

## **THE MENTALIZATION-FOCUSED APPROACH TO SELF PATHOLOGY**

Peter Fonagy, PhD, FBA, and Mary Target, PhD

### **1. INTRODUCTION**

Our approach to understanding self pathology in personality disorder assumes that the capacity to mentalize, that is, the capacity to conceive of mental states as explanations of behavior in oneself and in others, is a key determinant of self-organization. Along with contributory capacities of affect regulation and attention control mechanisms, the capacity for mentalization is acquired in the context of early attachment relationships. Disturbances of attachment relationships will therefore disrupt the normal emergence of these key social-cognitive capacities and create profound vulnerabilities in the context of social relationships. Ours is fundamentally a psychoanalytic approach but we have elaborated our model of social development on the basis of empirical observations as well as clinical work.

We define mentalization as a form of mostly preconscious imaginative mental activity, namely, perceiving and interpreting human behavior in terms of intentional mental states (e.g., needs, desires, feelings, beliefs, goals, purposes, and reasons). Mentalizing is imaginative because we have to imagine what other people might be thinking or feeling; an important indicator of high quality of mentalization is the awareness that we cannot know absolutely what is in someone else's mind. We suggest that a similar kind of imaginative leap is required to understand one's own mental experience, particularly in relation to emotionally charged issues. In order to conceive of others as having a mind, the individual needs a symbolic representational system for mental states and also must be able to selectively activate states of mind in line with particular intentions, which requires attentional control.

---

From the The Anna Freud Centre, London.

Address correspondence to Peter Fonagy, Sub-Department of Clinical Health Psychology, University College London, Gower Street, London WC1E 6BT; E-mail: p.fonagy@ucl.ac.uk

The work summarized in this paper is the result of the collaborative effort of a group of wonderfully talented individuals who have honored the author with their friendship over the past years. The ideas originate in collaborative work with the late George Moran, PhD, and were enriched and extended in collaboration with George Gergely, PhD. Major research contributions were made by Miriam Steele, PhD and Howard Steele, PhD, Helen Stein, PhD, and John Allen, PhD. The clinical implications of the approach were delineated by two exceptionally skilled psychiatrists, Efrain Bleiberg, MD and Anthony Bateman, MD. The coherence of the presentation of various versions of the work was ensured by Liz Allison, DPhil.

The ability to understand the self as a mental agent grows out of interpersonal experience, particularly primary object relationships (Fonagy, 2003). The baby's experience of himself as having a mind or self is not a genetic given; it evolves from infancy through childhood, and its development critically depends upon interaction with more mature minds, assuming these are benign, reflective, and sufficiently attuned. Mentalization involves both a self-reflective and an interpersonal component. It is underpinned by a large number of specific cognitive skills, including an understanding of emotional states, attention and effortful control, and the capacity to make judgments about subjective states as well as thinking explicitly about states of mind—what we might call mentalization proper. In combination, these functions enable the child to distinguish inner from outer reality and internal mental and emotional processes from interpersonal events.

This paper addresses the complex relation of attachment and mentalization. We discuss the role of mentalizing in the development of the agentic sense of self, and consider the contribution of attachment trauma to the development of psychopathology by virtue of undermining mentalizing capacity.

## **2. EVOLUTIONARY AND NEUROBIOLOGICAL LINKS BETWEEN ATTACHMENT AND MENTALIZATION**

### THE SELECTIVE ADVANTAGES OF ATTACHMENT

As we understand more about the interface of brain development and early psychosocial experience, it becomes clear that the evolutionary role of the attachment relationship goes far beyond giving physical protection to the human infant. Attachment ensures that the brain processes that come to subserve social cognition are appropriately organized and prepared to equip the individual for the collaborative and cooperative existence with others for which the brain was designed.

In our view the major selective advantage conferred by attachment to humans is the opportunity to develop social intelligence that nearness to concerned adults affords. Alan Sroufe (1996) and Myron Hofer (2004) played a seminal role 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 an organizer of physiological and brain regulation. More recent work has begun to articulate the associated biological pathways (e.g., Champagne et al., 2004; 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 are influenced by the infant's environment. The brain is experience-expectant (Siegel, 1999).

### THE SELECTIVE ADVANTAGES OF MENTALIZATION

Mentalization is arguably the evolutionary pinnacle of human intellectual achievement. But what has driven the selection processes of the two million or so years of human evolution towards a consciousness of mental states in self and 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 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., climatic variability, threat of predation, and availability of prey).

The evolutionary biologist Richard Alexander (1989) proposed a widely accepted model of how humans evolved their minds. He suggested that our exceptional intelligence evolved not to deal with the hostile forces of nature, but rather to deal with competition from other people. This further evolution occurred only *after* our species had already achieved relative dominance over their environment. At that point 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 cooperate. A kind of evolutionary arms race probably took place among ever more effective social groups. Skill in understanding and outsmarting other people was required to gain the upper hand. Thus, the ability to construct and manipulate thoughts about thoughts and feelings acquired a major reproductive advantage. The possibility of interpreting and anticipating behavior permits cooperation, offers competitive advantage, and continually selects for increasingly higher levels of social interpretive capacity.

### THE INTERPERSONAL INTERPRETIVE FUNCTION

The capacity to interpret human behavior (see Bogdan, 1997) requires the intentional stance: “treating the object whose behavior you want to predict as a rational agent with beliefs and desires” (Dennett, 1987 p. 15). We label the capacity to adopt this stance the interpersonal interpretive function (IIF), an evolutionary-developmental function of attachment. The IIF is a cluster of mental functions for processing and interpreting new interpersonal experiences that includes mentalization and the cluster of psychological processes on which effective mentalizing depends (Fonagy, 2003). Four emotional processing and control mechanisms contribute to the developmental unfolding of interpretive function: labelling and understanding affect, arousal regulation, effortful control, and specific mentalizing capacities (Fonagy & Target, 2002).

Since the mind needs to adapt to ever more challenging competitive conditions, the capacity for mentalization cannot be fixed by genetics or constitution. The social brain must continuously reach higher and higher levels of sophistication to stay on top. Evolution has charged attachment

relationships with ensuring the full development of the social brain. The capacity for mentalization, along with many other social-cognitive capacities, evolves out of the experience of social interaction with caregivers. Increased sophistication in social cognition evolved hand-in-hand with apparently unrelated aspects of development, such as increased helplessness in infancy, a prolongation of childhood, and the emergence of intensive parenting.

We have proposed a mechanism for this process rooted in dialectic models of self-development (Cavell, 1991; Davidson, 1983). Our approach explicitly rejects the classical Cartesian assumption that mental states are apprehended by introspection; on the contrary, mental states are discovered through contingent mirroring interactions with the caregiver (Gergely & Watson, 1999). Therefore, early disruption of affectional bonds will not only set up maladaptive attachment patterns (e.g., Waters, Merrick, Treboux, Crowell, & Albersheim, 2000), but will also undermine a range of capacities vital to normal social development. Understanding minds is difficult if one does not know what it is like to be understood as a person with a mind. Our argument may seem to place an excessive burden upon the caregiver-infant relationship, but we must remember that placing the social development of a human infant in the hands of one adult is a recent phenomenon compared to the previous average of four relatives who had a genetic stake in the child's survival (Hrdy, 2000). Recent neurobiological evidence discussed next buttresses the ecological view of attachment relationships as pivotally linked to mentalizing capacities.

#### THE NEUROBIOLOGY OF ATTACHMENT

The neurobiology of attachment is now fairly well understood. It is linked to the mesocorticolimbic dopaminergic reward circuit, which also plays a key role in mediating the process of physical (as well as emotional) addiction. It is highly unlikely that nature created a brain system specifically to subserve cocaine and alcohol abuse. It is more likely that addictions are the accidental by-product of the activation of a biological system underpinning the crucial evolutionary function of attachment (Insel, 1997; MacLean, 1990; Panksepp, 1998). Attachment can be thought of as an "addictive disorder" (Insel, 2003). Changes in attachment behavior, such as falling in love, which are stimulated by social/sexual activity, entail the activation of an oxytocin and vasopressin sensitive circuit within the anterior hypothalamus (MPOA) linked to the VTA and the nucleus accumbens (Insel, 2003). fMRI studies indicate specific activation of the same pathways in the brain of somebody seeing their own baby or partner, compared to another familiar baby or other people's partners (Nitschke et al., 2004).

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 brain regions in two systems both responsible for different as-

pects of cognitive regulation and control, but also including those associated with making social judgments and mentalizing. Bartels and Zeki (2004) suggest grouping these reciprocally active areas into two functional regions. The first (let us refer to it as system A) includes the middle prefrontal, inferior parietal and middle temporal cortices mainly in the right hemisphere, as well as the posterior cingulate cortex. These areas are specialized for attention and long-term memory (Cabeza & Nyberg, 2000), and have variable involvement in both positive (Maddock, 1999) and negative (Mayberg et al., 1999) emotions. Their role in both cognition and emotion suggests that these areas may be specifically responsible for integrating emotion and cognition (e.g., emotional encoding of episodic memories). Further, these areas may play a role in recalling emotion-related material and generating emotion-related imagery that may be relevant in relation to understanding the typology of attachment (Maddock, 1999).

The second set of areas deactivated by the activation of the attachment system includes the temporal poles, parietotemporal junction, amygdala, and mesial prefrontal cortex (let us call this system B). Activation of these areas is consistently linked to negative affect, judgments of social trustworthiness, moral judgments, "theory of mind" tasks, attention to one's own emotions, and in particular, they constitute the primary neural network underlying our ability to identify mental states (both thoughts and feelings) in other people (Frith & Frith, 2003; Gallagher & Frith, 2003). Mentalization pertains not just to states of mind in others but also reflecting on one's own emotional and belief states and consequently such tasks appear to be associated with activation in the same neural system (Gusnard, Akbudak, Shulman, & Raichle, 2001). Making judgments that involve mental states has been shown to be associated with activation of the same system. Thus, intuitive judgments of moral appropriateness (rather than moral reasoning) are linked (Greene & Haidt, 2002) as is assessment of social trustworthiness based on facial expressions (Winston, Strange, O'Doherty, & Dolan, 2002).

This suggests that being in an emotionally attached state inhibits aspects of social cognition, including mentalizing and the capacity accurately to see the attachment figure as a person. (Currently we are working to perform an independent replication of this study). The activation of the attachment system, mediated by dopaminergic structures of the reward system in the presence of oxytocin and vasopressin, inhibits neural systems that underpin the generation of negative affect. This is to be expected since a key function of the attachment system is to moderate negative emotions in the infant and presumably to continue to do so in later development (Sroufe, 1996). Equally consistent with expectations, is the suppression of social and moral judgments (probably mediated by the second of the two regulatory systems) associated with the activation of the attachment system. Judgments of social trustworthiness and morality serve to distance us from others, but become less relevant and may indeed inter-

ferre with our relationships with those to whom we are strongly attached (Belsky, 1999a; Simpson, 1999).

The configuration described by Bartels and Zeki has critical developmental implications. Attachment has been selected by evolution as the principal “training ground” for the acquisition of mentalization because attachment is a marker for shared genetic material, reciprocal relationships, and altruism. It is a non-competitive relationship in which the aim is not to outsmart and thus, learning about minds can be safely practiced. Missing out on early attachment experience (as for the Romanian orphans) creates a long term vulnerability from which the child may never recover—the capacity for mentalization is never fully established, leaving the child vulnerable to later trauma and unable to cope fully with attachment relationships (e.g., Rutter & O’Connor, 2004). More importantly, trauma, by activating attachment will often decouple the capacity for mentalization. This, of course, is further exacerbated when the trauma is attachment trauma.

#### IMPLICATIONS OF ATTACHMENT-MENTALIZATION RECIPROCITY

The apparently reciprocal relationship of mentalization and attachment may at first appear to contradict our earlier assumption that mentalization and secure attachment are positively correlated. Further scrutiny suggests greater complexity but no inconsistency. It is possible, taking an evolutionary perspective, that the parent’s capacity to mentalize the infant or child serves to reduce the child’s experienced need to monitor the parent for trustworthiness. This relaxation of the interpersonal barrier facilitates the emergence of a strong attachment bond. While at first sight the precocious emergence of theory of mind in children who were securely attached in infancy (e.g., Meins, 1997) may seem inconsistent with the inverse relationship between attachment and mentalization, it is to be expected that in individuals whose attachment is secure, there are likely to be fewer calls over time for the activation of the attachment system. This in turn, given the inhibitory effect of the activation of the attachment system on mentalization related brain activity, might account for the precocious development of mentalization.

The capacity for mentalization in the context of attachment is likely to be in certain respects independent of the capacity to mentalize about interpersonal experiences outside the attachment context (Fonagy & Target, 1997). Our specific measure of mentalization in the attachment context, reflective function (Fonagy, Target, Steele, & Steele, 1998) is predictive of behavioral outcomes that other measures of mentalization do not correlate with. For example, in a quasi-longitudinal study based on interviews and chart reviews with young adults, some of whom had suffered trauma, we found that the impact of trauma on mentalization in attachment contexts mediated outcome measured as the quality of adult romantic relation-

ships, but mentalization measured independently of the attachment context using the Reading the Mind in the Eyes test did not (Fonagy, Stein, Allen, & Fultz, 2003a). It seems that measuring mentalization in the context of attachment might measure a unique aspect of social behavior.

The key consideration is probably that securely attached children do not need to activate their attachment system as often and have greater opportunity to “practice” mentalization in the context of the child-caregiver relationship. Belsky’s (1999b) evolutionary model of attachment classification is helpful here. When resources are scarce and insecure attachment strategies are possibly most adaptive, children probably need to monitor the unpredictable caregivers’ mental states quite carefully, are forced to find alternative social contexts to acquire social cognitive capacities, and thus they deprive themselves of some developmental learning opportunities of understanding minds in abstract ways independent of physical reality.

### **3. THE DEVELOPMENT OF AN AGENTIVE SELF: THE SOCIAL ACQUISITION OF SOCIAL COGNITION**

#### **AN OVERVIEW OF THE MODEL OF CONTINGENT MIRRORING**

The evolutionary neurobiological speculations above imply that children’s caregiving environments play a key role in their development as social beings. How are we to conceive of the actions of these environmental influences? Our model relies on the child’s inbuilt capacity to detect aspects of his world that react contingently to his own actions. In his first months the child begins to understand that he is a physical agent whose actions can bring about changes in bodies with which he has immediate physical contact (Leslie, 1994). Developing alongside this is the child’s understanding of himself as a social agent. Through interactions with the caregiver (from birth) the baby learns that his behavior affects his caregiver’s behavior and emotions (Neisser, 1988). Both these early forms of self-awareness probably evolve through the workings of an innate contingency detection mechanism that enables the infant to analyze the probability of causal links between his actions and stimulus events (Watson, 1994). The child’s initial preoccupation with perfectly response-contingent stimulation (provided by the proprioceptive sensory feedback that the self’s actions always generate) allows him to differentiate his agentive self as a separate entity in the environment and to construct a primary representation of the bodily self.

At about 3–4 months, infants switch from preferring perfect contingency to preferring high, but imperfect contingencies thereafter (Bahrick & Watson, 1985)—the level of contingency that is characteristic of an attuned caregiver’s empathic mirroring responses to the infant’s displays of emotion. Repeated experience of such affect-reflective caregiver reactions is essential for the infant to begin to be able to differentiate his/her internal self-states: a process we termed “social biofeedback” (Gergely & Watson,

1996). A congenial and secure attachment relationship can vitally contribute to the emergence of early mentalization capacities allowing the infant to “discover” or “find” his/her psychological self in the social world (Gergely, 2001). The discovery of the representational or psychological self (what we may think of as full mentalization) is probably based in the same mechanism.

#### COMING TO UNDERSTAND AND REGULATE EMOTION AND BE SECURELY ATTACHED

Let us take the development of an understanding of affects as an example. We assume that at first infants are not introspectively aware of different emotion states. Rather, their representations of these emotions are primarily based on stimuli received from the external world. Babies learn to differentiate the internal patterns of physiological and visceral stimulation that accompany different emotions by observing their caregivers' facial or vocal mirroring responses to these (e.g., Legerstee & Varghese, 2001; e.g., Mitchell, 1993). First, the baby comes to associate his control 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. Second, 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. If the parent's affect expressions are not contingent on the infant's affect this will undermine the appropriate labelling of internal states which may, in turn, remain confusing, experienced as unsymbolized, and hard to regulate.

If the capacity to understand and regulate emotion is to develop, two conditions need to be met: (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, 1999). Consequently, two difficulties may arise: (a) in the case of incongruent mirroring the infant's representation of internal state will not correspond to a constitutional self state (nothing real) and a predisposition to a narcissistic structure might be established, perhaps analogous to Winnicott's notion of “false-self” (Winnicott, 1965); and (b) in cases of un-marked mirroring the caregiver's expression may be seen as externalization of the infant's experience and a predisposition to experiencing emotion through other people (as in a borderline personality structure) might be established (Fonagy, Gergely, Jurist, & Target, 2002). 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, making the child's experience seem contagious and escalating rather than regulating his state.



The secure caregiver soothes by combining mirroring with a display that is incompatible with the child's feelings (thus implying contact with distance and coping). This formulation of sensitivity has much in common with Bion's (1962) notion of the role of the mother's capacity to mentally "contain" the affect state that feels intolerable to the baby, and respond in a manner that acknowledges the child's mental state, yet serves to modulate unmanageable feelings (see below). Well-regulated affect in the infant parent couple is thought to be internalized by the child to form the basis of a secure attachment bond and internal working model (Sroufe, 1996). Ratings of the quality of reflective function of each parent during pregnancy were found independently to predict the child's later security of attachment in the London Parent-Child Project (Fonagy, Steele, Moran, Steele, & Higgitt, 1992). However, this finding is somewhat limited since only the AAI RF measure was examined in relation to infant attachment (Fonagy, Steele, Moran, Steele, & Higgitt, 1991; Fonagy, Steele, Steele, Higgitt, & Target, 1994). Thus, the parents' capacity to mentalize was measured in relation to their own childhood and their capacity to do likewise with their child had been assumed rather than observed.

#### EVIDENCE LINKING PARENTAL MENTALIZATION TO THE DEVELOPMENT OF WELL REGULATED AFFECT (SECURE ATTACHMENT)

Three programs of work, by Elizabeth Meins (Meins, Fernyhough, 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), took this forward, all looking at aspects of interactional narratives between parents and children. In the Meins study mentalization was assessed on the basis of the mothers' verbalization to a 6-month-old infant. One measure of maternal mind-mindedness (MMM) repeatedly used in Meins's studies was developed based on the question "Can you describe [child] for me?" with the codified categories of mental, behavioral, physical, and general (Meins & Fernyhough, 1999; Meins, Fernyhough, Russel, & Clark-Carter, 1998). This is an "off-line" measure of mentalizing, but Meins and colleagues have also developed a more "on-line" measure of MMM based on twenty minutes of free play between mothers and their 6-month-old babies, which are coded for appropriate mind-related comments, amongst other parameters. Mind-related comments were shown to be predictive of attachment security at 6 months (Meins, Fernyhough, Fradley, & Tuckey, 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 studies, the mothers provided commentaries on their own previously recorded playful interaction with their child. Both studies found that high levels of mentali-

zation of the child in the mothers' narratives were associated with secure infant-mother attachment. While both studies demonstrated that mentalization of the child in the context of the mother-child relationship, rather than global measures of sensitivity, was likely to predict the security of the attachment relationship, the studies assessed the quality of mentalization rather differently. The measure used in the Meins study aimed to assess the quality of the parents' thinking about the child in "real time" in the course of an interaction. The measure used in the Oppenheim studies was focused on a more reflective, "off-line" mentalizing capacity. Both measures were, however, "episodic" giving an indication of the parent's quality of mentalization of a particular moment of interaction. Neither was designed to measure the extent that mothers generally mentalize their relationship with their child (or rather their idea of their relationship with their idea of their child).

The Slade et al. (Slade et al., 2005) study extends previous observations by using an AAI-like autobiographical memory focused measure, the Parent Development Interview (PDI), rather than an episode of observed interaction as an index of mentalizing capacity. Strong relationships were found between attachment in the infant and the quality of mentalizing in the parent about the child. A measure such as the PDI estimates mentalization as an aggregate across many episodes of interaction and what might be assumed to be a prototype is drawn 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 gives access to these latter types of general autobiographical memories that are assumed to have a preferred level of entry to the autobiographical memory system (Conway & Holmes, 2004).

As commentary on the events remembered is part of the content scrutinized for level of mentalization, the PDI measure probably also incorporates an indication of the mother's "off-line" reflective mentalizing capacity picking up the mother's predominant stance towards the child as more or less an intentional being, perhaps reflecting many hundreds of interactions and thus providing greater accuracy of prediction. In this way it is able to index more than simple "mind-mindedness" (Meins et al., 2001), measured as the complexity of mental state terms and concepts used. High scorers on the PDI-RF scale are aware of the characteristics of mental functioning in their infants and grasp the complex interplay between their own mental states and the child's inner experience.

The Slade et al. (Slade et al., 2005) study includes ten infants with disorganized attachment classification, whose mothers' RF scores are a standard deviation below those who are secure. What do low RF parents do that might disorganize 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 is particularly sensitive to atypical behaviors associated with the disorganization of infant-mother attachment. Parents of disorganized infants were almost a standard deviation higher on this measure than parents of secure ones. There is a substantial correlation between AMBIANCE codings and RF—again, an effect size greater than 1. The size of the effect is somewhat surprising given the disparity of the domains of measurement; the AMBIANCE is a behavioral measure based on a single interaction and RF is coded from a narrative. There appears to be a strong relationship between the observed frequency of behaviors, such as demanding a show of affection from the infant, fearful behavior, or intrusive or negative behaviors, such as mocking or criticizing and narratives that, for example, show little appreciation that the infant's mind cannot be directly read, or depict her as having no feelings, thoughts, or wishes. This 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 organization of narratives about her. It is conceivable that the correlation is accounted for by the common neural basis that might underpin both tasks. The mentalizing system might provide input for the organization of both social interaction and person-centered autobiographical narrative.

This suggests that possibly Slade and her colleagues have closed the transmission gap identified a decade ago by Marinus van IJzendoorn (1995). A somewhat simplistic restatement of our current knowledge might go like this. Secure attachment history of the mother permits and enhances her capacity to explore her own mind and liberates 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 behaviors 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. In this context, then, disorganization of attachment is implicitly seen by Arietta Slade and her group as the consequence of an undermining of a mental self, or the disorganization of the self.

Affect regulation, the capacity to modulate emotional states, is closely related to mentalization, which plays a fundamental role in the unfolding of a sense of self and agency. In this account, affect regulation is a prelude to mentalization; yet, once mentalization occurs, the nature of affect regulation is transformed: not only does it allow adjustment of affect states, but more fundamentally it is used to regulate the self. This is 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 attachment figure. Social understanding is an emergent property of the

child's experience of referential interactions with the caregiver about an object, which will inevitably generate the discovery that others have differing beliefs about the world from one's own.

#### ESTABLISHING ATTENTIONAL CONTROL

Early attachment, which allows the child to internalize the mother's ability to divert the child's attention from something immediate to something else (Fonagy, 2001), serves to equip children with the capacity for attentional control. Longitudinal studies of self-regulation demonstrate that the capacity for effortful control is strongly related to a child's observed willingness to comply with maternal wishes (Kochanska, Coy, & Murray, 2001). Withholding an impulsive response is a prerequisite for mentalizing, as this requires the foregrounding of a distal second-order non-visible stimulus (mental state) in preference to what immediately impinges on the child (physical reality). The successful performance of theory of mind tasks, for example, must involve the inhibition of the child's prepotent responses to directly perceived aspects of current reality in favor of generating a response on the basis of less salient representations of reality attributed to other minds.

Attentional control is also linked to attachment. The major function of attachment is the control of distress and attentional processes must play a key role if the attachment system is to achieve this objective (Harman, Rothbart, & Posner, 1997). Michael Posner, amongst others, suggests that the interaction between infant and caregiver is likely to train the infant to control his distress through orienting the infant away from the source of distress by soothing and involving him in distracting activities. Self-regulation is taught (or more accurately, modelled) by the caregiver's regulatory activity. It has been suggested that joint-attention with caregiver serves a self-organizing function in early development (Mundy & Neal, 2001). Indeed, we have long known that intelligence remains related to early attachment security (e.g., Cicchetti, Rogosch, & Toth, 2000; e.g., Jacobsen & Hofmann, 1997). More recently, Jay Belsky and Pasco Fearon have drawn our attention to early attachment relationships as a possible organizer of attentional systems (Belsky & Fearon, 2002; Fearon & Belsky, 2004). In a study of almost 1,000 children, a positive relationship was found between attachment and attentional performance using a Continuous Performance Test (CPT) to measure attentional capacity at 54 months. Findings indicated that children with secure attachment appeared to be protected from the effects of cumulative social contextual risk (and male gender) on CPT attentional performance relative to their insecure counterparts. A further study of infants who were disorganized in their attachment found that these infants also had difficulties with social attention coordination in interactions with their caregiver (Schölmerich, Lamb, Leyendecker, & Fracasso, 1997). Cocaine-exposed children with disorganized attachment at 12 months showed the greatest dysfunctions of social attention coordina-

tion not only with the caregiver, but also with an experimenter (e.g., they initiated joint attention less often) (Claussen, Mundy, Mallik, & Willoughby, 2002). Evidence from late-adopted Romanian orphans with profound disorganizations of attachment suggests that quite severe attention problems are more common in this group than would be expected both in relation to other forms of disturbance and epidemiological considerations (e.g., Chugani *et al.*, 2001).

From the point of view of our model of the development of mentalization, we argue that an enfeebled attentional control system is a likely consequence of attachment disorganization, perhaps linked with enfeebled affect representation, and serves to undermine the development of mentalization as well as its appropriate functioning in later development. The prepotent response is to attribute one's own mental state to the other. Attentional control is essential if the child is to arrive at a differentiation of their own and others' thoughts, feelings, beliefs, and desires. The disruption of attentional control is likely to account for many instances where we encounter temporary and selective disruptions of mentalization. It is probable that trauma further undermines attention regulation and is associated with chronic failures of inhibitory control (Allen, 2001).

#### THE STAGES OF ACQUIRING MENTALIZATION (A THEORY OF MIND)

The emergence of mentalizing function follows a well-researched developmental line that identifies "fixation points":

(a) During the second half of the first year of life, the child begins to construct causal relations that connect actions to their agents on the one hand and to the world on the other. Infants around 9 months begin to look at actions in terms of the actor's underlying intentions (Baldwin, Baird, Saylor, & Clark, 2001). This is the beginning of their understanding of themselves as teleological agents who can choose the most efficient way to bring about a goal from a range of alternatives (Csibra & Gergely, 1998). At this stage agency is understood in terms of purely physical actions and constraints. Infants expect actors to behave rationally, given physically apparent goal states and the physical constraints of the situation that are already understood by the infant (Gergely & Csibra, 2003). There is no implication here that the infant has an idea about the mental state of the object. He/she is simply judging rational behavior in terms of the physical constraints that prevail and that which is obvious in terms of the physical end state which the object has reached. We have suggested a connection between the focus on understanding actions in terms of their physical as opposed to mental outcomes (a teleological stance) and the mode of experience of agency that we often see in the self-destructive acts of individuals with borderline personality disorder (BPD) (Fonagy, Target, & Gergely, 2000). Thus, slight changes in the physical world can trigger elaborate conclusions concerning states of mind. Patients frequently cannot accept

anything other than a modification in the realm of the physical as a true index of the intentions of the other.

(b) During the second year, children develop a mentalistic understanding of agency. They understand that they and others are intentional agents whose actions are caused by prior states of mind such as desires (Wellman & Phillips, 2000) and that their actions can bring about changes in minds, as well as bodies (e.g., by pointing Corkum & Moore, 1995). Shared imaginative play is enjoyable and exciting for toddlers and may be the basis for the development of collaborative, co-operative skills (Brown, Donelan-McCall, & Dunn, 1996). Fifteen-month-old children can distinguish between an action's intended goal and its accidental consequences (Meltzoff, 1995). At this stage the capacity for emotion regulation comes to reflect the prior and current relationship with the primary caregiver (Calkins & Johnson, 1998). Most importantly, children begin to acquire an internal state language and the ability to reason non-egocentrically about feelings and desires in others (Repacholi & Gopnik, 1997). Paradoxically, this becomes evident not only through the increase in joint goal directed activity, but also through teasing and provocation of younger siblings (Dunn, 1988). However, functional awareness of minds does not yet enable the child to represent mental states independent of physical reality and, therefore, the distinction between internal and external, appearance and reality is not yet fully achieved (Flavell & Miller, 1998), making internal reality sometimes far more compelling and at other times inconsequential relative to an awareness of the physical world. We have referred to these states as psychic equivalence and pretend modes respectively (see below).

(c) Around three to four years of age, understanding of agency in terms of mental causation begins to include the representation of epistemic mind states (beliefs). The young child thus understands himself as a representational agent, he knows that people do not always feel what they appear to feel, they show emotional reactions to an event that are influenced by their current mood or even by earlier emotional experiences which were linked to similar events (Flavell & Miller, 1998). The preschool child's mental states are representational in nature (Wellman, 1990). This transforms their social interactions so their understanding of emotions comes to be associated with empathic behavior (Zahn-Waxler, Radke-Yarrow, Wagner, & Chapman, 1992) and more positive peer relations (Dunn & Cutting, 1999). Most children come to understand that human behavior can be influenced by transient mental states (such as thoughts and feelings) as well as by stable characteristics (such as personality or capability) and this creates the basis for a structure to underpin an emerging self-concept (Flavell, 1999). They also come to attribute mistaken beliefs to themselves and to others, which enriches their repertoire of social interaction with tricks, jokes, and deception (Sodian & Frith, 1992; Sodian, Taylor, Harris, & Perner, 1992). A meta-analytic review in excess of 500 tests showed that, by and large, children younger than three fail the false-belief task and as the child's age increases they are increasingly likely to pass (Well-

man, Cross, & Watson, 2001), suggesting that mentalizing abilities take a quantum leap forward around age four. The early acquisition of false belief is associated with more elaborate capacity to pretend play (Taylor & Carlson, 1997), greater connectedness in conversation (Slomkowski & Dunn, 1996) and teacher rating of social competence (Lalonde & Chandler, 1995). Notably, also at this time the child shifts from a preference for playing with adults to playing with peers (Dunn, 1994). We understand this shift as bringing to a close the time when mentalization was acquired through the agency of an adult mind and opening a lifelong phase of seeking to enhance the capacity to understand self and others in mental state terms through linking with individuals who share one's interest and humor.

(d) In the sixth year, we see related advances such as the child's ability to relate memories of his intentional activities and experiences into a coherent causal-temporal organization, leading to the establishment of the temporally extended self (Povinelli & Eddy, 1995). Full experience of agency in social interaction can emerge only when actions of the self and other can be understood as initiated and guided by assumptions concerning the emotions, desires, and beliefs of both. Further theory of mind skills that become part of the child's repertoire at this stage include second order theory of mind (the capacity to understand mistaken beliefs about beliefs), mixed emotions (e.g., understanding being in a conflict), the way expectations or biases might influence the interpretation of ambiguous events, and the capacity for subtle forms of social deceptions (e.g., white lies). As these skills are acquired, the need for physical violence begins to decline (Tremblay, 2000; Tremblay, Japel, & Perusse, 1999) and relational aggression increases (Cote, Tremblay, Nagin, Zoccolillo, & Vitaro, 2002; Nagin & Tremblay, 2001).

#### RELATIONSHIP INFLUENCES ON THE ACQUISITION OF MENTALIZATION

Our claim that attachment relationships are vital to the normal acquisition of mentalization challenges nativist assumptions. The nativistic position assumes that children's social environments can trigger, but cannot determine the development of theory of mind (Baron-Cohen, 1995; Leslie, 1994). There is some evidence that the timetable of theory of mind development is fixed and universal (Avis & Harris, 1991). However, the bulk of the evidence is inconsistent with the assumption of a universal timetable. More recent studies find ample evidence for substantial cultural differences, not just in the rate of emergence of theory of mind skills, but also the order of their emergence (Wellman, Cross, & Watson, 2001). Many findings suggest that the nature of family interactions, the quality of parental control (e.g., Vinden, 2001), parental discourse about emotions (e.g., Meins et al., 2002), the depth of parental discussion involving affect (Dunn, Brown, & Beardsall, 1991), and parents' beliefs about parenting (e.g., Ruffman, Perner, & Parkin, 1999) are all strongly associated with

the child's acquisition of mentalization. The role of family members in this developmental achievement is further highlighted by the finding that the presence of older siblings in the family appears to improve the child's performance on a range of false-belief tasks (e.g., Ruffman, Perner, Naito, Parkin, & Clements, 1998). In sum, the ability to give meaning to psychological experiences evolves as a result of our discovery of the mind behind others' actions, which develops optimally in a relatively safe and secure social context.

Much that is known about correlates and predictors of early ToM development is consistent with the assumption that the attachment relationship plays an important role in the acquisition of mentalization. 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). 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 (e.g., Charman, Ruffman, & Clements, 2002).

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 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. 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, 1999). (2) While, as we have seen, broadly, positive emotion promotes the emergence of mentalization (Dunn, 1999), negative emotion can 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, humor 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).



#### **4. UNDERSTANDING THE RELATIONSHIP INFLUENCES ON MENTALIZATION**

##### **INTERSUBJECTIVITY BEYOND INFANCY**

The basic assumption of modern developmental theory is of a primary intersubjectivity—that knowledge about the world is shared knowledge. To paraphrase this, the evolutionary underpinnings of human culture require that the infant turns to others for essential information about the world (Gergely & Csibra, 2005). The idea of a shared consciousness in infancy is not new. A number of developmentalists have emphasized the key functions of such sharing (e.g., Hobson, 2002; Rochat & Striano, 1999). The sharing of minds established at this early stage is considered by many philosophers of mind (e.g., Cavell, 1994) and relational psychoanalysts (e.g., Mitchell, 2000) to be a stable characteristic of mental function. We have argued that the evidence for relational influences on mentalization is best explained by the assumption that the acquisition of theory of mind is part of an intersubjective process between the infant and caregiver (see Gopnik, 1993, for an elegant elaboration of such a model). In our view, the caregiver helps the child create mentalizing models, through complex linguistic and quasi-linguistic processes that involve non-verbal, as well as verbal aspects of social interaction within an attachment context (Brown, Hobson, Lee, & Stevenson, 1997).

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. For example, a recent study (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, implying that this sharing of attention and interest was indeed their goal. Infants of 12 months happily point just to inform an adult of the location of a misplaced object they have no direct interest in. Such declarative and informing motives are apparently "purely social" in their aims.

The small child assumes that what he knows is known by others and vice versa, that is, that the world is shared between all of us. Only slowly does the uniqueness of our own perspective differentiate so that a sense of mental self can develop. Infants possess by three months or so at the latest, a distinct sense of their integrity as physical beings. But in relation to what we know and understand about reality, we start with the assumption that knowledge is common and there is nothing unique about our own thoughts or feelings. 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). 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, frustration of their wishes seems malign or wilfully obtuse, rather than the result of a different point of view, alternative priorities, etc.

Mentalization evolves out of the child's biological predisposition to assume that his knowledge is shared by all. The child naturally turns to the caregiver to learn from her about the nature of the world, internal and external. Unconsciously and pervasively, the caregiver ascribes a mental state to the child with her behavior, treating the child as a mental agent. Ultimately, the child concludes that the caregiver's reaction to him makes sense given internal states of belief or desire within himself. This conclusion enables him to elaborate mental models of causation, and facilitates the development of a core sense of selfhood organized along these lines. We assume that this is mostly a mundane process, and that it is preconscious to both infant and parent—inaccessible to reflection or modification.

#### SUBJECTIVITY BEFORE MENTALIZATION

How does the child experience subjectivity before he recognizes that internal states are representations of reality? In describing the normal development of mentalizing in the child of two to five years (Fonagy & Target, 1996; Target & Fonagy, 1996), we suggest that there is a transition from a split mode of experience to mentalization. We hypothesize that the very young child equates the internal world with the external. What exists in the mind must exist out there and what exists out there must also exist in the mind. At this stage there is no room yet for alternative perspectives. "How I see it, is how it is." The toddler's or young pre-school child's insistence that "there is a Tiger under the bed" is not allayed by parental reassurance. This "psychic equivalence," as a mode of experiencing the internal world, can cause intense distress, since the experience of a fantasy as potentially real can be terrifying. The acquisition of a sense of pretend in relation to mental states is therefore essential. While playing, the child knows that internal experience may not reflect external reality (e.g., Bartsch & Wellman, 1989; Dias & Harris, 1990), but then the internal state is thought to have no implications for the outside world ("*pretend mode*").

Normally at around four years old, the child integrates these modes to arrive at *mentalization*, or *reflective mode*, in which mental states can be experienced as representations. Inner and outer reality can then be seen as linked, yet differing in important ways, and no longer have to be either equated or dissociated from each other (Gopnik, 1993). The child discovers that "seeing-leads-to-knowing;" if you have seen something in a box, you know something about what's in the box (Pratt & Bryant, 1990). They can begin to work out from gaze direction what a person is thinking about, thus making use of the eyes of another person to make a mentalistic inter-

pretation (Baron-Cohen & Cross, 1992). There are, however, circumstances under which pre-mentalistic forms of subjectivity re-emerge to dominate social cognition years after the acquisition of full mentalization. We shall consider these in section 5.

Mentalization normally comes about through the child's experience of his mental states being reflected on, prototypically through secure play with a parent or older child, which facilitates integration of the pretend and psychic equivalence modes. This interpersonal process is perhaps an elaboration of the complex mirroring the parent offered earlier. In playfulness, the caregiver gives the child's ideas and feelings (when he is "only pretending") a link with reality, by indicating an alternative perspective outside the child's mind. The parent or older child also shows that reality may be distorted by acting upon it in playful ways, and through this playfulness a pretend but real mental experience may be introduced.

If the child's capacity to perceive mental states in himself and others depends on his observation of the mental world of his caregiver, clearly children require a number of adults with an interest in their mental state, who can be trusted (i.e., with whom an attachment bond exists), to support the development of their subjectivity from a pre-mentalizing to a fully mentalizing mode. In this regard, in past initiatives, perhaps we have placed too much emphasis on parents (particularly mothers). It follows from the evolutionary model presented in section 2 and here that the child's brain is experience expectant from a range of benign adults willing to take the pedagogic stance towards their subjectivity. Thus, teachers, neighbors, older siblings, as well as parental figures, could play important roles in optimizing the child's capacity for mentalization. Children can perceive and conceive of their mental states to the extent that the behavior of those around them has implied that they have them. This can happen through an almost unlimited set of methods ranging from shared pretend playing with the child (empirically shown to be associated with early mentalization), and many ordinary interactions (such as conversations and peer interaction) will also involve shared thinking about an idea.

#### DISORGANIZED ATTACHMENT AND THE UNMENTALIZED (ALIEN) SELF

In children whose attachment is disorganized, mentalization may be evident, but it does not play the positive role in self-organization that it does in securely or even in insecurely attached children. The child with disorganized attachment is forced to look not for the representation of his own mental states in the mind of the other, but the mental states of that other which threaten to undermine his agentic sense of self. These mental states can create an alien presence within his self-representation, so unbearable that his attachment behavior becomes focused on re-externalizing these parts of the self onto attachment figures, rather than on the in-

ternalization of a capacity for containment of affects and other intentional states.

Disorganized infants, even if interpersonally perceptive, fail to integrate this emotional awareness with their self-organization. There may be a number of linked reasons for this: a) the child needs to use disproportionate resources to understand the parent's behavior, at the expense of reflecting on self-states; b) the caregiver of the disorganized infant is likely to be less contingent in responding to the infant's self-state, and further to show systematic biases in her perception and reflection of his state; c) the mental state of the caregiver of the disorganized infant may evoke intense anxiety through either frightening or fearful behavior towards the child, including inexplicable fear of the child himself. These factors combine, perhaps, to make children whose attachment system is disorganized become keen readers of the caregiver's mind under certain circumstances, but (we suggest) poor readers of their own mental states.

## **5. THE DECOUPLING OF MENTALIZATION IN THE PRESENCE OF ATTACHMENT TRAUMA**

### **TRAUMA RELATED LOSS OF THE CAPACITY TO CONCEIVE OF MENTAL STATES**

Adults with a history of childhood attachment trauma often seem unable to understand how others think or feel. We have hypothesized that childhood maltreatment undermines mentalization. When combined with the enfeebled affect representation, poor affect control systems and disorganized self structure that can result from a deeply insecure early environment, trauma has profound effects: (a) It inhibits playfulness which is essential for the adequate unfolding of the interpersonal interpretive function (Dunn, Davies, O'Connor, & Sturgess, 2000); (b) it interferes directly with affect regulation and attentional control systems (Arntz, Appels, & Sieswerda, 2000); (c) most importantly, in vulnerable individuals, it can lead to an unconsciously motivated failure of mentalization. This failure is a defensive adaptive manoeuvre: the child seeks to protect himself from the frankly malevolent and dangerous states of mind of the abuser by decoupling his capacity to conceive of mental states, at least in attachment contexts (Fonagy, 1991). (d) We believe that adult social functioning is impaired by childhood and adolescent adversity to the extent that adversity causes a breakdown of attachment related mentalization (Fonagy et al., 2003a). There is considerable evidence that maltreated children have specific mentalization deficits and that individuals with BPD are poor at mentalization following severe experiences of maltreatment (Fonagy et al., 1996). Children cannot learn words for feelings (Beeghly & Cicchetti, 1994), and adults have more difficulty recognizing facial expressions, the more severe their childhood maltreatment (Fonagy et al.,

2003a). What is the clinical picture like when trauma brings about a partial and temporary collapse of mentalization? We observed an apparent lack of imagination about the mental world of others, a naiveté or cluelessness about what others think or feel that can verge on confusion, and a corresponding absence of insight into the way that the traumatized person's own mind works.

Many maltreated children grow up into adequately functioning adults. While maltreatment places children at increased risk for developing psychopathology, only a small proportion will prospectively need mental health services (Widom, 1999). It is possible that early maltreatment reduces the individual's opportunity fully to develop mentalizing skills, leaving them with inadequate capacities to identify and avoid risks for further interpersonal trauma. In dysfunctional attachment contexts, particularly when children are victims of abuse, they may learn to interpret parental initiation of communicative attention-directing behaviors as a cue that potentially harmful interactions are likely to follow. In consequence, they may defensively inhibit the mentalistic interpretation of such cues; this may finally lead to the defensive disruption of their own metacognitive monitoring procedures in all subsequent intimate relationships (Fonagy, Target, Gergely, Allen, & Bateman, 2003b).

#### THE EQUATION OF INNER AND OUTER

The collapse of mentalization in the face of trauma entails a loss of awareness of the relationship between internal and external reality (Fonagy & Target, 2000). Modes of representing the internal world re-emerge that developmentally precede an awareness that thoughts, feelings, and wishes are part of the mind. The 2–3 year old as we saw, not yet experiencing his mind as truly representational, assumes in the mode of psychic equivalence that what he thinks also exists in the physical world. Post-traumatic subjective experience (the flashback) is similarly compelling, resistant to argument and feels dangerous until it becomes mentalized. Often survivors of trauma simply refuse to think about their experience because thinking about it means reliving it. Aspects of the notion of psychic equivalence evidently overlap with descriptions of paranoid-schizoid forms of thinking, particularly as formulated by Wilfred Bion in the "*Elements of Psychoanalysis*" (Bion, 1963), and symbolic equation as formulated by Hanna Segal (1957).

#### SEPARATION FROM REALITY

As we saw, the pretend mode is a developmental complement to psychic equivalence. Not yet able to conceive of internal experience as mental, the child's fantasies are dramatically divided off from the external world. Small children cannot simultaneously pretend (even though they know it is not real) and engage with normal reality; asking them if their pretend gun is a

gun or a stick spoils the game. Following trauma and the constriction of mentalization we see the intrusion of the pretend mode, particularly in dissociative experiences. In dissociated thinking, nothing can be linked to anything—the principle of the “pretend mode,” in which fantasy is cut off from the real world, is extended so that nothing has implications (Fonagy & Target, 2000). Patients report “blacking out,” “clamming up” or remembering their traumatic experiences only in dreams. The most characteristic feature of traumatization is the oscillation between psychic equivalence and pretend modes of experiencing the internal world.

#### “I BELIEVE IT WHEN I SEE IT”

A third pre-mentalistic aspect of psychic reality is the re-emergence of a teleological mode of thought. This mode of understanding the world antedates even language. Infants as young as 9 months are able to attribute goals to people and to objects that seem to behave purposefully, but these goals are not yet truly mental, they are tied to what is observable. The return of this teleological mode of thought is perhaps the most painful aspect of a subjectivity stripped of mentalization.

Following trauma, verbal reassurance means little. Interacting with others at a mental level has been replaced by attempts at altering thoughts and feelings through action. Trauma, certainly physical and sexual abuse, is by definition teleological. It is hardly surprising that the victim feels that the mind of another can only be altered in this same mode, through a physical act, threat, or seduction. Following trauma we all need physical assurances of security.

#### THE IMPACT OF ATTACHMENT TRAUMA ON MENTALIZATION: THE HYPERACTIVATION OF ATTACHMENT

Attachment is normally the ideal “training ground” for the development of mentalization because it is safe and non-competitive. This biological configuration, which is so adaptive in the context of normal development, becomes immensely destructive in the presence of attachment trauma. Attachment trauma hyperactivates the attachment system because the person to whom the child looks for reassurance and protection is the one causing fear. The devastating psychic impact of attachment trauma is the combined result of the inhibition of mentalization by attachment and the hyperactivation of the attachment system by trauma. This context demands extraordinary mentalizing capacities from the child, yet the hyperactivation of the attachment system will have inhibited whatever limited capacity he has.

The coincidence of trauma and attachment creates a biological vicious cycle. Trauma normally leads a child to try to get close to the attachment figure. Where the child depends on an attachment figure who maltreats them, there is a risk of an escalating sequence of further maltreatment,

increased distress, and an ever-greater inner need for the attachment figure. The inhibition of mentalization in a traumatizing, hyperactivated attachment relationship is always likely to lead to a prementalistic psychic reality, largely split into psychic equivalence and pretend modes. Because the memory of the trauma feels currently real, there is a constant danger of re-traumatization from inside. The traumatized child often begins to fear his own mind. The inhibition of mentalization is also clearly an intrapsychic adaptation to traumatic attachment. The frankly malevolent mental state of the abuser terrifies the helpless child. The parents' abuse undermines the child's capacity to mentalize, because it is no longer safe for the child, for example, to think about wishing, if this implies recognizing his parent's wish to harm him. Because he cannot use the model of the other to understand himself, diffusion of identity, and dissociation often follows.

#### THE IMPACT OF ATTACHMENT TRAUMA ON MENTALIZATION: THE BIOLOGY OF BEING FRAZZLED

The impact of trauma on mentalization is intermittent. As above, sometimes mentalization disappears because an attachment relationship intensifies. At other times, being stressed (for example, touching on a sensitive issue) can trigger what feel like wild, unjustified reactions. Eight years ago, in a hallmark paper entitled "The biology of being frazzled," Amy Arnsten (1998) explained why (see also Arnsten, Mathew, Ubriani, Taylor, & Li, 1999; Mayes, 2000). At the risk of simplifying highly complex pioneering neuroscientific work, Arnsten's Dual Arousal Systems Model delineates two complementary, independent arousal systems: the prefrontal and posterior cortical and subcortical systems. The system that activates frontal and pre-frontal regions inhibits the second arousal system that normally "kicks in" only at quite high levels of arousal, when pre-frontal activity goes "offline" and posterior cortical and subcortical functions (e.g., more automatic or motor functions) take over.

The switch-point between the two arousal systems may be shifted by childhood trauma. Undoubtedly, as mentalization is located in the pre-frontal cortex, this accounts for some of the inhibition of mentalization in individuals with attachment trauma, in response to increases in arousal that would not be high enough to inhibit mentalization in most of us. Anticipating some of the clinical implications of our thinking, in the light of this phenomenon, it is important to monitor the traumatized patient's readiness to hear comments about thoughts and feelings. As arousal increases, in part in response to interpretative work, traumatized patients cannot process talk about their minds. Interpretations of the transference at these times, however accurate they might be, are likely to be way beyond the capacity of the patient to hear. The clinical priority has to be work to reduce arousal so that the patient can again think of other perspectives (mentalize).

### THE IMPACT OF ATTACHMENT TRAUMA ON MENTALIZATION: PROJECTIVE IDENTIFICATION AS A MATTER OF LIFE AND DEATH

Bion's (1963) first element in his elements of psychoanalysis is "the essential feature of Melanie Klein's conception of projective identification . . . the dynamic relationship between container and contained" (p. 3). Edith Jacobson (1954) and Donald Winnicott (Winnicott, 1956) independently noted that the internalization of the representation of another before the boundaries of the self are fully formed undermines the creation of a coherent sense of self. The infant is forced to internalize the other not as an internal object but as a core part of his self. If the caregiver fails to contain the infant's anxieties, metabolize them, and mirror the self state, the infant, rather than gradually constructing a representation of his internal states, is forced to accommodate the object, an alien being, within his self representation. Such incoherencies in self-structure are not only features of profoundly neglected children. Since even the most sensitive caregiver is insensitive to the child's state of mind over 50% of the time, we all have alien parts to our self-structure. The illusion of self-coherence is normally maintained by the continuous narrative commentary on behavior that mentalization provides, preconsciously. This weaves our experiences together so that they make sense. In the absence of a robust mentalizing capacity, in the wake of trauma, alien fragments in the self-structure are likely to be clearly revealed in all of us.

Of course, these introjections in traumatized individuals are colored by the traumatic context in which they occur. What is internalized as part of the self is a caregiver with terrifying intentions. This can generate momentary experiences of unbearable psychic pain when in the mode of psychic equivalence the self feels attacked literally from within and almost overwhelmed by an experience of "badness" that reassurance cannot mitigate and from which, in a teleological mode of functioning, self-destruction might appear the only escape. In our view, this state is commonly the trigger for acts of self-harm and suicide.

The only way the person can deal with such introjects is by constantly externalizing these alien parts of the self-structure into an other. Through projective identification the persecutory parts are experienced as outside. It is then essential that the alien experiences are owned by another mind, so that another mind is in control of the parts of the self set upon its own destruction. Paradoxically, then, the need for projective identification is a matter of life and death for those with a traumatizing part of the self-structure, but the constellation creates a dependence on the object that has many features of addiction. Neuroscience is helpful here, in explaining that the triggering of the attachment system (by the need to find a container for traumatized, alien parts of the self) will once again inhibit mentalization. This reduces the chance of either alternative solutions being accepted or a non-teleological (non-physical) solution being found.



Maltreatment, or more broadly trauma, is seen as interacting with the domain- and situation-specific restrictions upon mentalization at two levels. First, maltreatment makes the young child reluctant to take the perspective of others, because of the actual threat within the intentional stance of the abuser, as well as the constraints upon self-development imposed by the parent's failure to understand and acknowledge the child's budding intentionality. Second, the child is deprived of the later resilience provided by the capacity to understand interpersonal situations (Fonagy et al., 1994). Thus, individuals traumatized by their family environment are vulnerable in terms of the long-term impact of the trauma, their reduced capacity to cope with it, and their difficulty in finding better relationships later. The outcome may be severe developmental psychopathology, ultimately entrenched personality disorder.

## **6. CONCLUSION**

We have considered the development of mentalization from both a phylogenetic and ontogenetic perspective. We have argued that mentalization has a selective advantage in enhancing collaboration in the context of attachment and competition with conspecifics in all other contexts. Against this evolutionary background we argued that the preferred context for the ontogenetic development of mentalization is one where the child can have trust in the person who has the child's mind in mind. The child's sense of an agentive self, underpinned by the capacity for mentalization, takes shape in this interpersonal context. Some of the brain mechanisms underpinning mentalization are inhibited when the mesocorticolimbic dopaminergic system that mediates attachment and social affiliation is activated. We have argued that this is likely to be part of the evolutionary design that privileges close relationships for the safe exploration of intersubjective space.

Mentalization is acquired alongside a range of associated cognitive capacities necessary for conceptualizing mental states. Affect representation and regulation and attentional control are important aspects of this development. The quality of children's relationships with those from whom they acquire an understanding of minds is likely to be crucial to all these. The creation of an integrated sense of agentive self depends upon a contingently, but not too accurately, mirroring relational context. Incongruent and poorly marked mirroring is assumed by us to create the kind of incoherence and disorganization within the self-structure that could account for the controlling disturbed behavior of kindergarten-aged children with a history of disorganized attachment. Before mentalization is fully acquired, subjectivity is dominated by the equation of the status of internal and external and the complement of this state, an experience of dissociation between internal and external. Disturbed attachment organization is likely to be associated with persistence of these non-mentalized ways of representing subjectivity.

We have suggested that the psychological consequences of trauma, in an attachment context and perhaps beyond, entail a decoupling of mentalization and a re-emergence of non-mentalizing modes of representing internal reality. This is pernicious because the immediacy of a memory experienced in the non-mentalizing mode of psychic equivalence has the capacity to re-traumatize again and again. This further inhibits mentalization and makes the experience ever more real. Trauma in the attachment context is most pernicious because the biological basis of attachment assumes trust. Part of this is the safety of not having to mentalize, of knowing that others are thinking for us, that we need not monitor our own or others' thinking. Trauma inevitably activates the attachment system. This activation (probably for evolutionary reasons) temporarily inhibits areas of the brain concerned with both remembering and mentalization. This is why mentalization comes to be so readily abandoned in the face of trauma, particularly attachment trauma. Unmentalized trauma endures and compromises mental function. Of course, it also interferes with new relationships. The self being destroyed from within, by identification with the aggressor, is an imperative for projective identification, drawing the other closer and selecting relationships that will retraumatize. To escape from the grip of trauma, the individual needs help to recover mentalization.

## REFERENCES

- 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.
- Allen, J. G. (2001). *Traumatic relationships and serious mental disorders*. Chichester: Wiley.
- 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.
- Arntz, A., Appels, C., & Sieswerda, S. (2000). Hypervigilance in borderline disorder: A test with the emotional Stroop paradigm. *J Personal Disord*, *14*, 366–373.
- Avis, J., & Harris, P. (1991). Belief-desire reasoning among Baka children: Evidence for a universal conception of mind. *Child Development*, *62*, 460–467.
- Bahrnick, L. R., & Watson, J. S. (1985). Detection of intermodal proprioceptive-visual contingency as a potential basis of self-perception in infancy. *Developmental Psychology*, *21*, 963–973.
- Baldwin, D. A., Baird, J. A., Saylor, M. M., & Clark, M. A. (2001). Infants parse dynamic action. *Child Development*, *72*, 708–717.
- Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Cambridge, MA: Bradford, MIT Press.
- Baron-Cohen, S., & Cross, P. (1992). Reading the eyes: Evidence for the role of perception in the development of a theory of mind. *Mind and Language*, *6*, 173–186.
- Bartels, A., & Zeki, S. (2000). The neural basis of romantic love. *Neuroreport*, *11*, 3829–3834.
- Bartels, A., & Zeki, S. (2004). The neural correlates of maternal and romantic love. *Neuroimage*, *21*, 1155–1166.
- Bartsch, K., & Wellman, H. M. (1989). Young children's attribution of action to beliefs and desires. *Child Development*, *60*, 946–964.

- 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.
- Belsky, J. (1999a). Interactional and contextual determinants of attachment security. In J. Cassidy & P. R. Shaver (Eds.), *Handbook of attachment: Theory, research and clinical applications* (pp. 249–264). New York: Guilford.
- Belsky, J. (1999b). Modern evolutionary theory and patterns of attachment. In J. Cassidy & P. R. Shaver (Eds.), *Handbook of attachment: Theory, research and clinical applications* (pp. 141–161). New York: Guilford.
- Belsky, J., & Fearon, R. M. (2002). Infant-mother attachment security, contextual risk, and early development: A moderation analysis. *Development and Psychopathology*, 14, 293–310.
- Bion, W. R. (1962). A theory of thinking. *International Journal of Psychoanalysis*, 43, 306–310.
- Bion, W. R. (1963). *Elements of psychoanalysis*. London: Heinemann.
- Birch, S. A., & Bloom, P. (2003). Children are cursed: An asymmetric bias in mental-state attribution. *Psychol Sci*, 14, 283–286.
- Bogdan, R. J. (1997). *Interpreting minds*. Cambridge, MA: MIT Press.
- Bronfman, E., Parsons, E., & Lyons-Ruth, K. (1999). *Atypical Maternal Behavior Instrument for Assessment and Classification (AMBLANCE): Manual for coding disrupted affective communication, version 2. Unpublished manuscript*. Cambridge, MA: Harvard Medical School.
- Brown, J. R., Donelan-McCall, N., & Dunn, J. (1996). Why talk about mental states? The significance of children's conversations with friends, siblings, and mothers. *Child Development*, 67, 836–849.
- Brown, R., Hobson, R. P., Lee, A., & Stevenson, J. (1997). Are there "autistic-like" features in congenitally blind children? *Journal of Child Psychology and Psychiatry*, 38, 693–703.
- Cabeza, R., & Nyberg, L. (2000). Neural bases of learning and memory: Functional neuroimaging evidence. *Curr Opin Neurol*, 13, 415–421.
- Calkins, S., & Johnson, M. (1998). Toddler regulation of distress to frustrating events: Temperamental and maternal correlates. *Infant Behavior & Development*, 21, 379–395.
- Cavell, M. (1991). The subject of mind. *International Journal of Psycho-Analysis*, 72, 141–154.
- 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. *Journal of Neuroscience*, 24, 4113–4123.
- Charman, T., Ruffman, T., & Clements, W. (2002). Is there a gender difference in false belief development? *Social Development*, 11, 1–10.
- Chugani, H. T., Behen, M. E., Muzik, O., Juhasz, C., Nagy, F., & Chugani, D. C. (2001). Local brain functional activity following early deprivation: A study of postinstitutionalized Romanian orphans. *Neuroimage*, 14, 1290–1301.
- Cicchetti, D., Rogosch, F. A., & Toth, S. L. (2000). The efficacy of toddler-parent psychotherapy for fostering cognitive development in offspring of depressed mothers. *Journal of Abnormal Child Psychology*, 28, 135–148.
- Claussen, A. H., Mundy, P. C., Mallik, S. A., & Willoughby, J. C. (2002). Joint attention and disorganised attachment status in infants at risk. *Development and Psychopathology*, 14, 279–291.
- 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.
- Conway, M. A., & Holmes, A. (2004). Psycho-social stages and the accessibility of autobiographical memories across the life cycle. *Journal of Personality*, 72, 461–480.
- Corkum, V., & Moore, C. (1995). Develop-

- ment of joint visual attention in infants. In C. Moore & P. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 61–83). New York: Erlbaum.
- Cote, S., Tremblay, R. E., Nagin, D., Zoccolillo, M., & Vitaro, F. (2002). The development of impulsivity, fearfulness, and helpfulness during childhood: Patterns of consistency and change in the trajectories of boys and girls. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *43*, 609–618.
- Csibra, G., & Gergely, G. (1998). The teleological origins of mentalistic action explanations: A developmental hypothesis. *Developmental Science*, *1*, 255–259.
- Cutting, A. L., & Dunn, J. (2002). The cost of understanding other people: Social cognition predicts young children's sensitivity to criticism. *Journal of Child Psychology and Psychiatry*, *43*, 849–860.
- Davidson, D. (1983). *Inquiries into truth and interpretation*. Oxford: Oxford University Press.
- Dennett, D. (1987). *The intentional stance*. Cambridge, Mass: MIT Press.
- Dias, M. G., & Harris, P. L. (1990). The influence of the imagination on reasoning by young children. *British Journal of Developmental Psychology*, *8*, 305–318.
- Dunn, J. (1988). *The beginnings of social understanding*. Oxford: Basil Blackwell Ltd. and Cambridge, MA: Harvard University Press.
- Dunn, J. (1994). Changing minds and changing relationships. In C. Lewis & P. Mitchell (Eds.), *Children's early understanding of mind: Origins and development* (pp. 297–310). Hove, UK: Lawrence Erlbaum.
- 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* (Vol., pp. 229–242). Mahwah, NJ: Lawrence Erlbaum Associates.
- Dunn, J., & Brown, J. (2001). Emotion, pragmatics and developments in emotion understanding in the preschool years. In D. Bakhurst & S. Shanker (Eds.), *Jerome Bruner: Language, culture, self*. Thousand Oaks, CA: Sage.
- 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., & Cutting, A. (1999). Understanding others, and individual differences in friendship interactions in young children. *Social Development*, *8*, 201–219.
- Dunn, J., Davies, L. C., O'Connor, T. G., & Sturgess, W. (2000). Parents' and partners' life course and family experiences: Links with parent-child relationships in different family settings. *Journal of Child Psychology and Psychiatry*, *41*, 955–968.
- Fearon, R. M., & Belsky, J. (2004). Attachment and attention: Protection in relation to gender and cumulative social-contextual adversity. *Child Development*, *75*, 1677–1693.
- Flavell, J., & Miller, P. (1998). Social cognition. In D. Kuhn & R. Siegler (Eds.), *Cognition, perception, and language. Handbook of child psychology* (5 ed., Vol. 2, pp. 851–898). New York: Wiley.
- Flavell, J. H. (1999). Cognitive development: Children's knowledge about the mind. *Annual Review of Psychology*, *50*, 21–45.
- 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). *Early intervention and the development of self-regulation*. Paper presented at the meeting of the Australian Association for Infant Mental Health, Perth, Australia.
- Fonagy, P. (2003). The development of psychopathology from infancy to adulthood: The mysterious unfolding of disturbance in time. *Infant Mental Health Journal*, *24*, 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., Leigh, T., Steele, M., Steele, H., Kennedy, R., Mattoon, G., et al. (1996). The relation of attachment status, psychiatric classification, and response to psychotherapy. *Journal of Consulting and Clinical Psychology*, *64*, 22–31.
- Fonagy, P., Steele, H., Moran, G., Steele, M.,

- & Higgitt, A. (1991). The capacity for understanding mental states: The reflective self in parent and child and its significance for security of attachment. *Infant Mental Health Journal*, *13*, 200–217.
- Fonagy, P., Steele, M., Moran, G. S., Steele, H., & Higgitt, A. (1992). The integration of psychoanalytic theory and work on attachment: The issue of intergenerational psychic processes. In D. Stern & M. Ammaniti (Eds.), *Attaccamento E Psiconalis* (pp. 19–30). Bari, Italy: Laterza.
- Fonagy, P., Steele, M., Steele, H., Higgitt, A., & Target, M. (1994). Theory and practice of resilience. *Journal of Child Psychology and Psychiatry*, *35*, 231–257.
- Fonagy, P., Stein, H., Allen, J., & Fultz, J. (2003a). *The relationship of mentalization and childhood and adolescent adversity to adult functioning*. Paper presented at the Biennial Meeting of the Society for Research in Child Development, Tampa, FL.
- Fonagy, P., & Target, M. (1996). Playing with reality: I. Theory of mind and the normal development of psychic reality. *International Journal of Psycho-Analysis*, *77*, 217–233.
- Fonagy, P., & Target, M. (1997). Attachment and reflective function: Their role in self-organization. *Development and Psychopathology*, *9*, 679–700.
- Fonagy, P., & Target, M. (2000). Playing with reality III: The persistence of dual psychic reality in borderline patients. *International Journal of Psychoanalysis*, *81*, 853–874.
- Fonagy, P., & Target, M. (2002). Early intervention and the development of self-regulation. *Psychoanalytic Inquiry*, *22*, 307–335.
- Fonagy, P., Target, M., & Gergely, G. (2000). Attachment and borderline personality disorder: A theory and some evidence. *Psychiatric Clinics of North America*, *23*, 103–122.
- Fonagy, P., Target, M., Gergely, G., Allen, J. G., & Bateman, A. (2003b). The developmental roots of borderline personality disorder in early attachment relationships: A theory and some evidence. *Psychoanalytic Inquiry*, *23*, 412–459.
- Fonagy, P., Target, M., Steele, H., & Steele, M. (1998). *Reflective-Functioning Manual, version 5.0, for Application to Adult Attachment Interviews*. London: University College London.
- 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.
- Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of “theory of mind.” *Trends in Cognitive Science*, *7*, 77–83.
- Gergely, G. (2001). The obscure object of desire: “Nearly, but clearly not, like me. Contingency preference in normal children versus children with autism.” In J. Allen, P. Fonagy, & G. Gergely (Eds.), *Contingency perception and attachment in infancy, special issue of the bulletin of the Menninger Clinic* (pp. 411–426). New York: Guilford.
- 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. (2005). Social construction of the cultural mind: Imitative pedagogy as a mechanism for social learning. *Interaction Studies*, *6*, 463–481.
- 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.
- 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.
- Greene, J., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Science*, *6*, 517–523.
- 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*, 299–311.

- 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. *Proceedings of the National Academy of Sciences USA*, *98*, 4259–4264.
- Harman, C., Rothbart, M. K., & Posner, M. I. (1997). Distress and intention interactions in early infancy. *Motivation and Emotion*, *21*, 27–43.
- Hobson, P. (2002). *The cradle of thought: Explorations of the origins of thinking*. Oxford: Macmillan.
- 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.
- Hrdy, S. B. (2000). *Mother nature*. New York: Ballentine Books.
- 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., & 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.
- 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? *Physiology & Behavior*, *79*, 351–357.
- Jacobsen, T., & Hofmann, V. (1997). Children's attachment representations: Longitudinal relations to school behavior and academic competency in middle childhood and adolescence. *Developmental Psychology*, *33*, 703–710.
- Jacobson, E. (1954). The self and the object world: Vicissitudes of their infantile cathexes and their influence on ideational affective development. *The Psychoanalytic Study of the Child*, *9*, 75–127.
- 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 (Berlin)*, *181*(1), 8–15.
- Kochanska, G., Coy, K. C., & Murray, K. T. (2001). The development of self-regulation in the first four years of life. *Child Development*, *72*, 1091–1111.
- 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.
- Lalonde, C., & Chandler, M. J. (1995). False belief understanding goes to school: On the social-emotional consequences of coming early or late to a first theory of mind. *Cognition and Emotion*, *9*, 167–185.
- 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. (1994). TOMM, ToBy, and agency: Core architecture and domain specificity. In L. Hirschfeld & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 119–148). New York: Cambridge University Press.
- Liszkowski, U., Carpenter, M., Henning, A., Striano, T., & Tomasello, M. (2004). Twelve-month-olds point to share attention and interest. *Developmental Science*, *7*, 297–307.
- 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 in Neuroscience*, *22*, 310–316.
- 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. *American Journal of Psychiatry*, *156*, 675–682.
- Mayes, L. C. (2000). A developmental perspective on the regulation of arousal states. *Seminars in Perinatology*, *24*, 267–279.

- Meins, E. (1997). *Security of attachment and the social development of cognition*. London: Psychology Press.
- Meins, E., & Fernyhough, C. (1999). Linguistic acquisitional style and mentalizing 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 mentalizing 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 Development, 74*, 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., Fernyhough, 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. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology, 31*, 838–850.
- Mitchell, R. W. (1993). Mental models of mirror self-recognition: Two theories. *New Ideas in Psychology, 11*, 295–325.
- Mitchell, S. A. (2000). *Relationality: From attachment to intersubjectivity*. Hillsdale, NJ: Analytic Press.
- 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.
- Nagin, D. S., & Tremblay, R. E. (2001). Parental and early childhood predictors of persistent physical aggression in boys from kindergarten to high school. *Archives of General Psychiatry, 58*, 389–394.
- Neisser, U. (1988). Five kinds of self-knowledge. *Philosophical Psychology, 1*, 35–59.
- 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.
- 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*, 583–592.
- 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.
- Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. Oxford: Oxford University Press.
- Plotsky, P. M., Thirivikraman, 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, 181*(1), 8–15.
- Povinelli, D. J., & Eddy, T. J. (1995). The unduplicated self. In P. Rochat (Ed.), *The self in infancy: Theory and research* (pp. 161–192). Amsterdam: Elsevier.
- Pratt, C., & Bryant, P. E. (1990). Young children understand that looking leads to knowing (so long as they are looking into a single barrel). *Child Development, 61*, 973–982.
- Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology, 33*, 12–21.
- Rochat, P., & Striano, T. (1999). Social-cognitive development in the first year. In P. Rochat (Ed.), *Early social cognition*. Mahwah, NJ: Lawrence Erlbaum.
- Ruffman, T., Perner, J., Naito, M., Parkin, L., & Clements, W. (1998). Older (but not younger) siblings facilitate false belief understanding. *Developmental Psychology, 34*, 161–174.
- Ruffman, T., Perner, J., & Parkin, L. (1999). How parenting style affects false belief understanding. *Social Development, 8*, 395–411.

- Rutter, M., & O'Connor, T. G. (2004). Are there biological programming effects for psychological development? Findings from a study of Romanian adoptees. *Developmental Psychology, 40*, 81–94.
- Schechter, D. S., Coots, 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 post-traumatic stress and reflective functioning. *Attachment and Human Development, 7*, 313–331.
- Schölmerich, A., Lamb, M. E., Leyendecker, B., & Fracasso, M. P. (1997). Mother-infant teaching interactions and attachment security in Euro-American and Central-American immigrant families. *Infant Behavior and Development, 20*, 165–174.
- Segal, H. (1957). Notes on symbol formation. *International Journal of Psycho-Analysis, 38*, 391–397.
- Siegel, D. J. (1999). *The developing mind: Toward a neurobiology of interpersonal experience*. New York: Guilford.
- Simpson, J. A. (1999). Attachment theory in modern evolutionary perspective. In J. Cassidy & P. R. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications* (pp. 115–140). New York: Guilford.
- Slade, A. (2005). Parental reflective functioning: An introduction. *Attachment and Human Development, 7*, 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*, 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.
- Sodian, B., & Frith, U. (1992). Deception and sabotage in autistic, retarded and normal children. *J Child Psychol Psychiatry, 33*, 591–605.
- Sodian, B., Taylor, C., Harris, P. L., & Perner, J. (1992). Early deception and the child's theory of mind: False trails and genuine markers. *Child Development, 62*, 468–483.
- Stroufe, L. A. (1996). *Emotional development: The organization of emotional life in the early years*. New York: Cambridge University Press.
- Sutton, J., Smith, P. K., & Swettenham, J. (1999). Social cognition and bullying: Social inadequacy or skilled manipulation? *British Journal of Developmental Psychology, 17*, 435–450.
- Target, M., & Fonagy, P. (1996). Playing with reality II: The development of psychic reality from a theoretical perspective. *International Journal of Psycho-Analysis, 77*, 459–479.
- Taylor, M., & Carlson, S. M. (1997). The relation between individual differences in fantasy and theory of mind. *Child Development, 68*, 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 Development, 65*, 1581–1604.
- Tremblay, R. E. (2000). The origins of violence. *ISUMA(Autumn)*, 19–24.
- Tremblay, R. E., Japel, C., & Perusse, D. (1999). The search for the age of onset of physical aggression: Rousseau and Bandura revisited. *Criminal Behavior and Mental Health, 9*, 8–23.
- 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.
- 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.
- Waters, E., Merrick, S. K., Treboux, D., Crowell, J., & Albersheim, L. (2000). Attachment security from infancy to early adulthood: A 20 year longitudinal study. *Child Development, 71*, 684–689.
- 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.
- 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 Development, 72*, 655–684.
- Wellman, H. M., & Phillips, A. T. (2000). Developing intentional understandings. In L. Moses, B. Male & D. Baldwin (Eds.), *Intentionality: A key to human understanding*. Cambridge, MA: MIT Press.
- Widom, C. S. (1999). Posttraumatic stress disorder in abused and neglected children grown up. *Am J Psychiatry, 156*, 1223–1229.
- Winnicott, D. W. (1956). Mirror role of mother and family in child development. In D. W. Winnicott (Ed.), *Playing and reality* (pp. 111–118). London: Tavistock.
- Winnicott, D. W. (1965). *The maturational process and the facilitating environment*. London: Hogarth Press.
- Winston, J. S., Strange, B. A., O'Doherty, J., & Dolan, R. J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature Neuroscience, 5*, 277–283.
- Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., & Chapman, M. (1992). Development of concern for others. *Developmental Psychology, 28*, 126–136.
- 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. *Journal of Neuroscience, 25*, 1493–1502.