Title: Mutual regulation, mentalization and therapeutic action: A reflection on the contributions of Ed Tronick to developmental and psychotherapeutic thinking

Head: Mutual regulation and mentalization

Key words: Mutual regulation, mentalization, therapeutic action, attachment theory

Author: Peter Fonagy, PhD FBA
             Psychoanalysis Unit
             Research Department of Clinical, Educational and Health Psychology
             University College London
             Gower Street
             London WC1E 6BT

                   Tel: +44 (0)20 7679 1943
             Fax: +44 (0)20 7916 8502
             p.fonagy@ucl.ac.uk

Bio: Peter Fonagy, PhD, FBA, is Freud Memorial Professor of Psychoanalysis and Head of the Research Department of Clinical, Educational and Health Psychology at University College London; Chief Executive of the Anna Freud Centre, London; and Consultant to the Child and Family Program at the Menninger Department of Psychiatry and Behavioral Sciences at the Baylor College of Medicine. His work integrates empirical research with psychoanalytic theory, and his clinical interests center around borderline psychopathology, violence, and early attachment relationships.

Abstract: This paper celebrates Ed Tronick's contribution to psychodynamic developmental psychology. Tronick's ideas have implicitly and explicitly influenced late 20th-century and early 21st-century psychoanalytic thinking, and his contribution is here acknowledged in relation to the work on the developmental ideas concerning mentalizing. Tronick's mutual regulation model (MRM) is outlined and examined from the point of view of the attachment-theory-based mentalization model of interpersonal interaction and psychopathology. The article identifies common ground between clinical implications of Tronick's MRM and the mentalization model as independent but complementary ways of expanding ideas originating in attachment theory, and highlights ways in which the MRM provides greater clarity for several concepts which are frequently used in writings about mentalizing such as the background of safety and the holding environment.

Mutual regulation, mentalization and therapeutic action: A reflection on the contributions of Ed Tronick to developmental and psychotherapeutic thinking

Introduction

Ed Tronick's contribution to psychodynamic developmental psychology has been colossal. It stands alongside the work of other giants of the field like Renee Spitz, John Bowlby, Dan Stern, and a tiny handful of others who made the past few decades of psychoanalytic scholarship exciting and innovative. In this tribute to Tronick's work I will start from my own secure base, attachment theory, and try to show how Tronick's influence provides an appropriate counterpoint to the hegemony of attachment-theory-inspired speculation in developmental psychoanalysis. I will then selectively compare Tronick's mutual regulation model (MRM) with our mentalization model (MM). I will end the comparison by identifying overlaps in a review of the therapeutic process from both perspectives.

Attachment theory was almost unique among psychoanalytic theories in bridging the gap between general psychology and clinical psychodynamic theory. John Bowlby was dissatisfied with prevailing views of both in the first half of the 20th century, at least as far as those concerned the origin of affectional bonds. At that time, psychology and psychoanalysis were uncomfortably close. Both Freudian psychoanalytic and Hullian learning theory stressed that the emotional bond to the primary caregiver was a secondary drive, based on the gratification of oral needs. Yet evidence was already available that in the animal kingdom at least, the young of the species could become attached to adults who did not feed them (Lorenz, 1935). This led Bowlby (1958) to the discovery that the human infant enters the world predisposed to participate in social interaction. Developmental psychology has since made this discovery something of a truism (e.g., Meltzoff, 1995; Watson, 1994). Around the mid-point of the last century, however, Bowlby's determination to give central place to the infant's biological proclivity to form attachments—and to initiate, maintain, and terminate

interaction with the caregiver and use this person as a "secure base" for exploration and self enhancement—flew in the face of prevailing psychological and psychoanalytic dogma. Neither Anna Freud nor ego psychology had a place for attachment theory. Strangely, British object relations theory never fully embraced Bowlby's work. It was overly influenced by Melanie Klein's complex speculative and improbable model of the baby's psyche and thus probably found Bowlby's natural science construction barren and clinically of little value. To this day, John Bowlby and attachment theory are not formally taught in the British Psychoanalytic Institute.

Following on from over 30 years of Bowlby's work, the 1980s saw the beginning of an integration of relational approaches augmented and modified by the intersubjectivist vision of philosophically oriented psychoanalysts such as Stolorow (1997) and great infant researchers such as Robert Emde (Emde, Kubicek, & Oppenheim, 1997). Relational theory posited that the infant is oriented towards the outside world from birth; the baby's mind is already organized and becomes increasingly complex and integrated as it meets a fully supportive caregiving environment and interacts in dyadic structures such as the infant– parent relationship.

It was in this context, building on these developments, that Ed Tronick has made his important and lasting conceptual contributions to the field. In particular, he has deepened our understanding of the uniqueness of attachment relationships and proposed an important model for how these relationships increasingly differentiate themselves (Tronick, 2003). This work has taken us beyond attachment theory's early, reductive views on proximity, and has helped us to conceptualize the nuances of the mother–infant relationship and the mechanisms through which attachment is achieved and maintained.

In particular, by pointing to the uniqueness of each attachment bond, Tronick's contribution has challenged the "view that the child's relationship with the mother is

3

paradigmatic for later coming relationships" (Tronick, 2008). At the same time, however, Tronick's model radically revised our understanding of how relationships do influence one another (Tronick, 2005, 2007). Bowlby's model of relationship representations was overly influenced by the then prevailing cognitivist model of mind excessively colored by the mind-computer metaphor (Fonagy & Target, 2007b). Tronick's model updated the notion of the internal working model, replacing it with a theory of relationship representations that is at once subtler and more dynamic. Tronick's model, focused on dyadic meaning-making, came into its own when applied to understanding the therapeutic relationship and its effects.

Our own work, developing in parallel but constantly benefitting from Tronick's empirical and theoretical work, also aims to link affectional bonds to meaning-making to provide an alternative to Bowlby's understanding of the function and importance of attachment relationships. Our model, based around the concept of mentalization, is in many ways complementary to Tronick's. At the same time, the difference in emphasis between our respective models creates an important dialogue that we believe can help deepen understandings of attachment, therapeutic action, and the development of the human mind.

Going beyond the reductionism of attachment theory

Bowlby's attachment theory, like classical psychoanalysis, has a biological focus (see especially Bowlby, 1969). Attachment readily reduces to a "molecular" level of infant behaviors, such as smiling and vocalizing, that alert the caregiver to the child's interest in socializing, and bring him or her into close proximity with the child. Bowlby's critical contribution was his unwavering focus on the infant's need for an unbroken (secure) early attachment to the mother. Bowlby emphasized the survival value of attachment in enhancing safety through proximity to the caregiver in addition to feeding, learning about the environment and social interaction, as well as protection from predators.

While giving central place to parenting in general and the infant–mother relationship in particular, Bowlby's early work was, however, silent on the mechanisms by which "maternal deprivation" might be expected to generate adverse consequences. Similarly, in the first volume of the *Attachment and Loss* trilogy, Bowlby (1969) was not yet clear about how attachment behavior functioned beyond the termination of the system once physical proximity was ensured. Proximity was the set goal of the attachment system; its measurement was simple and purely behavioral. The absence of the attachment figure generates the biological need; her return and presence turns it off. Perhaps it should be no surprise then that psychoanalysts were horrified by this apparently simplistic approach, which bore the hallmarks of the worst excesses of behaviorist reductionism.

But Bowlby's system, like Freud's, evolved—although, as with Freud, Bowlby's critics were often apparently "imprinted" with the initial model; their attitude of hostility did not permit noting the change in Bowlby's view. In the second volume of the *Attachment and Loss* trilogy, Bowlby established the set goal of the attachment system as maintaining *the caregiver's accessibility and responsiveness*, which he covered with a single term: "availability" (Bowlby, 1973, p. 202). In fact, it was not until the third section of the book that he addressed the critical role of appraisal in the operation of the attachment system. Here he asserts that "availability" means confident expectation, gained from "tolerably accurately" (p. 202) represented experience, over a significant time period, that the attachment figure will be available. The attachment behavioral system thus came to be underpinned by a set of cognitive mechanisms, discussed by Bowlby as representational models or following the pioneering psychologist Craik (1943) as internal working models.

The positing of a representational system underpinning attachment permitted a far more sophisticated consideration of individual differences (Bowlby, 1973, 1980). Given the power of the biological forces driving the human attachment system, Bowlby assumed that

almost all human beings will become attached. Non-attachment, or the disorganization of the attachment system, was not in John Bowlby's dictionary; attachment could be only secure or insecure. Secure attachment implied a representational system where the attachment figure was seen as accessible and responsive when needed. Anxious attachment implied a somewhat dysfunctional system where the responsiveness of the caregiver was not assumed and the child adopted strategies for circumventing the perceived unresponsiveness of the attachment figure (Ainsworth, Blehar, Waters, & Wall, 1978). What is remarkable in this formulation is Bowlby's anticipation of what later empirical research demonstrated to be the key mechanism communicating to the infant that he/she should prioritize secure over insecure strategies in interpersonal interactions (Belsky & Fearon, 2008). Bowlby was prescient in assuming that caregiver responsiveness was critical in determining the security of the attachment system: "the extent to which the mother has permitted clinging and following, and all the behavior associated with them, or has refused them" (Bowlby, 1958, p. 370).

Thus, the central feature of the internal working model concerned the infant encoding interactions in terms of what they implied about the expected availability of the attachment figure. Bowlby also envisioned a complementary working model of the self. The key feature of this was how acceptable or unacceptable the child felt in the eye of the attachment figure. A child whose internal working model of the caregiver was focused around rejection was expected to evolve a complementary working model of the self as unlovable, unworthy, and flawed. Although not explicitly stated by Bowlby, these models of the attachment figure and the self were somewhat transactional, interactive models representing self–other relationships. The explanatory power of Bowlby's model rested in his proposal that these cognitive models encoding expectation were capable of providing prototypes for all later relationships.

Bowlby's original concept has been thoughtfully elaborated by some of the greatest minds in the attachment field, and no attempt to review these exhaustively can be undertaken here (see Bretherton & Munholland, 2008). In our analysis of Bowlby's model (Fonagy, 1998; Fonagy & Target, 2007b) we identified four representational systems that are implied by his reformulations: (1) expectations of interactive attributes of early caregivers, which are created in the first year of life and subsequently elaborated; (2) event representations, by which general and specific memories of attachment-related experiences are encoded and retrieved; (3) autobiographical memories, by which specific events are conceptually connected because of their relation to a continuing personal narrative and developing self-understanding; and (4) understanding of the psychological characteristics of other people (inferring and attributing causal motivational mind states such as desires and emotions and epistemic mind states such as intentions and beliefs) and differentiating these from those of the self. Implied in this analysis is the recognition that the attachment system serves an epistemic function beyond Bowlby's emphasis on a model of interpersonal expectancy. Specifically, our formulation links the notion of "coherence", which, following Mary Main's groundbreaking work, became the hallmark of the secure internal working model (Main, Hesse, & Goldwyn, 2008), to the parent–infant relationship via the construct of "mentalization", the capacity with which the secure attachment relationship endows the infant (Fonagy, Gergely, & Target, 2007). The caregiver's capacity to mentalize the infant—to accurately assign to him/her feelings, wishes and beliefs—underpins the construct of sensitivity and, through a process of internalization, forms the basis of a second-order, metacognitive understanding of internal states (Fonagy, Steele, Steele, Moran, & Higgitt, 1991). In synergy with Tronick, we have suggested that the evolutionary function of the dyadic relationship between parent and human infant goes way beyond ensuring the safety of the latter, as the attachment system was "captured" by evolution to provide a platform for the

transmission of cultural knowledge across generations, most particularly understanding about the nature of subjectivity and the symbolic functioning of the human mind (Fonagy, Gergely, Jurist, & Target, 2002).

In our view, the major evolutionary advantage of attachment in humans is the opportunity it gives the infant to develop social intelligence. Alan Sroufe (1996) and Myron Hofer (2003) played a seminal role in extending the scope of attachment theory from an account of the developmental emergence of a set of social expectations to a far broader conception of attachment as an organizer of physiological and brain regulation. Our work simply extended their ideas. Attachment ensures that the brain processes that serve social cognition are appropriately organized and prepared to enable us to live and work with other people.

We first developed our concept of mentalization in the context of a large empirical study in which the security of infant attachment with each parent turned out to be strongly predicted by the parent's ability to think about and understand their childhood relationship to their own parents in terms of states of mind (Fonagy, Steele, Moran, Steele, & Higgitt, 1991). We proposed that there was a vital synergy between attachment processes and the development of the child's ability to understand interpersonal behavior in terms of mental states (Fonagy et al., 2002; Fonagy, Redfern, & Charman, 1997).

Alongside this empirical research, inspiration for the development of the concept of mentalization also came from psychoanalytic work with borderline patients (Fonagy & Bateman, 2006). In an early paper effectively co-authored with George Moran, we identified the repudiation of a concern with mental states as a key aspect of borderline psychopathology (Fonagy, 1991). The basic suggestion is that the capacity for representing self and others as thinking, believing, wishing, or desiring does not simply arrive at age 4 as an inevitable consequence of maturation; rather, it is a developmental achievement that is profoundly

rooted in the quality of early relationships (Fonagy, 1989). Its liability to disappear under stress in borderline conditions was seen as an appropriate focus for psychoanalytically oriented psychological intervention (Fonagy, 1989).

A second line of analytic inspiration came from work with children undertaken as part of a project to construct a manual for child analysis and subsequent work in developmental science by Mary Target and Peter Fonagy (Fonagy et al., 2002; Fonagy & Target, 1996, 2000, 2007a; Target & Fonagy, 1996). This work helped us to think more deeply about the normal development of thinking or mentalizing capacity, and the more primitive modes of thought that precede its emergence. In trying to map the emergence of mentalization on the basis of material from records of child analysis and clinical and research work with children in other contexts, we came up with a heuristic map of the emergence of mentalization that turned out to be extremely valuable in understanding some qualitative aspects of the thinking of patients in borderline states. In particular, we noticed that the types of thinking that many have identified as a hallmark of borderline personality disorder (Higgitt & Fonagy, 1992) were not dissimilar to the ways young children normally tend to treat their internal experience.

Tronick's model of the roots of meaning in interaction

Tronick's thinking also aims at going forward to locate a coherent self-structure beyond the simplistic notion of expectation. But, while our model leaves a gap in relation to the mechanisms whereby mental states are internalized through the process of parent–infant interaction, Tronick's model is detailed and specific with regard to the way in which meaning and coherence emerge out of interactive processes focused on mutual regulation. Tronick's thinking rapidly overtook and has now gone substantially beyond the schematic, proximity-focused attachment models that once dominated the field. His (Tronick, 2007; Tronick, 1989) mutual regulation model (MRM) of infant–adult interaction looks at the subtle, non-

verbal, micro-regulatory and social-emotional processes that unfold in mother–infant interactions. Tronick understood the infant on the model of an open thermodynamic system that must constantly take in energy and work towards coherence in order to stave off dissipation. The MRM postulates that infants have "self-organizing neurobehavioral capacities" and "biopsychological processes" that allow the infant to "organize behavioral states" and make "sense of themselves and their place in the world" (Tronick, 2007). At the same time, however, Tronick pointed to the limits of the self-organizing capacities and stated that they need to be supplemented by a "larger dyadic regulatory system" in which the infant participates with the caregiver. In this way Tronick brings together the notion of sensitivity with the overriding construct of meaning-making. Regulation in the MRM is accomplished through the operation of a communication system in which the infant communicates its regulatory status to the caregiver, who responds to the meaning of the communication. This communication is expressed through the totality of the infant's and caregiver's biopsychological processes—including the words or sounds each uses, their momentary changes in facial expression, the quality of their touch, and even changes in their body odor (Tronick, 2008).

Successful mutual regulation between the partners is achieved when an infant and caregiver together generate, communicate, and integrate meaningful elements of consciousness. This creates a synchrony in implicit relational knowing, meaning that each can anticipate and "know" the moves of the other. This "knowing" is initially of a pattern of physiological responses or activations rather than of intentional states, although quite clearly it can be the platform for knowing of intention given developmental time. The parent–infant collaboration results in a singular, organized dyadic state that is believed to be more than the sum of its parts. Assuming a six-month-old infant to be capable of apprehending another's state of consciousness, Tronick argues that a mutual mapping of each other's state of

consciousness does take place. Each individual's sense of self is augmented by the consciousness, meanings, and representation of the other, as well as by representations of the relationship as a whole. This leads to what Tronick terms a state of "co-creativity" (Tronick, 2003) in which infant and caregiver shape their relationship through a process of mutual regulation. Co-creativity is not seen as a mechanistic series of steps, nor is it assumed to be an end state, but rather, more realistically, it is described as part of the unique and continuous unfolding of the parent–child relationship.

Also constitutive to the uniqueness of each relationship, Tronick importantly points out, are the "time-activity contexts" in which co-creativity unfolds (Tronick, 2003). These contexts consist of the mutually regulated, shared experiences of the dyad, which are organized around certain times and behaviors (such as feedings, diaper changings, putting to bed times, etc.). Together, the number, frequency, and qualities of these interactions make up what Tronick terms that "thickness" of the relationship (Tronick, 2003)—that is, the depth of knowing how to be together in various ways and in various contexts that informs and supports further co-creativity. Furthermore, Tronick notes that often, looked at from the outside, the meaning of a piece of behavior between individuals is "cryptic": it cannot be understood without being experienced (Tronick, 2003). This explains both the skepticism that many clinicians feel about using film and video to "capture" the relationship between parent and child (e.g., Green, 2000) and also why much of the knowing in one relationship—that is, its thickness—is non-transferable to other relationships (Tronick, 2003).

The limitation of Tronick's model from our standpoint rests in its emphasis on idiosyncrasy. From our point of view, what needs explaining is the way the parent–infant relationship provides an evolutionarily protected and privileged foundation for the transmission of human culture. The MRM, while explaining much about relationship experience and perhaps even the development of consciousness, does little to teach us about

how knowledge of the mind can become part of an intergenerational process (Fonagy & Target, 2007a). Recently, however, Tronick has made an important addition to his conceptualization of the time-activity context. This is a reflection on how culture permeates the activities of the mother–infant dyad (Tronick, 2007). "The ways [the mother] moves and talks and interacts with her child as well as with others", says Tronick, "are not simply expressions of her uniqueness; they are a manifestation of a cultured uniqueness. Her 'cultured way' is transferred to her child, affecting what the child experiences and how the child experiences him- or herself in the world" (Tronick, 2007, pp. 7–8). Thus, the unique quality and thickness of any relationship is now considered by Tronick to be influenced by a shared cultural context (which is, in turn, uniquely manifested in any particular individual).

Rejecting attachment theory and other narrative models that claim all relationships to be influenced by a parental prototypic model, Tronick argues against model forms for relationships. Instead, he brought a radical new perspective, dynamic systems theory, to the study of parent–infant interaction. He proposed that not only is each relationship with another person unique, but so too is each interaction with each person. At the same time, however, Tronick emphasizes the interconnectedness of relationship representations both within individual relationships and across relationships. He accounts for this by proposing, based on the work of Freeman (1994), that every single interaction a person has, although itself unique, can potentially alter our memory of past interactions with that or any other person (Tronick, 2007). Relational activation patterns (RAPs) are micro-interaction patterns that accumulate and reside in experiential relational space (ERS). Together, these RAPs enable implicit relational knowing. Typically, the first of these influential RAPs is the infant–mother relationship, although each one is fluid and changes with each interaction. In spite of their fluid nature, RAPs are also stable because of the individual's continuing perception of the ERS—that is, the context of each RAP. The infant recognizes a familiar

"place" in a relationship or interaction (e.g., "now we are separating"). As each interaction subtly alters the RAPs, it creates new possibilities for co-creativity and for ways of being together.

So, how can the coherence of relationship narratives emerge from RAPs? There may even be something oxymoronic between the dynamic systems theory construction of the mother–infant relationship and the notion of coherence that emerges out of the attachment theory–expectancy hypothesis. Tronick's solution is both elegant and creative, going beyond our mentalization-based formulation as well as the classical Main–Hesse theory (Hesse, 2008). Tronick (2007) suggests that miscommunication and "messiness" lie at the heart of the development of the self and self-regulation. Miscommunication creates negative affect, but, when interactive errors are repaired, the negative is replaced by positive affect in both infant and mother. These intense experiences in the dyadic relationship generate "coherence" of mother and infant, deepening their dyadic state of consciousness. This state expands the awareness of each, changing their RAPs and therefore their interpretation of all relationships, both old and new. In disruptions and repair of ongoing regulations, expectations are violated and ensuing efforts to resolve these breaches are hypothesized to underpin the generation of further psychic structure (see also Blatt & Behrends, 1987; Blatt & Luyten, 2009).

The mentalization model and Tronick's discoveries

The mentalization model (MM) of attachment that we have proposed, as we have tried to show, is in most ways complementary to and benefits from Tronick's formulations. Like Tronick's conjectures, our theories were partly developed in response to the limits of attachment theory, and see the evolutionary role of attachment as going far beyond giving physical protection to the human infant.

As summarized above, both MRM and MM maintain that none of us is born with the capacity to regulate our own emotional reactions or represent intentionality. The MRM is

specific about how the caregiver understands and responds to the newborn infant's signals of moment-to-moment changes in his/her state, such that a dyadic regulatory system gradually evolves. The infant learns that mutual regulatory processes will protect him/her from overwhelming emotional arousal while in the caregiver's presence, and when he/she starts to feel overwhelmed, the infant will seek or signal to the caregiver in the hope of understanding, soothing, and the recovery of homeostasis.

As with the MRM, the MM has also developed over recent years and perhaps now more effectively complements Tronick's interactional scheme. We have come to conceive of mentalization as a multidimensional construct whose core processing dimensions are underpinned by distinct neural systems (Luyten, Fonagy, Lowyck, & Vermote, 2012). Thus, mentalization involves both a self-reflective and an interpersonal component, it is based on both observing others and reflecting on their mental states, it is both implicit and explicit, and it concerns both feelings and cognitions (Fonagy & Luyten, 2009; Luyten et al., 2012). When they are working together in optimal combination, the neural systems underpinning these components enable the child to represent causal mental states, distinguish inner from outer reality, infer others' mental states from subtle behavioral and contextual cues, moderate behavior and emotional experience, and construct representations of his/her own mental states from perceptible cues (arousal, behavior, and context). In this regard the MRM and MM have moved toward common ground.

By contrast, we may argue that the way the interpersonal dynamic system between caregiver and infant is constructed is conceptualized in a distinct way by the two models, at least insofar as the MM increasingly emphasizes stable individual predispositions in relation to interpersonal encounters (Fonagy & Luyten, 2009). Increasing evidence suggests that the formation of attachment relationships is supported by at least two neurobiological systems: (1) linking attachment experiences to reward and pleasure, motivating the caregiver (and in

all likelihood the infant as well) to seek experiences of closeness; and (2) linking enhanced social understanding to the attachment context, with closer bonds triggering biological systems that are likely to enhance sensitivity to social cues. Given the availability of a neurobiological pathway, what can the developmental psychology literature tell us about the link between attachment and mentalization? If attachment underpins the emergence of mentalization, we would expect securely attached children to outperform insecurely attached children in this domain (measured as passing theory of mind tasks earlier). Many studies indeed support this hypothesis (see Fonagy & Luyten, 2009 for a review). Generally, it seems that the expectation of understanding from the caregiver leading to a secure attachment pattern generates better mentalization and interpersonal understanding, possibly because both functions may be subject to similar social (contextual or cultural) influences. Two decades of research have confirmed that parenting is the key determinant of attachment security. Mindful parenting probably enhances both attachment security and mentalization. These overlapping attributes appear to be associated with both secure attachment and mentalization in the child (see Sharp, Fonagy, & Goodyer, 2006).

Tronick's work, however, reminds us not to treat causality too linearly in this context. While in the MM these correlations have always been understood as parent-to-child effects, they can be just as readily explained as child-to-parent effects, or in the MRM as part of a dynamic system. For example, less power-assertive parenting may be associated with mentalization (Pears & Moses, 2003) not because this may facilitate independent thinking, but because less mentalizing children may be more likely to elicit controlling parenting behavior. It may also be that the same aspects of family functioning that facilitate secure attachment also facilitate the emergence of mentalizing. Indeed, the process of acquiring mentalization is so ordinary and normal that it may be more appropriate to consider secure

attachment as providing an environment that is free of obstacles to its development rather than providing active and direct facilitation.

Recent research on early development makes it clear that infants have genetically inbuilt "healthy" social expectations (e.g., Baillargeon, Scott, & He, 2010). So, how can we integrate nativist and environmentalist perspectives of the development of subjectivity? Mentalization will develop along biologically set paths as long as the child's social experience is developmentally "good enough". By this we mean that the environment has to comply with biologically preset expectations about responsiveness and fit in with biologically prepared mechanisms that evolved to transmit human culture. The environment has to be chronologically consistent with neural development (i.e., matching the developing brain's capacity to integrate new information and provide the social stimuli that the brain requires for its development). Violations of these inbuilt expectations are toxic because they not only "teach" inappropriate content but also undermine mechanisms for the social acquisition of knowledge and the emergence of an agentive sense of self. The severe early neglect of the Romanian children adopted into the UK offers an example (Colvert, Rutter, Beckett, et al., 2008; Colvert, Rutter, Kreppner, et al., 2008).

So, how do we understand the psychological mechanisms that may mediate the impact of early deprivation on the development of subjectivity against a genetically programmed readiness of social cognition? In more recent formulations of our model (Fonagy et al., 2007) we have incorporated the natural pedagogy theory advanced by the Hungarian developmentalists Csibra and Gergely (2009, 2011). These authors have suggested that all humans are born with a cue-driven social-cognitive adaptation of mutual design dedicated to ensure the efficient transfer of relevant cultural knowledge. Parents are predisposed to "teach" and infants to "learn" new and culturally relevant information from each other. Our communication system has been adapted for the efficient transmission of

shared cultural knowledge from one generation to the next. Csibra and Gergely assume that specific communicational (ostensive) cues, such as special types of intonation (so-called "motherese") and, particularly, marked contingent responding to an infant, function as cues to alert the infant that the parent or caregiver is about to transmit information that has relevance beyond the particular situation.

This insight—at first sight apparently trivial but actually very profound—is of enormous significance in understanding the aspects of interpersonal interaction between child and parent that are prioritized by the developmental process and may link closely to Tronick's description of the MRM process. For all of us the most important aspect of human culture concerns our subjectivity, which, as we have become increasingly aware, is not a culturally invariant absolute but a social construct shaped by the social experiences of infancy (Gergely & Csibra, 2005). Attachment to the caregiver may be important precisely because through sensitivity and the contingent interaction that are the markers of caregiving behavior, as Tronick beautifully describes, and the root of the attachment response, a sense of epistemic trust is created in the child. Feeling securely attached to a particular adult also tells the child that the information relayed by that person may be trusted and learned. We have known for many years that secure attachment represents a substantial cognitive advantage for the child (Thompson, 2008). Secure attachment is engendered by ostensive cues (consistently contingent responding) and establishes a "communication superhighway" for the infant's learning, including learning about the self. Contingent responding, such as early mirroring by the parent of the child's affect, is exaggerated or manifestative because it is part of a process of "teaching" the child about his/her subjectivity. Learning about internal states depends on the caregiver's capacity to offer appropriate cues to indicate that the experience is generalizable. In other words, it is not episodic but semantic. The cues—the intonation of motherese, contingent responding, eye contact, the child being addressed by his/her own

name—activate the learning system, which in the domain of learning about one's subjective experience is most commonly transmitted within attachment relationships. The infant is biologically prepared to receive this information. The weakening of the epistemic trust in insecure attachment reduces the child's attention to the caregiver and limits opportunities for finding oneself in the other. The destruction of this trust in severe neglect and frank abuse destroys the opportunity to learn about oneself.

Some intriguing clinical implications follow. The mind is found within the other, not within itself, and thus it is obvious that evolution has in a sense "prepared" children's brains for psychological therapy. Children are eager to learn about the opaque mental world from those around them, but they are prepared to learn most readily about minds in conditions of epistemic trust. Therapists ignore the people to whom the individual naturally turns for knowledge at their peril. Therapy is not just about the *what* but also the *how* of learning. Indeed, opening the child's mind so that he/she can once again trust the social world by changing expectations will destroy the quality of a therapeutic relationship unless it is nested in ample signals of epistemic trust, contingent responding, respecting of identity, and appropriate non-verbal characteristics of the relationship.

Therapeutic process and action

Tronick's conceptual innovations have led him to develop a model of the therapeutic relationship that shares some elements with our mentalization-based view of therapeutic action. However, the two models also diverge in some important ways. For Tronick, the "patient and therapist co-create dyadic states of consciousness of mature minds" (Tronick, 2003, p. 486). There is thus a relation, albeit "very far from a strict relation" (Tronick, 2007, p. 14), between the MRM of infant–adult interactions and the mutual regulation that occurs in patient–therapist interactions. In particular, Tronick conceives of the therapeutic relationship as involving adult capacities and levels of maturity. This means that the therapeutic focus

should not just be on intentional states and affects—words, symbols, representations, and insight do count for adults—but also requires an affective, nonverbal background where ostensive cues of interpersonal proximity and creative co-consciousness may be more important than verbal content since, as in the process of establishing epistemic trust, the "reparation of messiness rather than synchrony might be a key change-inducing process in therapy" (Tronick, 2007, p. 14).

Tronick's model explains why the process of change is not necessarily an easy one (Fonagy & Target, 2002). "Change for a patient", Tronick (2007, p. 15) wrote, "means risking dissipation and experiencing fear or even the terror of annihilation and the dissolution of the self. But change also means hope". In order to facilitate change, a therapist must engage in what Tronick describes as a two-sided "scaffolding" of the patient. The therapist should help regulate the affect that threatens the patient's willingness to risk change at the same time as co-creating dyadic states of consciousness in which the patient and therapist can make new, more complex and coherent meanings. These meanings, Tronick stresses, are communicated at various levels and are received in various states. In alert, interpretative states, the patient is more likely to detect the meaning in cognitions. In order to apprehend meanings from preconscious or unconscious levels, something like reverie states (Ogden, 1997) may be needed. Yet other biopsychological levels could be required to reach meaning, requiring the use of something like Downing's (2003) body work. Traditionally, analysts have paid scant attention to the dyadic regulation of affect and tended to see the therapist's contribution as key and minimize the patient's receptivity as a determinant. Tronick's model suggests a far more "collaborative stance" in relation to the clinical situation.

We have found that our MM has similar implications for the treatment of patients, particularly those with BPD (Fonagy, 1998). Attention to the patient's level of arousal is a crucial part of our therapeutic technique, which recognizes that a high level of arousal is

inconsistent with balanced mentalizing (Bateman & Fonagy, 2004b). In crucial ways, however, BPD patients may be unlike the mature adults Tronick considers in his model (see also Allen, Fonagy, & Bateman, 2008; Bateman & Fonagy, 2004a; Bateman & Fonagy, 2006). Nevertheless, our formulation, like Tronick's, implies the need to abandon the overvaluation of specific techniques in favor of a generic therapeutic stance that cuts across theoretical modalities. The overarching expectation emerging from our model is that given the generic nature of mentalization as a mental function, most treatments of BPD will be effective to the extent that they include important components that facilitate mentalizing, even though this capacity may be addressed using different languages by various models of therapy, such as "mindfulness", "validation", "self-states", and so on. Similar conclusions would follow from the model of therapy that Tronick's work implies.

Both models suggest that the overall aim of treatment should be to stimulate a patient's attachment and involvement with treatment whilst helping them maintain some capacity for reflection (Tronick's two-sided scaffolding analogy). In the MM, therapists are also trained to be alert to the patient's level of arousal and to avoid situations where a patient may be expected to talk of mental states that he/she cannot link to subjectively felt reality. Thus, with regard to dynamic therapies of individuals with BPD, this often implies that there should be: (a) a de-emphasis of "deep" unconscious interpretations in favor of conscious or near-conscious content; (b) a modification of therapeutic aim, especially with severely disturbed patients, from insight to recovery of mentalization (i.e., achieving representational coherence and integration); (c) careful eschewing of descriptions of complex mental states (e.g., conflict, ambivalence, unconscious) that are incomprehensible to a person whose mentalizing is vulnerable; and (d) avoidance of extensive discussion of past trauma, except in the context of reflecting on current perceptions of mental states of maltreating figures and changes in mental state from being a victim in the past in contrast to one's experiences now.

In the MM we add specific reference to a desirable stance on the part of the therapist that is harder to locate within the MRM. The key task of mentalization-based treatment is to promote curiosity about the way mental states motivate and explain the actions of the self and others. Therapists can achieve this through judicious use of an "inquisitive stance", highlighting their own interest in the mental states that underpin behavior, qualifying their own understanding and inferences (showing respect for the opaqueness in mental states), and showing the patient how such information can help them to make sense of their experiences. It is not for the therapist to "tell" patients about how they feel, what they think, or how they should behave, or what the underlying conscious or unconscious reasons may be for their difficulties. Any therapeutic approach that moves towards claiming to "know" how patients "are", how they should behave and think, and "why they are the way they are" is likely to be harmful to patients whose capacity to mentalize is vulnerable.

The therapist's mentalizing therapeutic stance should include: (a) humility deriving from a sense of "not knowing"; (b) whenever possible taking time to identify differences in perspectives; (c) legitimizing and accepting different perspectives; (d) active questioning of the patient in relation to their experience, asking for detailed descriptions of experience ("what" questions) rather than explanations ("why" questions); and (e) eschewing the need to understand what makes no sense (i.e., saying explicitly that something is unclear). An important component of this stance is monitoring one's own mistakes as a therapist. Acknowledgement of mistakes models honesty and courage and tends to lower the patient's arousal because the therapist is taking responsibility. These moments also offer invaluable opportunities to explore how mistakes can arise out of inaccurate assumptions about opaque mental states and how misunderstandings can lead to upsetting experiences. In this context, it is important to be aware that the therapist is constantly at risk of losing his/her own capacity to mentalize in the face of a non-mentalizing patient. Consequently, we consider therapists'

occasional enactments as an acceptable concomitant of the therapeutic alliance, and something that simply has to be "owned up to". As with other instances of breaks in mentalizing, such incidents require that the process is "rewound" and the incident explored. Hence, in this collaborative patient–therapist relationship the two partners involved have a joint responsibility to understand enactments.

Conclusion

Tronick's important conceptualizations of the MRM and RAPs have brought a much-needed subtlety to areas of attachment theory. In particular, Tronick's model has contributed to an understanding of the unique personal and shared cultural contexts in which attachment relationships unfold. This includes the "messiness" of interpersonal communications as well as the greater cohesion allowed through reparation and co-creativity. Furthermore, Tronick's MRM allows us to operationalize concepts such as the "holding environment" (Winnicott, 1965) and "background of safety" (Sandler, 1960), as well as aspects of reflective understanding within a mentalization model of the parent–infant relationship. Over the past decades, Tronick has taken us far closer to a genuinely relational model of change in psychoanalytic treatment.

Many concepts from the MM, which we have worked on in parallel to Tronick, seem congruent with his insights. Indeed, like Tronick's MRM, mentalization is probably best considered as a complex multi-component capacity with a variety of determinants, some of which are genetic while others are more influenced by environmental facilitation and interference. Each of the correlates of secure attachment may interface with one or more of a range of neuropsychologically defined components of mentalizing. The MRM and MM can be considered as complementary to each other, with both emphasizing the dynamic interplay of state changes as determinants of interpersonal interaction, but while the MRM points to such co-regulation as the material of self-development and therapeutic change—its very

essence, if you will—the MM is less far advanced along this road and has thus far considered state changes only as moderators of the dynamic system of interaction between parent and child, and between patient and therapist. Tronick's contribution to our understanding of contingent interaction between these dyads has been an inspiration—not just to those working within the framework of mentalization, but to all of us who struggle to use the human relationship as the mediator of therapeutic change.

References

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the Strange Situation*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Allen, J. G., Fonagy, P., & Bateman, A. W. (2008). *Mentalizing in clinical practice*. Washington, DC: American Psychiatric Publishing.

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences, 14*, 110-118.

Bateman, A., & Fonagy, P. (2004a). *Psychotherapy for borderline personality disorder: Mentalization-based treatment*. Oxford, UK: Oxford University Press.

Bateman, A. W., & Fonagy, P. (2004b). Mentalization-based treatment of BPD. *Journal of Personality Disorders, 18*, 36-51.

Bateman, A. W., & Fonagy, P. (2006). *Mentalization based treatment for borderline personality disorder: A practical guide*. Oxford, UK: Oxford University Press.

Belsky, J., & Fearon, P. R. M. (2008). Precursors of attachment security. In J. Cassidy & P. R. Shaver (Eds.), *Handbook of attachment theory and research* (2nd ed., pp. 295-316). New York, NY: Guilford Press.

Blatt, S. J., & Behrends, R. S. (1987). Internalization, separation-individuation, and the nature of therapeutic action. *International Journal of Psycho-Analysis, 68*, 279-297.

Blatt, S. J., & Luyten, P. (2009). A structural-developmental psychodynamic approach to psychopathology: Two polarities of experience across the life span. *Development and Psychopathology, 21*, 793-814.

Bowlby, J. (1958). The nature of the child's tie to his mother. *International Journal of Psycho-Analysis, 39*, 350-373.

Bowlby, J. (1969). *Attachment and loss, Vol. 1: Attachment*. London, UK: Hogarth Press and Institute of Psycho-Analysis.

Bowlby, J. (1973). *Attachment and loss, Vol. 2: Separation: Anxiety and anger*. London, UK: Hogarth Press and Institute of Psycho-Analysis.

Bowlby, J. (1980). *Attachment and loss, Vol. 3: Loss: Sadness and depression*. London, UK: Hogarth Press and Institute of Psycho-Analysis.

Bretherton, K., & Munholland, K. A. (2008). Internal working models in attachment relationships: A construct revisited. In J. Cassidy & P. R. Shaver (Eds.), *Handbook of attachment: Theory, research and clinical applications* (2nd ed.). New York, NY: Guilford Press.

Colvert, E., Rutter, M., Beckett, C., Castle, J., Groothues, C., Hawkins, A., . . . Sonuga-Barke, E. J. (2008). Emotional difficulties in early adolescence following severe early deprivation: Findings from the English and Romanian adoptees study. *Development and Psychopathology, 20*, 547-567.

Colvert, E., Rutter, M., Kreppner, J., Beckett, C., Castle, J., Groothues, C., . . . Sonuga-Barke, E. J. (2008). Do theory of mind and executive function deficits underlie the adverse outcomes associated with profound early deprivation? Findings from the English and Romanian adoptees study. *Journal of Abnormal Child Psychology, 36*, 1057-1068.

Craik, K. (1943). *The nature of explanation*. Cambridge, UK: Cambridge University Press.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*, 148-

153.

Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 366*, 1149-1157.

Downing, G. (2003). Video-Mikroanalyse-Therapie: Einige Grundlagen und Prinzipien. In H. Scheuerer-Englisch, G. J. Suess & W.-K. Pfeifer (Eds.), *Wege zur Sicherheit: Bindungswissen in Diagnostik und Intervention* (pp. 51-68). Giessen, Germany: Psychosozial-Verlag.

Emde, R., Kubicek, L., & Oppenheim, D. (1997). Imaginative reality observed during early language development. *International Journal of Psycho-Analysis, 78*, 115-133.

Fonagy, P. (1989). On tolerating mental states: Theory of mind in borderline patients. *Bulletin of the Anna Freud Centre, 12*, 91-115.

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. (1998). An attachment theory approach to treatment of the difficult patient. *Bulletin of the Menninger Clinic, 62*, 147-169.

Fonagy, P., & Bateman, A. (2006). Progress in the treatment of borderline personality disorder. *British Journal of Psychiatry, 188*, 1-3.

Fonagy, P., Gergely, G, Jurist, E., & Target, M. (2002). *Affect regulation, mentalization and the development of the self*. New York, NY: Other Press.

Fonagy, P., Gergely, G., & Target, M. (2007). The parent–infant dyad and the construction of the subjective self. *Journal of Child Psychology and Psychiatry, 48*, 288-328.

Fonagy, P., & Luyten, P. (2009). A developmental, mentalization-based approach to the understanding and treatment of borderline personality disorder. *Development and*

Psychopathology, 21, 1355-1381.

Fonagy, P., Redfern, S., & Charman, T. (1997). The relationship between belief-desire reasoning and a projective measure of attachment security (SAT). *British Journal of Developmental Psychology, 15*, 51-61.

Fonagy, P., Steele, M., Moran, G. S., Steele, H., & Higgitt, A. C. (1991). Measuring the ghost in the nursery: A summary of the main findings of the Anna Freud Centre/University College London parent-child study. *Bulletin of the Anna Freud Centre, 14*, 115-131.

Fonagy, P., Steele, M., Steele, H., Moran, G. S., & Higgitt, A. C. (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, 12*, 201-218.

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. (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. (2007a). Playing with reality: IV. A theory of external reality rooted in intersubjectivity. *International Journal of Psychoanalysis, 88*, 917-937.

Fonagy, P., & Target, M. (2007b). The rooting of the mind in the body: New links between attachment theory and psychoanalytic thought *Journal of the American Psychoanalytic Association, 55*, 411-456.

Freeman, W. (1994). *Societies of brains*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Gergely, G., & Csibra, G. (2005). The social construction of the cultural mind: Imitative learning as a mechanism of human pedagogy. . *Interaction Studies, 6*, 463-481.

Green, A. (2000). Science and science fiction in infant research. In J. Sandler, A.-M. Sandler

& R. Davies (Eds.), *Clinical and observational psychoanalytic research: Roots of a controversy* (pp. 41-73). London, UK: Karnac Books.

Hesse, E. (2008). The Adult Attachment Interview: Protocol, method of analysis, and empirical studies. In J. Cassidy & P. R. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications* (Second ed., pp. 552-598). New York, NY: Guilford Press.

Higgitt, A., & Fonagy, P. (1992). The psychotherapeutic treatment of borderline and narcissistic personality disorder. *British Journal of Psychiatry, 161*, 23-43.

Hofer, M. A. (2003). The emerging neurobiology of attachment and separation: How parents shape their infant's brain and behavior. In S. W. Coates, J. L. Rosenthal & D. S. Schechter (Eds.), *September 11: Trauma and human bonds* (pp. 191-210). Hillsdale, NJ: Analytic Press.

Lorenz, K. (1935). Der Kumpan in der Umwelt des Vogels [Companionship in bird life]. In C. H. Schiller (Ed.), *Instinctive behavior: Development of a modern concept* (pp. 83-128). New York, NY: International Universities Press.

Luyten, P., Fonagy, P., Lowyck, B., & Vermote, R. (2012). Assessment of mentalization. In A. W. Bateman & P. Fonagy (Eds.), *Handbook of mentalizing in mental health practice* (pp. 43-66). Washington, DC: American Psychiatric Publishing.

Main, M., Hesse, E, & Goldwyn, R. . (2008). Studying differences in language usage in recounting attachment history: An introduction to the AAI. In H. Steele & M. Steele (Eds.), *Clinical applications of the Adult Attachment Interview*. New York, NY: Guilford Press.

Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology, 31*, 838-850.

Ogden, T. (1997). *Reverie and interpretation: Sensing something human*. Northvale, NJ:

Jason Aronson.

Pears, K. C., & Moses, L. J. (2003). Demographics, parenting, and theory of mind in preschool children. *Social Development, 12*, 1-20.

Sandler, J. (1960). The background of safety. *International Journal of Psycho-Analysis, 41*, 191-198.

Sharp, C., Fonagy, P., & Goodyer, I. M. (2006). Imagining your child's mind: Psychosocial adjustment and mothers' ability to predict their children's attributional response styles. *British Journal of Developmental Psychology, 24*, 197-214.

Sroufe, L. A. (1996). *Emotional development: The organization of emotional life in the early years*. New York, NY: Cambridge University Press.

Stolorow, R. D. (1997). Review of "A dynamic systems approach to the development of cognition and action". *International Journal of Psycho-Analysis, 78*, 620-623.

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.

Thompson, R. A. (2008). Early attachment and later development: Familiar questions, new answers. In J. Cassidy & P. R. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications* (2nd ed., pp. 348-365). New York, NY: Guilford Press.

Tronick, E. (2005). Why is connection with others so critical? The formation of dyadic states of consciousness and the expansion of individuals' states of consciousness: Coherence governed selection and the co-creation of meaning out of messy meaning making. In J. Nadel & D. Muir (Eds.), *Emotional development* (pp. 293-315). Oxford, UK: Oxford University Press.

Tronick, E. (2007). *The neurobehavioral and social-emotional development of infants and children*. New York, NY: W.W. Norton & Company.

Tronick, E. (2008). *Meaning making and the dyadic expansion of consciousness model*. Paper presented at the Festschrift in Honor of Arnold Modell. March 8, 2008, Boston, MA, Boston Psychoanalytic Society and Institute.

Tronick, E. Z. (1989). Emotions and emotional communication in infants. *American Psychologist, 44*, 112-119.

Tronick, E. Z. (2003). "Of course all relationships are unique": How co-creative processes generate unique mother-infant and patient-therapist relationships and change other relationships. *Psychoanalytic Inquiry, 23*, 473-491.

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.

Winnicott, D. W. (1965). *The maturational process and the facilitating environment*. London, UK: Hogarth Press.