simulation to folk psychology: The case for development. *Mind and Language*, 7, 120-144.
- Harris, P. L. (1999). Individual differences in understanding emotions: The role of attachment status and emotional discourse. *Attachment and Human Development*, 1(3), 307-324.
- Harris, P. L., & Kavanaugh, R. D. (1993). Young children's understanding of pretence. *Monographs of the Society for Research in Child Development*, 58(1).
- Hobson, P. (2002). *The cradle of thought: Explorations of the origins of thinking*. Oxford: Macmillan.
- Hobson, R. P. (1993). *Autism and the development of mind*. London: Lawrence Erlbaum.
- Hobson, R. P., & Bishop, M. (2003). The pathogenesis of autism: insights from congenital blindness. *Philos Trans R Soc Lond B Biol Sci*, 358(1430), 335-344.
- Hobson, R. P., Lee, A., & Brown, R. (1999). Autism and congenital blindness. *J Autism Dev Disord*, 29(1), 45-56.
- Hofer, M. A. (2004). The Emerging Neurobiology of Attachment and Separation: How Parents Shape Their Infant's Brain and Behavior. In S. W. Coates & J. L. Rosenthal (Eds.), *September 11- "When the Bough Broke", Attachment Theory, Psychobiology, and Social Policy: An Integrated Approach to Trauma*. New York: Analytic Press.
- Holloway, R. L. (1996). Evolution of the human brain. In A. Lock & C. R. Peters (Eds.), *Handbook of human symbolic evolution* (pp. 74-116). New York: Oxford University Press.

- Howes, C., & Espinosa, M. P. (1985). The consequences of child abuse for the formation of relationships with peers. *International Journal of Child Abuse and Neglect*, 9, 397-404.
- Hrdy, S. B. (2000). *Mother nature*. New York: Ballentine Books.
- Hughes, C., & Cutting, A. (1999). Nature, nurture and individual differences in early understanding of mind. *Psychological Science*, 10, 429-432.
- Hughes, C., Deater-Deckard, K., & Cutting, A. (1999). 'Speak roughly to your little boy?' Sex differences in the relations between parenting and preschoolers' understanding of mind. *Social Development*, 8, 143-160.
- Hughes, C., & Dunn, J. (2002). 'When I say a naughty word.' Children's accounts of anger and sadness in self, mother and friend: Longitudinal findings from ages four to seven. *British Journal of Developmental Psychology*, 20, 515-535.
- Hughes, C., Jaffee, S. R., Happe, F., Taylor, A., Caspi, A., & Moffitt, T. E. (2005). Origins of individual differences in theory of mind: from nature to nurture? *Child Dev*, 76(2), 356-370.
- Hughes, C., & Leekham, S. (2004). What are the links between theory of mind and social realtions? Review, reflections and new directions for studies of typical and atypical development. *Social Behavior*, 13, 590-619.
- Hughes, M., Dote-Kwan, J., & Dolendo, J. (1998). A closer look at the cognitive play of preschoolers with visual impairments in the home. *Exceptional Children*, 64, 451-462.
- Insel, T. (1997). A neurobiological basis of social attachment. *American Journal of Psychiatry*, 154, 726-735.
- Insel, T. R. (2003). Is social attachment an addictive disorder? *Physiol Behav*, 79(3), 351-357.
- Jackson, P. L., Meltzoff, A. N., & Decety, J. (2005). How do we perceive the pain of others? A window into the neural processes involved in empathy. *Neuroimage*, 24(3), 771-779.
- Jacobsen, T., & Hofmann, V. (1997). Children's attachment representations: longitudinal relations to school behavior and academic competency in middle childhood and adolescence. *Dev Psychol*, 33(4), 703-710.
- Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., & Jasnow, M. D. (2001). Rhythms of Dialogue in Infancy. *Monographs of the Society for Research in Child Development*, 66(2).
- Jaworski, J. N., Francis, D. D., Brommer, C. L., Morgan, E. T., & Kuhar, M. J. (2005). Effects of early maternal separation on ethanol intake, GABA receptors and metabolizing enzymes in adult rats. *Psychopharmacology (Berl)*.
- Jenkins, J., & Astington, J. W. (1996). Cognitive factors and family structure associated with theory of mind development in young children. *Developmental Psychology*, 32, 70-78.
- Jenkins, J. M., & Astington, J. W. (2000). Theory of mind and social behavior: Causal models tested in a longitudinal study. *Merrill-Palmer Quarterly*, 46, 203-220.
- Jenkins, J. M., & Oatley, K. (2004). The space in between: The development of joint thinking and planning *Behavioral and Brain Sciences*, 27(112-113).
- Jenkins, J. M., Turrell, S. L., Kogushi, Y., Lollis, S., & Ross, H. S. (2003). A longitudinal investigation of the dynamics of mental state talk in families. *Child Dev*, 74(3), 905-920.

- Kagan, J. (1992). The conceptual analysis of affects. In T. Shapiro & R. N. Emde (Eds.), *Affects: Psychoanalytic Perspectives*. Madison: International Universities Press.
- Kamerawi, K., Kato, M., Kanda, T., Ishiguro, H., & Hiraki, K. (2005). Six-and-a-half-month-old children positively attribute goals to human action and to humanoid-robot motion. *Cognitive Development, 20*, 303-320.
- Kelley, C., & Jacoby, L. (1996). Adult egocentrism: subjective experience versus analytic bases for judgement. *Journal of Memory and language, 35*, 157-175.
- Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition, 89*(1), 25-41.
- Keysers, C., & Gazzola, V. (2006). Chapter 21 Towards a unifying neural theory of social cognition. *Prog Brain Res, 156*, 379-401.
- Király, I., Csibra, G., & Gergely, G. (2004). *The role of communicative-referential cues in observational learning during the second year (poster)*. Paper presented at the 14th Biennial International Conference on Infant Studies, Chicago, May.
- Király, I., Jovanovic, B., Prinz, W., Aschersleben, G., & Gergely, G. (2003). The early origins of goal attribution in infancy. *Consciousness & Cognition, 12*, 752-769.
- Klimes-Dougan, B., & Kistner, J. . (1990). Physically abused preschoolers' responses to peers' distress. . *Developmental Psychology, 25*, 516-524.
- Kobak, R., Cassidy, J., Lyons-Ruth, K., & Ziv, Y. (2006). Attachment, stress and psychopathology: A developmental pathways model. In D. Cicchetti & D. J. Cohen (Eds.), *Development and Psychopathology (second edition). Vol. 1: Theory and Method* (pp. 334-369). New York: Wiley.
- Kobayashi, C., Glover, G. H., & Temple, E. (2006). Cultural and linguistic influence on neural bases of 'Theory of Mind': an fMRI study with Japanese bilinguals. *Brain Lang, 98*(2), 210-220.
- Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science, 297*(5582), 846-848.
- Koós, O., & Gergely, G. (2001). The 'flickering switch' hypothesis: A contingency-based approach to the etiology of disorganized attachment in infancy. In J. Allen, P. Fonagy & G. Gergely (Eds.), *Contingency Perception and Attachment in Infancy, Special Issue of the Bulletin of the Menninger Clinic* (pp. 397-410). New York: Guilford.
- Koren-Karie, N., Oppenheim, D., Dolev, S., Sher, S., & Etzion-Carasso, A. (2002). Mother's insightfulness regarding their infants' internal experience: Relations with maternal sensitivity and infant attachment. *Developmental-Psychology, 38*, 534-542.
- Kruger, A., & Tomasello, M. (1996). Cultural learning and learning culture. In D. Olson & N. Torrance (Eds.), *The Handbook of education and human development* (pp. 369–387). Oxford: Blackwell.
- Laible, D. J., & Thompson, R. A. (1998). Attachment and emotional understanding in pre-school children. *Developmental Psychology, 34*, 1038-1045.
- LeDoux, J. E. (2000). Emotion circuits in the brain. *Annu Rev Neurosci, 23*, 155-184.
- Lee, K., Olson, D. R., & Torrance, N. (1999). Chinese children's understanding of false beliefs: The role of language. *Journal of Child Language, 26*, 1-21.
- Legerstee, M., & Varghese, J. (2001). The role of maternal affect mirroring on social expectancies in 2-3 month-old infants. *Child Development, 72*, 1301-1313.

- Leslie, A. M. (1987). Pretense and representation: The origins of 'Theory of Mind'. *Psychological Review*, 94, 412-426.
- Leslie, A. M. (2000). "Theory of Mind" as a mechanism of selective attention. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (2nd ed., pp. 1235-1247). Cambridge, Massachusetts: The MIT Press.
- Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in 'theory of mind'. *Trends in Cognitive Sciences*, 8, 528-533.
- Levine, B. (1999). Self-regulation and autothetic consciousness. In E. Tulving (Ed.), *Memory, consciousness and the brain: the Talinn Conference* (pp. 200-214). Philadelphia: Psychology Press.
- Lewis, C., Freeman, N. H., Kyriakidou, C., Maridaki-Kassotaki, K., & Berridge, D. (1996). Social influences on false belief access: Specific sibling influences or general apprenticeship? *Child Development*, 67, 2930-2947.
- Lewis, M., Alessandri, S. M., & Sullivan, M. W. (1990). Violation of expectancy, loss of control and anger expressions in young infants. *Developmental Psychology*, 26(5), 745-751.
- Lewis, M., & Brooks, J. (1978). Self-knowledge and emotional development. In M. D. Lewis & L. A. Rosenblum (Eds.), *The Development of Affect*, (pp. 205-226). New York: Plenum Press.
- Lewis, M., & Michaelson, L. (1983). *Children's Emotions and Moods: Developmental Theory and Measurement*. New York: Plenum Press.
- Lewis, M. D., & Granic, I. (2000). *Emotion, Development, and Self-Organization: Dynamic Systems Approaches to Emotional Development*. Cambridge: Cambridge University Press.
- Lewis, V., Norgate, S., Collis, G., & Reynolds, R. (2000). The consequences of visual impairment for children's symbolic and functional play. *British Journal of Developmental Psychology*, 18, 449-464.
- Lim, M. M., Murphy, A. Z., & Young, L. J. (2004). Ventral striatopallidal oxytocin and vasopressin V1a receptors in the monogamous prairie vole (*Microtus ochrogaster*). *J Comp Neurol*, 468(4), 555-570.
- Lim, M. M., Wang, Z., Olazabal, D. E., Ren, X., Terwilliger, E. F., & Young, L. J. (2004). Enhanced partner preference in a promiscuous species by manipulating the expression of a single gene. *Nature*, 429(6993), 754-757.
- Lim, M. M., & Young, L. J. (2004). Vasopressin-dependent neural circuits underlying pair bond formation in the monogamous prairie vole. *Neuroscience*, 125(1), 35-45.
- Liszkowski, U. (2006). Infant Pointing at Twelve Months: Communicative Goals, Motives, and Social-Cognitive Abilities. . In N. Enfield & S. Levinson (Eds.), *The roots of human sociality: Culture, cognition, and interaction*. (pp. 153-178). Berg: Oxford.
- Liszkowski, U., Carpenter, M., Henning, A., Striano, T., & Tomasello, M. (2004). Twelvemonth-olds point to share attention and interest. *Developmental Science*, 7, 297-307.
- Lohmann, H., & Tomasello, M. (2003). The role of language in the development of false belief understanding: A training study. *Child Development*, 74, 1130-1144.
- Lotze, M., Heymans, U., Birbaumer, N., Veit, R., Erb, M., Flor, H., et al. (2006). Differential cerebral activation during observation of expressive gestures and motor acts. *Neuropsychologia*, 44(10), 1787-1795.

- Luo, Y., & Baillargeon, R. (2005). Can a self-propelled box have a goal? Psychological reasoning in 5-month-old infants. *Psychological Science, 16*, 601-608.
- Macfie, J., Cicchetti, D., & Toth, S. L. (2001). The development of dissociation in maltreated preschool-aged children. *Development and Psychopathology, 13*, 233-254.
- Macfie, J., Toth, S. L., Rogosch, F. A., Robinson, J., Emde, R. N., & Cicchetti, D. (1999). Effect of maltreatment on preschoolers' narrative representations of responses to relieve distress and of role reversal. *Developmental Psychology, 35*, 460-465.
- Macguire, M., & Dunn, J. (1997). Friendships in early childhood, and social understanding. *International Journal of Behavioural Development, 21*, 669-686.
- MacLean, P. (1990). *The triune brain in evolution: role in paleocerebral functions*. New York: Plenum.
- Maddock, R. J. (1999). The retrosplenial cortex and emotion: new insights from functional neuroimaging of the human brain. *Trends Neurosci, 22*(7), 310-316.
- Mahler, M. S., Pine, F., & Bergman, A. (1975). *The Psychological Birth of the Human Infant: Symbiosis and Individuation*. New York: Basic Books.
- Main, M., & George, C. (1985). Responses of abused and disadvantaged toddlers to distress in agemates: A study in the daycare setting. *Developmental Psychology, 21*, 407-412.
- Malloch, S. (1999). Mother and infants and communicative musicality. In I. Deliège (Ed.), *Rhythms, musical narrative, and the origins of human communication: Musicae Scientiae, 1999-2000* (pp. 29-57). Liege: European Society for the Cognitive Sciences of Music.
- Maridaki-Kassotaki, K., Lewis, C., & Freeman, N. H. (2003). Lexical choice can lead to problems: What false belief tests tell us about Greek alternative verbs of agency. *Journal of Child Language, 30*, 1-20.
- Maughan, A., & Cicchetti, D. (2002). Impact of child maltreatment and interadult violence on children's emotion regulation abilities and socioemotional adjustment. *Child Development, 73*(5), 1525-1542.
- Mayberg, H. S., Liotti, M., Brannan, S. K., McGinnis, S., Mahurin, R. K., Jerabek, P. A., et al. (1999). Reciprocal limbic-cortical function and negative mood: converging PET findings in depression and normal sadness. *Am J Psychiatry, 156*(5), 675-682.
- Mayes, L. C. (2000). A developmental perspective on the regulation of arousal states. *Seminars in Perinatology, 24*, 267-279.
- McAlpine, L. M., & Moore, C. L. (1985). The development of social understanding in children with visual impairments. *Journal of Visual Impairment and Blindness, 89*, 349-358.
- McCabe, K., Houser, D., Ryan, L., Smith, V., & Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences, 98*, 11832-11835.
- Meins, E., & Fernyhough, C. (1999). Linguistic acquisitional style and mentalising development: The role of maternal mind-mindedness. *Cognitive Development, 14*, 363-380.
- Meins, E., Fernyhough, C., Russel, J., & Clark-Carter, D. (1998). Security of attachment as a predictor of symbolic and mentalising abilities: a longitudinal study. *Social Development, 7*, 1-24.

- Meins, E., Fernyhough, C., Wainwright, R., Clark-Carter, D., Das Gupta, M., Fradley, E., et al. (2003). Pathways to understanding mind: construct validity and predictive validity of maternal mind-mindedness. *Child Dev*, *74*(4), 1194-1211.
- Meins, E., Fernyhough, C., Wainwright, R., Das Gupta, M., Fradley, E., & Tuckey, M. (2002). Maternal mind-mindedness and attachment security as predictors of theory of mind understanding. *Child Development*, *73*, 1715-1726.
- Meins, E., Ferryhough, C., Fradley, E., & Tuckey, M. (2001). Rethinking maternal sensitivity: Mothers' comments on infants mental processes predict security of attachment at 12 months. *Journal of Child Psychology and Psychiatry*, *42*, 637-648.
- Meltzoff, A. M. (2005). Imitation and other minds: The “like me” hypothesis. In S. Hurley & N. Chater (Eds.), *Perspectives on imitation: From neuroscience to social science (vol. 2)* (pp. 55-77). Cambridge, MA: MIT Press.
- Meltzoff, A. N. (1990). Foundations for developing a concept of self: The role of imitation in relating self to other and the value of social mirroring, social modeling and self practice in infancy. In D.Cicchetti & M.Beehly (Eds.), *The Self in Transition: Infancy to Childhood*. Chicago: University of Chicago Press.
- Meltzoff, A. N., & Gopnik, A. (1993). The role of imitation in understanding persons and developing a theory of mind. In S. Baron-Cohen, H. Tager-Flusberg & D. Cohen (Eds.), *Understanding Other Minds: Perspectives from Autism* (pp. 335-366). New York: Oxford University Press, Inc.
- Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, *198*, 75-78.
- Meltzoff, A. N., & Moore, M. K. (1989). Imitation in newborn infants: Exploring the range of gestures imitated and the underlying mechanisms. *Developmental Psychology*, *25*, 954-962.
- Meltzoff, A. N., & Moore, M. K. (1998). Infant intersubjectivity: broadening the dialogue to include imitation, identity and intention. In S. Braten (Ed.), *Intersubjective Communication and Emotion in Early Ontogeny* (pp. 47-62). Paris: Cambridge University Press.
- Milham, M. P., Banich, M. T., Webb, A., Barad, V., Cohen, N. J., Wszalek, T., et al. (2001). The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on nature of conflict. *Brain Res Cogn Brain Res*, *12*(3), 467-473.
- Miller, B. L., Chang, L., Mena, I., Boone, K., & Lesser, I. M. (1993). Progressive right frontotemporal degeneration: clinical, neuropsychological and SPECT characteristics. *Dementia* *4*, 204–213.
- Miller, E. K., & Cohen, J. D. (2001). An integration of theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202.
- Minter, M., Hobson, R. P., & Bishop, M. (1998). Congenital visual impairment and “theory of mind.” *British Journal of Developmental Psychology*, *16*, 183-196.
- Mitchell, R. W. (1993). Mental models of mirror self-recognition: Two theories. *New Ideas in Psychology*, *11*, 295-325.
- Mithen, S. (2005). *The Prehistory of the Mind: The Cognitive Origins of Art, Religion and Science*. London: Thames and Hudson.
- Moll, J., de Oliveira-Souza, R., & Eslinger, P. J. (2003). Morals and the human brain: a working model. *Neuroreport* *14*, 299–305.

- Moll, J., Zahn, R., de Oliveira-Souza, R., Krueger, F., & Grafman, J. (2005). The neural basis of human moral cognition. *Nature Reviews Neuroscience*, 6, 799-809.
- Molnar-Szakacs, I., Kaplan, J., Greenfield, P. M., & Iacoboni, M. (2006). Observing complex action sequences: The role of the fronto-parietal mirror neuron system. *Neuroimage*.
- Morgan, G., & Kegl, J. (2006). Nicaraguan Sign Language and Theory of Mind: The issue of critical periods and abilities. *Journal of Child Psychology and Psychiatry*, 47, 811-819.
- Morton, J., & Frith, U. (1995). Causal modeling: A structural approach to developmental psychology. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology. Vol. 1: Theory and methods* (pp. 357-390). New York: John Wiley.
- Moses, L. J., Baldwin, D. A., Rosicky, J. G., & Tidball, G. (2001). Evidence for referential understanding in the emotions domain at twelve and eighteen months. *Child Dev*, 72(3), 718-735.
- Mumme, D. L., & Fernald, A. (2003). The infant as onlooker: learning from emotional reactions observed in a television scenario. *Child Dev*, 74(1), 221-237.
- Mundy, P., & Neal, R. (2001). Neural plasticity, joint attention, and a transactional social-orienting model of autism. In L. Masters Glidden (Ed.), *International review of mental retardation: Autism (Vol 23)* (pp. 139-168). San Diego, CA: Academic Press.
- Nelson, C. A. (1987). The recognition of facial expressions in the first two years of life: mechanisms of development. *Child Development*, 58, 889-909.
- Nelson, K. (2004). Toward a collaborative community of minds. *Behavioral and Brain Sciences*, 27, 119-120.
- Nelson, K. (2005). Language pathways into the community of minds. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for theory of mind* (pp. 26-49). Oxford: Oxford University Press.
- Newton, P., Reddy, V., & Bull, R. (2000). Children's everyday deception and performance on false-belief tasks. *British Journal of Developmental Psychology*, 18, 297-317.
- Nimchinsky, E. A., Gilissen, E., Allman, J. M., Perl, D. P., Erwin, J. M., & Hof, P. R. (1999). A neuronal morphologic type unique to humans and great apes. *Proc Natl Acad Sci U S A*, 96(9), 5268-5273.
- Nitschke, J. B., Nelson, E. E., Rusch, B. D., Fox, A. S., Oakes, T. R., & Davidson, R. J. (2004). Orbitofrontal cortex tracks positive mood in mothers viewing pictures of their newborn infants. *Neuroimage*, 21(2), 583-592.
- O'Connor, T. G., Croft, C., & Steele, H. (2000). The contributions of behavioural genetic studies to attachment theory. *Attach Hum Dev*, 2(1), 107-122.
- O'Connor, T. G., & Croft, C. M. (2001). A twin study of attachment in preschool children. *Child Dev*, 72(5), 1501-1511.
- O'Connor, T. G., & Hirsch, N. (1999). Intra-individual differences and relationship-specificity of mentalising in early adolescence. *Social Development*, 8, 256-274.
- Olson, D. R. (1988). On the origins of beliefs and other intentional states in children. In J. W. Astington, P. L. Harris & D. R. Olson (Eds.), *Developing Theories of Mind*. New York: Cambridge University Press.

- Onishi, K., & Baillargeon, R. (2005). Do 15-month-olds understand false beliefs? *Science*, 308, 255-258.
- Onishi, K. H., Baillargeon, R., & Leslie, A. M. (in press). 15-month-old infants detect violations of pretend scenarios. *Acta Psychologica (special issue)*.
- Ontai, L. L., & Thompson, R. A. (2002). Patterns of attachment and maternal discourse effects on children's emotion understanding from 3 to 5 years of age. *Social Development*, 11(4), 433-450.
- Oppenheim, D., & Koren-Karie, N. (2002). Mothers' insightfulness regarding their children's internal worlds: The capacity underlying secure child-mother relationships. *Infant-Mental-Health-Journal*, 23, 593-605.
- Oppenheim, D., Koren-Karie, N., Etzion-Carasso, A., & Sagi-Schwartz, A. (2005, April). *Maternal insightfulness but not infant attachment predicts 4 year olds' theory of mind (poster)*. Paper presented at the Biennial meeting of the Society for Research in Child Development, Atlanta, Georgia.
- Oztop, E., Kawato, M., & Arbib, M. (2006). Mirror neurons and imitation: a computationally guided review. *Neural Netw*, 19(3), 254-271.
- Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. Oxford: Oxford University Press.
- Pears, K. C., & Fisher, P. A. (2005). Emotion understanding and theory of mind among maltreated children in foster care. *Development and Psychopathology*, 17(1), 47-65.
- Pears, K. C., & Moses, L. J. (2003). Demographics, parenting, and theory of mind in preschool children. *Social Development*, 12, 1-20.
- Perner, J., & Lang, B. (2000). Theory of mind and executive function: Is there a developmental relationship? In S. Baron-Cohen, H. Tager-Flusberg & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience* (pp. 150-181). New York: Oxford University Press.
- Perner, J., Ruffman, T., & Leekman, S. R. (1994). Theory of mind is contagious: You catch it from your sibs. *Child Development*, 65, 1228-1238.
- Perry, R. J., & al. (2001). Hemispheric dominance for emotions, empathy and social behaviour: evidence from right and left handers with frontotemporal dementia. *Neurocase* 7, 145-160.
- Peterson, C., & Slaughter, V. (2003). Opening widows into the mind: Mothers' preference for mental state explanations and children's theory of mind. *Cognitive Development*, 18, 399-429.
- Peterson, C. C. (2000). Kindred spirits - Influences of siblings' perspectives on theory of mind. *Cognitive Development*, 15, 435-445.
- Peterson, C. C. (2004). Theory-of-mind development in oral deaf children with cochlear implants or conventional hearing aids. *Journal of Child Psychology and Psychiatry*, 45, 1-11.
- Peterson, C. C., Peterson, J. L., & Webb, J. (2000). Factors influencing the development of a theory of mind in blind children. *British Journal of Developmental Psychology*, 18, 431-447.
- Peterson, C. C., & Slaughter, V. P. (2006). Telling the story of Theory of Mind: Deaf and hearing children's narratives and mental state understanding. *British Journal of Developmental Psychology*, 24, 151-179.
- Petitto, L. A., Holowka, S., Sergio, J. E., & Ostry, D. (2001). Language rhythms in baby hand movements. *Nature*, 413, 35-36.
- Petitto, L. A., & Marentette, P. F. (1991). Babbling in the manual mode: Evidence for the ontogeny of language. *Science*, 251, 1493-1496.

- Phillips, A. T., Wellman, H. M., & Spelke, E. S. (2002). Infants' ability to connect gaze and emotional expression to intentional action. *Cognition*, *85*, 53-78.
- Plotsky, P. M., Thrivikraman, K. V., Nemeroff, C. B., Caldji, C., Sharma, S., & Meaney, M. J. (2005). Long-Term Consequences of Neonatal Rearing on Central Corticotropin-Releasing Factor Systems in Adult Male Rat Offspring. *Neuropsychopharmacology*.
- Pollak, S. D., Klorman, R., Thatcher, J. E., & Cicchetti, D. (2001). P3b reflects maltreated children's reactions to facial displays of emotion. *Psychophysiology*, *38*(2), 267-274.
- Pollak, S. D., & Sinha, P. (2003). Effects of early experience on children's recognition of facial displays of emotion. *Developmental Psychology*, *38*(5), 784-791.
- Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. *Dev Psychopathol*, *17*(3), 715-734.
- Posner, M. I., & Rothbart, M. K. (2000). Developing mechanisms of self-regulation. *Development and Psychopathology*, *12*, 427-441.
- Premack, D. (1984). Pedagogy and aesthetics as sources of culture. In M. S. Gazzaniga (Ed.), *Handbook of cognitive neuroscience* (pp. 15-35). New York: Plenum Press.
- Premack, D., & Premack, A. (2003). *Original intelligence. Unlocking the mystery of who we are*. New York: McGraw-Hill.
- Raikes, H. A., & Thompson, R. A. (2006). Family emotional climate, attachment security, and young children's emotion knowledge in a high-risk sample. *British Journal of Developmental Psychology*, *24*(1), 89-104.
- Reddy, V., & Morris, P. (2004). Participants don't need theories: Knowing minds in engagement. *Theory and Psychology*, *14*(5), 649-667.
- Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2004). The neural correlates of theory of mind within interpersonal interactions. *Neuroimage*, *22*(4), 1694-1703.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annu Rev Neurosci*, *27*, 169-192.
- Rizzolatti, G., Ferrari, P. F., Rozzi, S., & Fogassi, L. (2006). The inferior parietal lobule: where action becomes perception. *Novartis Found Symp*, *270*, 129-140; discussion 140-125, 164-129.
- Rochat, P., & Striano, T. (1999). Social-cognitive development in the first year. In P. Rochat (Ed.), *Early social cognition*. Mahwah, NJ.: Lawrence Erlbaum.
- Rogosch, F. A., Cicchetti, D., & Aber, J. L. (1995). The role of child maltreatment in early deviations in cognitive and affective processing abilities and later peer relationship problems. *Development and Psychopathology*, *7*, 591-609.
- Ruffman, T., Perner, J., Naito, M., Parkin, L., & Clements, W. (1998). Older (but not younger) siblings facilitate false belief understanding. *Developmental Psychology*, *34*(1), 161-174.
- Ruffman, T., Perner, J., & Parkin, L. (1999). How parenting style affects false belief understanding. *Social Development*, *8*, 395-411.
- Ruffman, T., Slade, L., Rowlandson, K., Rumsey, C., & Garnham, A. (2003). How language relates to belief, desire, and emotion understanding. *Cognitive Development*, *18*, 139-158.

- Saarela, M. V., Hlushchuk, Y., Williams, A. C., Schurmann, M., Kalso, E., & Hari, R. (2006). The Compassionate Brain: Humans Detect Intensity of Pain from Another's Face. *Cereb Cortex*.
- Sabbagh, M. A., & Callanan, M. A. (1998). Metarepresentation in action: 3-, 4-, and 5-year-olds' developing theories of mind in parent-child conversations. *Developmental Psychology*, *34*, 491-502.
- Safran, J. D., & Muran, J. C. (1996). The resolution of ruptures in the therapeutic alliance. *J Consult Clin Psychol*, *64*(3), 447-458.
- Sato, W., & Yoshikawa, S. (2006). Spontaneous facial mimicry in response to dynamic facial expressions. *Cognition*.
- Satpute, A. B., & Lieberman, M. D. (2006). Integrating automatic and controlled processes into neurocognitive models of social cognition. *Brain Res*, *1079*(1), 86-97.
- Saxe, R. (2005). Against simulation: the argument from error. *Trends Cogn Sci*, *9*(4), 174-179.
- Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding other minds: linking developmental psychology and functional neuroimaging. *Annu Rev Psychol*, *55*, 87-124.
- Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind". *Neuroimage*, *19*(4), 1835-1842.
- Saxe, R., & Wexler, A. (2005). Making sense of another mind: the role of the right temporo-parietal junction. *Neuropsychologia*, *43*(10), 1391-1399.
- Scaife, M., & Bruner, J. (1975). The capacity for joint visual attention in the infant. *Nature*, *253*, 265-266.
- Schechter, D. S., Coats, T., Zeanah, C. H., Davies, M., Coates, S., Trabka, K., et al. (2005). Maternal mental representations of the child in an inner-city clinical sample: Violence-related posttraumatic stress and reflective functioning. *Attachment and Human Development*, *7*(3), 313-331.
- Schilbach, L., Wohlschlaeger, A. M., Kraemer, N. C., Newen, A., Shah, N. J., Fink, G. R., et al. (2006). Being with virtual others: Neural correlates of social interaction. *Neuropsychologia*, *44*(5), 718-730.
- Schneider-Rosen, K., & Cicchetti, D. (1991). Early self-knowledge and emotional development: Visual self-recognition and affective reactions to mirror self-image in maltreated and non-maltreated toddlers. *Developmental Psychology*, *27*, 481-488.
- Schultz, R. T., Grelotti, D. J., Klin, A., Kleinman, J., Van der Gaag, C., Marois, R., et al. (2003). The role of the fusiform face area in social cognition: implications for the pathobiology of autism. *Philos Trans R Soc Lond B Biol Sci*, *358*(1430), 415-427.
- Schutz-Bosbach, S., Mancini, B., Aglioti, S. M., & Haggard, P. (2006). Self and other in the human motor system. *Curr Biol*, *16*(18), 1830-1834.
- Semendeferi, K., & Damasio, H. (2000). The brain and its main anatomical subdivisions in living hominoids using magnetic resonance imaging. *Journal of Human Evolution*, *38*, 317-332.
- Sharp, C., & Fonagy, P. (submitted). The parent's capacity to treat the child as a psychological agent: Constructs, measures and implications for developmental psychopathology.
- Sharp, C., Fonagy, P., & Goodyer, I. (2006). Imagining your child's mind: Psychosocial adjustment and mothers' ability to predict their children's

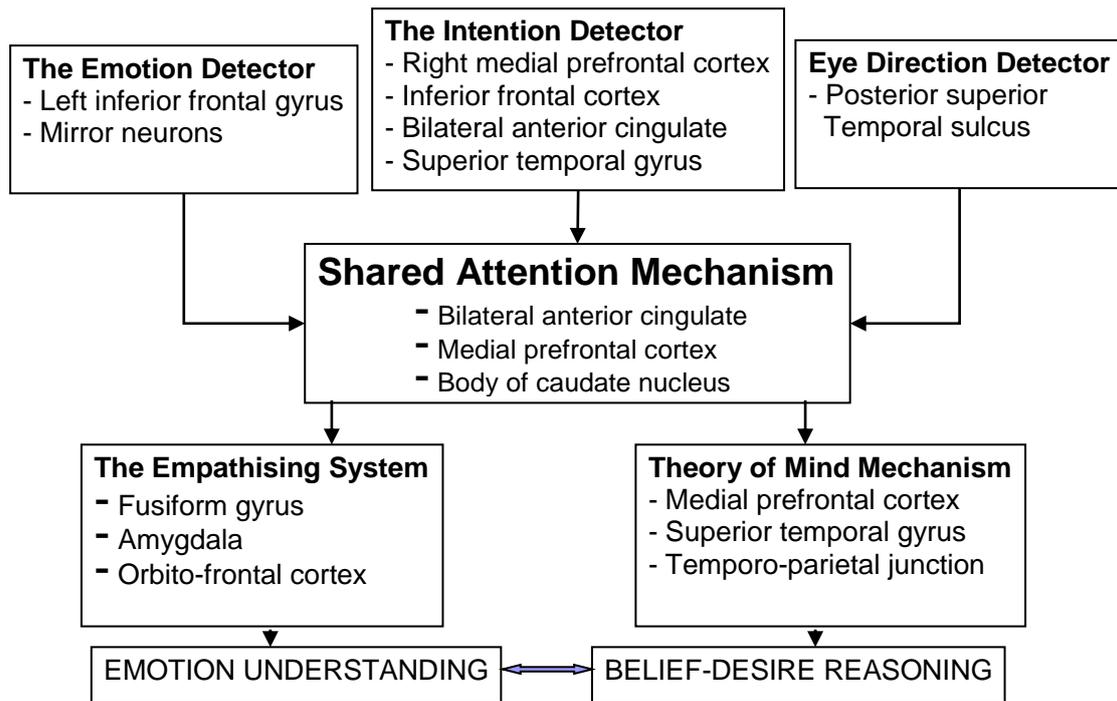
- attributional response styles. *British Journal of Developmental Psychology*, 24(1), 197-214.
- Shatz, M., & al. (2003). The influence of language and socioeconomic status on children's understanding of false beliefs. *Developmental Psychology*, 39, 717-729.
- Shipman, K. L., & Zeman, J. (1999). Emotional understanding: a comparison of physically maltreating and nonmaltreating mother-child dyads. *J Clin Child Psychol*, 28, 407-417.
- Siegal, M., & Patterson, C. C. (in press). Language and theory of mind in atypically developing children: Evidence from studies of deafness, blindness, and autism. In C. Sharp, P. Fonagy & I. Goodyer (Eds.), *Social cognition and developmental psychology*. Oxford: Oxford University Press.
- Siegal, M., & Varley, R. (2002). Neural systems involved in "theory of mind". *Nat Rev Neurosci*, 3(6), 463-471.
- Siegel, D. J. (1999). *The developing mind: Toward a neurobiology of interpersonal experience*. New York: Guilford.
- Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, 303(5661), 1157-1162.
- Slade, A. (2005). Parental reflective functioning: An introduction. *Attachment and Human Development*, 7(3), 269-281.
- Slade, A., Grienenberger, J., Bernbach, E., Levy, D., & Locker, A. (2005). Maternal reflective functioning, attachment and the transmission gap: A preliminary study. *Attachment and Human Development*, 7(3), 283-298.
- Slomkowski, C., & Dunn, J. (1996). Young children's understanding of other people's beliefs and feelings and their connected communication with friends. *Developmental Psychology*, 32, 442-447.
- Smith, M., & Walden, T. (1999). Understanding feelings and coping with emotional situations: A comparison of maltreated and nonmaltreated preschoolers. . *Social Development*, 8(1), 93-116.
- Sodian, B., & Thoermer, C. (2004). Infants' understanding of looking, pointing, and reaching as cues to goal-directed action. *Journal of Cognition and Development*, 5, 289-316.
- Sommerville, J. A., & Woodward, A. L. (2005). Pulling out the structure of intentional action: The relation between action processing and production in infancy. *Cognition*, 95, 1-30.
- Southgate, V., Senju, A., & Csibra, G. (in press-a). Action anticipation through attribution of false belief by two-year-olds. *Psychological Science*.
- Southgate, V., Senju, A., & Csibra, G. (in press-b). Anticipatory looking reveals attributions of false belief by two-year-olds. *Psychological Science*.
- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and Cognition*. Oxford: Blackwell.
- Sroufe, L. A. (1979). Socioemotional development. In J. Osofsky (Ed.), *Handbook of Infant Development* (pp. 462-516). New York: Wiley.
- Sroufe, L. A. (1996). *Emotional development: The organization of emotional life in the early years*. New York: Cambridge University Press.
- Stams, G. J. M., Juffer, F., & van IJzendoorn, M. H. (2002). Maternal sensitivity, infant attachment, and temperament in early childhood predict adjustment in middle childhood: The case of adopted children and their biologically unrelated parents. *Developmental Psychology*, 38(5), 806-821.

- Stanley, C., Murray, L., & Stein, A. (2004). The effect of postnatal depression on mother-infant interaction, infant response to the Still-face perturbation, and performance on an Instrumental Learning task. *Development & Psychopathology*, *16*(1), 1-18.
- Steele, M., Steele, H., Croft, C., & Fonagy, P. (1999). Infant mother attachment at one year predicts children's understanding of mixed emotions at 6 years. *Social Development*, *8*, 161-178.
- Stern, D. N. (1985). *The Interpersonal World of the Infant: A View from Psychoanalysis and Developmental Psychology*. New York: Basic Books.
- Stern, D. N. (1995). Self/other differentiation in the domain of intimate socio-affective interaction: Some considerations. In P. Rochat (Ed.), *The self in infancy: Theory and research* (pp. 419-429). Amsterdam: Elsevier.
- Striedter, G. F. (2006). Precis of principles of brain evolution. *Behav Brain Sci*, *29*(1), 1-12; discussion 12-36.
- Surian, L., Caldi, S., & Sperber, D. (in press). Attribution of beliefs by 13-month-old infants. *Psychological Science*.
- Sutton, J., Smith, P. K., & Swettenham, J. (1999a). Bullying and 'theory of mind': A critique of the 'social skills deficit' view of anti-social behaviour. *Social Development*, *8*, 117-127.
- Sutton, J., Smith, P. K., & Swettenham, J. (1999b). Social cognition and bullying: Social inadequacy or skilled manipulation? *British Journal of Developmental Psychology*, *17*, 435-450.
- Swain, J. E., Lorberbaum, J. P., Kose, S., & Strathearn, L. (in press). Brain basis of early parent-infant interactions: psychology, physiology, and in vivo functional neuroimaging studies. *Journal of Child Psychology and Psychiatry*.
- Sylvester, C. Y., Wager, T. D., Lacey, S. C., Hernandez, L., Nichols, T. E., Smith, E. E., et al. (2003). Switching attention and resolving interference: fMRI measures of executive functions. *Neuropsychologia*, *41*(3), 357-370.
- Symons, D. K., 24, . (2004). Mental state discourse, theory of mind, and the internalization of self-other understanding. *Developmental Review*, *24*, 159-188.
- Tager-Flusberg, H. (2001). A re-examination of the Theory of Mind hypothesis of autism. In J. Burack, T. Charman, N. Yirmiya & P. Zelazo (Eds.), *The development of autism: Perspectives from theory and research* (pp. 173-194). Mahwah, NJ: Lawrence Erlbaum Associates.
- Taylor, M., & Carlson, S. M. (1997). The relation between individual differences in fantasy and theory of mind. *Child Dev*, *68*(3), 436-455.
- Taylor, M., Esbensen, B. M., & Bennett, R. T. (1994). Children's understanding of knowledge acquisition: the tendency for children to report that they have always known what they have just learned. *Child Dev*, *65*(6), 1581-1604.
- Thompson, R. A. (2000). The legacy of early attachments. *Child Development*, *71*, 145-152.
- Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. Dunham (Eds.), *Joint Attention: Its Origins and Role in Development* (pp. 103-130). New York, N.Y.: Lawrence Erlbaum.
- Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.
- Tomasello, M., Call, J., & Hare, B. (2003). Chimpanzees understand psychological states - the question is which ones and to what extent. *Trends Cogn Sci*, *7*(4), 153-156.

- Tomasello, M., Carpenter, M., Call, J., Behne, T., & H., M. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28(5), 675-691; discussion 691-735.
- Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural learning. *Behavioral and Brain Sciences*, 16, 495-552.
- Toth, S. L., Cicchetti, D., Macfie, J., Maughan, A., & Vanmeenen, K. (2000). Narrative representations of caregivers and self in maltreated pre-schoolers. *Attachment and Human Development*, 2, 271-305.
- Toth, S. L., Maughan, A., Manly, J. T., Spagnola, M., & Cicchetti, D. (2002). The relative efficacy of two interventions in altering maltreated preschool children's representational models: implications for attachment theory. *Dev Psychopathol*, 14(4), 877-908.
- Tranel, D., Bechara, A., & Denburg, N. L. (2002). Asymmetric functional roles of right and left ventromedial prefrontal cortices in social conduct, decision-making, and emotional processing. *Cortex* 38, 589-612.
- Trevarthen, C. (1979). Communication and cooperation in early infancy: A description of primary intersubjectivity. In M. M. Bullowa (Ed.), *Before Speech: The Beginning of Interpersonal Communication* (pp. 321-347). New York: Cambridge University Press.
- Trevarthen, C. (1993). The self born in intersubjectivity: an infant communicating. In U. Neisser (Ed.), *The Perceived Self* (pp. 121-173). New York: Cambridge University Press.
- Trevarthen, C. (2001). The neurobiology of early communication: Intersubjective regulations in human brain development. In F. A. Kalverboer & A. Gramsbergen (Eds.), *Handbook on Brain and Behavior in Human Development* (pp. 841-882). Dordrecht: Kluwer Academic.
- Trevarthen, C. (2005). Stepping away from the mirror: Pride and shame in adventures of companionship - Reflections on the nature and emotional needs of infant intersubjectivity. In C. S. Carter, L. Ahnert, K. E. H. Grossman, S B, M. E. Lamb, S. W. Porges & al (Eds.), *Attachment and Bonding: A New Synthesis (Dahlem Workshop Report 92)* (pp. 55-84). Cambridge, MA: MIT Press.
- Trevarthen, C., & Aitken, K. J. (1994). Brain development, infant communication, and empathy disorders: Intrinsic factors in child mental health. Special issue: Neural plasticity, sensitive periods, and psychopathology. *Development and Psychopathology*, 6, 597-633.
- Trevarthen, C., & Aitken, K. J. (2001). Infant intersubjectivity: research, theory, and clinical applications. *Journal of Child Psychology and Psychiatry*, 42, 3-48.
- Trevarthen, C. A., K J, Vandekerckhove, M., Delafield-Butt, J., & Nagy, E. (2006). Collaborative regulations of vitality in early childhood: Stress in intimate relationships and postnatal psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental Psychopathology (second edition). Vol. 2: Developmental Neuroscience* (pp. 65-126). New York: Wiley.
- Tronick, E. (1989). Emotions and emotional communication in infants. *American Psychologist*, 44, 112-119.
- Tronick, E., 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 Child Psychiatry*, 17, 1-13.
- Tronick, E. Z. (2005). Why is connection with others so critical? The formation of dyadic states of consciousness: Coherence governed selection and the co-creation of meaning out of messy meaning making. In J. Nadel & D. Muir

- (Eds.), *Emotional Development* (pp. 293-315). Oxford: Oxford University Press.
- Tröster, H., & Bambring, M. (1994). Play behavior and play materials in blind and sighted infants and preschoolers. *Journal of Visual Impairment and Blindness*, 88, 421-433.
- Tröster, H., & Bambring, M. (1992). Early social-emotional development in blind infants. *Child: Care, Health and Development*, 18, 207-227.
- Tulving, E. (2002). Episodic memory: from mind to brain. *Annu Rev Psychol*, 53, 1-25.
- Turati, C., Simion, F., Milini, I., & Umiltà, C. (2002). Newborns' preference for faces: What is crucial? *Developmental Psychology*, 38, 875-882.
- Tzourio-Mazoyer, N., De Schonen, S., Crivello, F., Reutter, B., Aujard, Y., & Mazoyer, B. (2002). Neural correlates of woman face processing by 2-month-old infants. *Neuroimage*, 15(2), 454-461.
- van IJzendoorn, M. H. (1995). Adult attachment representations, parental responsiveness, and infant attachment: A meta-analysis on the predictive validity of the Adult Attachment Interview. *Psychological Bulletin*, 117, 387-403.
- van IJzendoorn, M. H., & van Vliet-Visser, S. (1988). The relationship between quality of attachment in infancy and IQ in kindergarten. *J Genet Psychol*, 149(1), 23-28.
- Vinden, P. G. (2001). Parenting attitudes and children's understanding of mind: A comparison of Korean American and Anglo-American families. *Cognitive Development*, 16, 793-809.
- Vogele, K., Bussfeld, P., Newen, A., Herrmann, S., Happe, F., Falkai, P., et al. (2001). Mind reading: neural mechanisms of theory of mind and self-perspective. *Neuroimage*, 14(1 Pt 1), 170-181.
- Vollm, B. A., Taylor, A. N., Richardson, P., Corcoran, R., Stirling, J., McKie, S., et al. (2006). Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. *Neuroimage*, 29(1), 90-98.
- Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, Mass: Harvard University Press.
- Wagner, L., & Carey, S. (2005). 12-month-old infants represent probable ending of motion events. *Infancy*, 7, 73-83.
- Watson, J. S. (1972). Smiling, cooing, and "the game". *Merrill-Palmer Quarterly*, 18, 323-339.
- Watson, J. S. (1994). Detection of self: The perfect algorithm. In S. Parker, R. Mitchell & M. Boccia (Eds.), *Self-Awareness in Animals and Humans: Developmental Perspectives* (pp. 131-149): Cambridge University Press.
- Watson, J. S. (2001). Contingency perception and misperception in infancy: Some potential implications for attachment. *Bulletin of the Menninger Clinic*, 65, 296-320.
- Wegner, D. M., & Wheatley, T. (1999). Apparent mental causation: Sources of the experience of will. *American Psychologist*, 54(7), 480-492.
- Wellman, H. (1990). *The Child's Theory of Mind*. Cambridge, Mass: Bradford Books/MIT Press.
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Dev*, 72(3), 655-684.

- Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Dev*, 75(2), 523-541.
- Werner, H., & Kaplan, B. (1963). *Symbol formation*. New York: Wiley.
- Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., & Rizzolatti, G. (2003). Both of us disgusted in My insula: the common neural basis of seeing and feeling disgust. *Neuron*, 40(3), 655-664.
- Wildgruber, D., Riecker, A., Hertrich, I., Erb, M., Grodd, W., Ethofer, T., et al. (2005). Identification of emotional intonation evaluated by fMRI. *Neuroimage*, 24(4), 1233-1241.
- Williams, J. H., Waiter, G. D., Gilchrist, A., Perrett, D. I., Murray, A. D., & Whiten, A. (2006). Neural mechanisms of imitation and 'mirror neuron' functioning in autistic spectrum disorder. *Neuropsychologia*, 44(4), 610-621.
- Williams, J. H., Waiter, G. D., Perra, O., Perrett, D. I., & Whiten, A. (2005). An fMRI study of joint attention experience. *Neuroimage*, 25(1), 133-140.
- Williams, L., Fonagy, P., Target, M., Fearon, P., Sargent, J., Bleiberg, E., et al. (2006). Training psychiatry residents in Mentalization-Based Therapy. In J. Allen & P. Fonagy (Eds.), *Handbook of mentalisation based treatments*. London: John Wiley.
- Winston, J. S., Strange, B. A., O'Doherty, J., & Dolan, R. J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nat Neurosci*, 5(3), 277-283.
- Wittgenstein, L. (1969). *The Blue and Brown Books*. Oxford: Blackwell.
- Wood, J. N., & Grafman, J. (2003). Human prefrontal cortex: processing and representational perspectives. *Nature Rev. Neurosci.*, 4, 139-147.
- Woodward, A. L. (2005). The infant origins of intentional understanding. In R. V. Kail (Ed.), *Advances in child development and behavior (Vol. 33)* (pp. 229-262). Amsterdam: Elsevier.
- Woolfe, T., Want, S. C., & Siegal, M. (2002). Signposts to development: Theory of mind in deaf children. *Child Development*, 73, 768-778.
- Zhang, T. Y., Chretien, P., Meaney, M. J., & Gratton, A. (2005). Influence of naturally occurring variations in maternal care on prepulse inhibition of acoustic startle and the medial prefrontal cortical dopamine response to stress in adult rats. *J Neurosci*, 25(6), 1493-1502.
- Zilles, K., Dabringhaus, A., Geyer, S., Amunts, K., Qu, M., Schleicher, A., et al. (1996). Structural asymmetries in the human forebrain and the forebrain of non-human primates and rats. *Neurosci Biobehav Rev*, 20(4), 593-605.



*Legend for Figure 1:*

A schematised version of Baron-Cohen's (2005) model of the social brain. The Emotion Detector, Intention Detector and Eye Direction Detector may be in place from birth and are thought to be fully functioning before 9 months, the Shared Attention Mechanism by 2nd half of first year, emotion understanding and belief-desire reasoning requiring appreciation of false beliefs, pretence and differentiation of self and other's perspective is not fully functional until 3-4 years although implicit emotion understanding belief desire reasoning appears to be in the child's repertoire during the 2<sup>nd</sup> year.