

Fonagy, P; Bateman, AW; (2006) Mechanisms of change in mentalization-based treatment of BPD. *Journal of Clinical Psychology* , 62 (4) 411 - 430. [10.1002/jclp.20241](https://doi.org/10.1002/jclp.20241).

## RESEARCH ARTICLE

# Mechanisms of change in mentalisation-based therapy with

# BPD

*Peter Fonagy\* and Antony W Bateman\*\**

*\*University College London and The Anna Freud Centre, Sub-Department of Clinical Health Psychology, University College London, Gower Street, London WC1E 6BT, [p.fonagy@ucl.ac.uk](mailto:p.fonagy@ucl.ac.uk)*

*\*\*St. Ann's Hospital, North London, and University College London*

## **ABSTRACT**

There are very few less contentious issues than the assertion that attachment plays a role in psychotherapy. Concepts such as the therapeutic alliance speak directly to the importance of activating the attachment system, normally in relation to the therapist in individual therapy and in relation to other family members in family based intervention, if therapeutic progress is to be made. In group therapy the attachment process may be activated by group membership. The past decade of neuroscientific research has helped us to understand some key processes which attachment entails at brain level. The paper will outline this progress and link it to recent findings on the relationship between the neural systems underpinning attachment and other processes such as making social judgements, theory of mind, access to long term memory. These findings allow intriguing speculations currently undergoing empirical tests on the neural basis of individual differences in attachment as well as the nature of psychological disturbances associated with profound disturbances of the attachment system. Crucial to this paper, we will explore the paradoxical brain state created by psychotherapy with powerful clinical implications for the maximisation of therapeutic benefit from the talking cure.

## Overview

Borderline personality disorder (BPD) is a complex and serious mental disorder that is characterized by a pervasive pattern of difficulties with emotion regulation, impulse control, and instability both in relationships and in self-image with a mortality rate, associated with suicide, that is 50 times that of the general population (Skodol, Gunderson, Pfohl et al., 2002).<sup>1</sup> The dysfunction of self-regulation is particularly apparent in the context of social relationships (e.g. Posner et al., 2002).

The regulation of emotion and the catastrophic reaction to the loss of intensely emotionally invested social ties together place borderline personality disorder in the domain of attachment. A number of theorists have drawn on Bowlby's ideas in explanation of borderline pathology (Holmes, 2003). Gunderson (1996) carefully described typical patterns of borderline dysfunction in terms of exaggerated reactions of the insecurely attached infant, for example clinging, fearfulness about dependency needs, terror of abandonment, constant monitoring of the proximity of the caregiver. Lyons-Ruth (1999) focused on the disorganisation of the attachment system in infancy as predisposing to later borderline pathology. Crittenden (1997) has incorporated in her representation of adult attachment disorganisation the specific style of borderline individuals deeply ambivalent and fearful of close relationships. We (Fonagy, Target, & Gergely, 2000) have also used the framework of attachment theory but emphasise the role of attachment in the development of symbolic function and the way in which insecure disorganised

---

<sup>1</sup> A more fully referenced version of this paper is available from the authors on request.

attachment may generate vulnerability in the face of further turmoil and challenges.

Further support for the central role of attachment in the disorder comes from the evidence that psychotherapy is the most effective treatment modality although pharmacotherapy may enhance its effects. However the mechanism of change remains unknown. Given the strong suggestion of abnormal (disorganized) attachment processes and the consequent instability in emotions and relationships in BPD, we suggest that the mechanism mediating change is indeed via the improved regulation of neuropsychological systems underpinning the organization of interpersonal relationships (Bateman & Fonagy, 2004).

### **Mechanisms of change: causing change in causal mechanisms**

Discussing the therapeutic action of psychotherapy with BPD assumes that psychotherapy is indeed therapeutic. While evidence for this claim is gradually accumulating (Roth & Fonagy, 2005) we do not know how psychotherapy effects change and given the relatively rapid rate of spontaneous improvement (Zanarini, Frankenburg, Hennen, & Silk, 2003) in BPD, the observations of symptomatic improvements in uncontrolled studies should be considered with great caution. If we were to start from first principles, an understanding about the processes underpinning psychopathology would inform treatment innovation, which in turn would then be subject to empirical investigation. Kazdin (2004) has outlined a radical and rigorous programme for psychotherapy research based on these principles. The first stage follows from the proposition that treatment should reflect what we know about the processes that directly bear on the onset and course of a

clinical problem. In Kazdin's model, demonstrating that a specific process was present in a sizeable proportion of individuals with a specific presentation would be the basis of treatment development. Further, rather than assuming that all individuals with the same presentation would be equally responsive to treatment, further work would aim to detect subtypes of a dysfunction, multiple pathways to the same presentation, as well as risk and protective factors. The second step asks questions about the processes by which a treatment method achieves change firstly by specifying the processes or factors responsible for change, then developing measures of these processes and finally showing that these processes change before therapeutic change occurs. On this basis manualization becomes feasible, on the presumption that the manual now includes a high dosage of 'effective' ingredients. Evaluation of outcome can then follow, along with process-outcome studies that aim to examine moderator variables (helping us to discover more about what actually does work for whom).

Following Kazdin's proposal we shall first consider a comprehensive model of biological and psychosocial factors that come together in BPD. Each of the components and the evidence associated with dysfunction in the domains specified will be outlined. Secondly we will consider how the therapeutic technique we favour, mentalisation based treatment (MBT), links to components of the proposed mechanisms underlying the pathology of BPD. Finally we offer preliminary (inevitably somewhat speculative) ideas as to how psychotherapeutic intervention might be seen to address these dysfunctions. In so doing we consider psychotherapy more broadly than our own approach as it is inherently unlikely that specific mechanisms exist for each of the

therapeutic models apparently relatively successfully used in this clinical context.

### **BPD as a deficit in attachment related mentalisation**

Our view of BPD and attachment has undergone a number of changes as a result of accumulating data. Fifteen years ago we suggested that a better understanding of the features of borderline personality disorder could be achieved if we assumed that patients with this diagnosis had a limited capacity to mentalize, that is, to comprehend and use their knowledge of their own and others' states of mind (Fonagy, 1989). Our original theory of mentalisation dysfunction suggested that the apparent inability to process mental states effectively and appropriately in an attachment context was a defensive reaction to physical or sexual abuse leading to a de-coupling of mental processes underpinning thinking about feelings and thoughts in self and others (Fonagy, 1991). Later on we added that constitutional factors were also likely to be involved (Bateman & Fonagy, 2004).

Building on accumulating evidence from developmental psychopathology, the mentalisation theory of borderline personality disorder suggests that individuals either constitutionally vulnerable (Torgersen et al., 2000) and/or exposed to neglect in early relationships (Battle et al., 2004) where their emotional experience is not adequately mirrored by the caregiver (Crandell, Patrick, & Hobson, 2003), develop with an enfeebled or fragile capacity to represent affect and effortfully control their attentional capacities (Posner et al., 2002). These individuals when confronted with trauma (Battle et al., 2004) are more likely to react by decoupling their capacity to deal with their own or others' mental states comprehensively, particularly in an

attachment context (Fonagy et al., 1996). We may think of this as an adaptive decoupling, a deliberate avoidance of the state of mind of the perpetrator of maltreatment. The child cannot be expected to take fully on board what must be the frankly hostile intentional stance of the abuser. It may be easier not to think about mental states as a whole. There may be more to this than an adaptation. Early trauma may cause changes in the neural mechanisms of arousal which lead to a relatively ready triggering of the arousal system underpinning posterior cortical activation while taking the frontal mentalising parts of the brain 'offline' in response to relatively mild emotional stimuli (Arnsten, 1998).

Whatever the immediate cause of the decoupling, its consequence is the re-emergence of modes of thinking about internal states that antedate the fully-fledged mentalising capacity of the adult. In previous work we discussed three of these that we claim are relatively readily observable in typical patterns of thinking of individuals with BPD: the psychic equivalent, the pretend and the teleological mode of representing the internal world (Fonagy, Gergely, Jurist, & Target, 2002). All these are modes of representing subjectivity that developmentally antedate the emergence of full mentalization (Gopnik, 1993; Flavell & Miller, 1998) Children of two or three years of age tend to assume a direct correspondence between what is in their mind and what is physically true. A small child's unshakable belief that there is a tiger under the bed is an example of psychic equivalence. At other times children can contemplate an internal reality as long as no connection is made to the real world, as in early pretend play. While psychic equivalence makes subjective experience too real, the pretend mode severs its connection with

reality and at an extreme it is tantamount to dissociation. Finally the teleological mode refers to a unique dependence on what is observable in thinking about intentions. Often individuals with BPD seek proof in physical action as part of a full confirmation of a subjective state. For example one is being loved only if one is physically touched.

Within our theoretical frame of reference, dysfunctional attachment relationships are not just the consequence of the difficulty in holding a stable and consistent representation of others' and one's own mind in mind (Liotti, 2002), but also cause developmental distortions in self-organisation.. We have suggested a model that might explain this striking sequence.

Representations of emotions cannot emerge without interaction with another person who mirrors the infant's experience and whose reflections of this experience are internalised by the infant. When the caregiver's mirroring is not congruent, the infant organises internal experience by internalising the caregiver, rather than the caregiver mirroring the child (as suggested by Winnicott, 1956). Thus such second-order representations of internal states are by definition 'alien'. They do not match the constitutional state of the self. Consequently the self-organisation will evolve in a somewhat flawed manner.

A further factor, namely trauma, contributes to the foregrounding of self-fragmentation (identity diffusion) in BPD. We suggest that traumatised individuals can use discontinuities within the self, in other words the alien parts of the self, to adapt to incomprehensible assault from someone connected to them by attachment bonds by 'identifying with the aggressor' (Freud, 1936). The tragic consequence of this manner of self-protection is the modification of self-organisation so that the self now incorporates the abusive



intention. This generates momentary experiences of unbearable psychic pain when the self feels attacked literally from within and almost overwhelmed by an experience of 'badness' that is hardly mitigated by reassurance.

Experienced in the mode of psychic equivalence the feeling of badness translates directly into 'actual badness' from which self-destruction might appear the only escape. In our view, this state is commonly the trigger for acts of self-harm and suicide.

In summary we are proposing that a constitutionally vulnerable individual who experiences developmental trauma in an attachment context becomes psychologically vulnerable in later attachment contexts as a result of instability of the self. In an attempt to cope the individual decouples their mind from others minds and relies on earlier psychological mechanisms to organise the experience and in doing so reveals fragments of the self.

If our suggestions are correct the enhancement of mentalisation and the reduction of the predominance of non-mentalising modes of experiencing internal reality represent the path to cure.

### **The nature of the therapeutic intervention and the process of change**

While there have been several attempts to describe mentalisation focused therapy from a psychoanalytic standpoint for children (e.g. Bleiberg, Fonagy, & Target, 1997) and adults (e.g. Fonagy et al., 2002), the first comprehensive description of the therapeutic approach is to be found in Bateman and Fonagy's description of the partial hospital based treatment of BPD (see Bateman & Fonagy, 2004). There is limited but good quality evidence for the effectiveness of the approach (Bateman & Fonagy, 1999, 2001, 2003) with further studies under way. We believe that psychotherapy

has the potential to recreate an interactional matrix of attachments in which mentalization develops and flourishes. The therapist's mentalizing in a way that fosters the patient's mentalizing is seen as a critical facet of the therapeutic relationship and the essence of the mechanism of change. The crux of the value of psychotherapy is the experience of another human being having the patient's mind in mind.

The key features of this psychotherapeutic approach may be summarised as follows: (1) The therapist is asked exclusively to focus on patients' current mental state (their thoughts, feelings, wishes and desires) with the aim of building up representations of internal states, (2) They are asked to avoid situations where patient talks of mental states that they cannot link to subjectively felt reality. Thus there is deviation from traditional psychodynamic technique in that (a) there is a de-emphasis of 'deep' unconscious concerns in favour of conscious or near conscious content and less focus on past as it is represented in the present, (b) the aim of therapy is not insight but the recovery of mentalization: achieving representational coherence and integration for intentional states, (c) therapists avoid describing complex mental states (such as conflict, ambivalence, unconscious) and are asked to make "small interpretations" referring to ideation that is only slightly beyond the boundaries of the patient's conscious thinking; (3) In this way the therapy creates a transitional area of relatedness where thoughts and emotions can be "played with"; (4) The inevitable enactments over the course of the treatment are not interpreted or understood in terms of their unconscious meaning but in terms of the situation and affects immediately prior to the enactment. For a fuller discussion of the key features of MBT and

its similarities and differences from other therapeutic approaches see Bateman and Fonagy (2004).

## **The biological basis of therapeutic change in MBT**

### The role of neuroscience in psychological therapies

Returning to our original contention that change results from improvement in regulation of neuropsychological systems, we must consider the biological basis of the change caused by the therapeutic approach outlined above. Although BPD is one of the most investigated of the PDs, its neurobiological basis is relatively unknown even though research is rapidly identifying neural correlates of complex subjective states (Adolphs, 2003), for example concern about the mental states of another person (Frith & Frith, 2003). Natural recovery from BPD (Zanarini et al., 2003), sometimes dramatic (Gunderson et al., 2003) and substantial treatment related improvements (e.g. Linehan et al., 2002) suggest that the psychoneurobiology of BPD should relate to a reversible brain state rather than permanent abnormalities or alterations of function.

On the basis of the key dysfunctions of BPD, namely impulsivity and interpersonal problems, there are two regions of brain function which link together psychological processes of particular relevance to psychotherapy for BPD and biological processes: (a) reward circuits and attachment and (b) dysfunctions of interpersonal relatedness and mentalisation. We shall briefly consider knowledge in these areas in turn and consider the links between them.

## The biology of reward and attachment

Our suggestion is that the poor functioning of the reward system in individuals with BPD is linked to the dysfunction of the attachment system by their shared neural basis. Individuals with BPD make 'impulsive choices' that result in appetitive reward in the short term, but which have the potential to be self-damaging in the long run. Examples of this include gambling, irresponsible spending, binge eating, substance abuse, unsafe sex, and reckless driving. These impulsive choices all require the individual to assign greater value to immediate, short-term reward than to long-term rewards like safety and security. We may think of this kind of impulsivity as characterized by actions that are poorly conceived and prematurely expressed (Daruna & Barnes, 1993).

The simplest analogues of this type of impulsive behavior have been investigated within laboratory studies (Rogers et al., 1999). Individuals with BPD make poor estimations of likely outcomes in estimating the likelihood of monetary reward (Bazanis et al., 2002). The pattern of maladaptive performance was similar to individuals with focal lesions to the orbitofrontal cortex (Rogers et al., 1999), a region previously associated with reward anticipation and valuation. A further study showed BPD patients choose short-delay small rewards over longer-delay, larger rewards (Dougherty, Bjork, Huckabee, Moeller, & Swann, 1999). Recent animal work implicates the nucleus accumbens (Nac) in temporal bias toward short-term over long-term rewards (Cardinal, Pennicott, Sugathapala, Robbins, & Everitt, 2001). This, combined with the dopaminergic reward system of the Nac being implicated in an animal model of impulsivity suggest that the Nac specifically and the

mesencephalic dopaminergic reward system (MDRS) in general may be involved in bias toward short-term reward, and may be characteristic of the maladaptive valuation and decision-making of individuals with BPD.

The mesencephalic dopaminergic reward system has also been implicated in the process of drug addiction and the neurobiology of attachment. The pathway of interest includes the ventral tegmental area (VTA) that projects directly and indirectly via the amygdala/bed nucleus of the stria terminalis to the nucleus accumbens. From here projections are to the ventral pallidum and from there to the thalamus. The thalamic projections go to the prefrontal and cingulate cortex which are thought to activate cells that ultimately feedback to the VTA (Everitt & Wolf, 2002). It appears that, broadly speaking, drugs that lead to dopamine release in this system, such as psychostimulants, are addictive (Koob & Le Moal, 1997) although the neurochemical basis of addiction and substance use is probably more complicated than dopamine release (Insel, 2003). Given that it is unlikely that nature created a brain system specifically to serve drug and alcohol abuse, it seems most likely that addiction is the accidental by-product of the activation of a biological system which plays a crucial evolutionary function. MacLean (1990) speculated that substance abuse and drug addiction were attempts to replace opiates or endogenous factors normally provided by social attachments. Similarly, Panksepp (1998) suggested a common neurobiology to mother–infant, infant–mother, and male–female attachment relationships linked to the mesocorticolimbic dopaminergic reward circuit. More recently, Insel (2003) summarised relevant data that seems to answer in the affirmative the question: “Is social attachment an addictive disorder?”

There is good evidence, reviewed by Insel (2003) that mesocorticolimbic pathways also mediate mother–infant interactions in rats: (a) dopamine is released and c-Fos (a proto-oncogene that encodes a 55,000 mol. wt phosphoprotein, Fos, which is thought to assist in the regulation of "target genes" containing an AP-1 binding site) is activated in the nucleus accumbens when maternal females are exposed to their pups (Stack, Balakrishnan, Numan, & Numan, 2002); (b) lesions of the VTA and the nucleus accumbens will reduce the females' approach and interaction with pups (Lee, Li, Watchus, & Fleming, 1999); (c) cocaine or c-flupenthixol (a non-specific dopamine antagonist) injected directly into the nucleus accumbens causes a decrease in pup retrieval (Vernotica, Rosenblatt, & Morrell, 1999); (d) Insel and his group demonstrated that the neuropeptides oxytocin (OT) and vasopressin (AVP) are released by socio-sexual experience and may serve as an important link by which parturition, copulation, and lactation can activate this reward circuit leading to the suggestion that OT and AVP activity mark the generic reward states as specifically attachment-related (Young, Lim, Gingrich, & Insel, 2001).

There is further good evidence that the mesocorticolimbic pathways also mediate pair bonding in rodents (Insel & Young, 2000). Prairie and pine voles form partner preferences and pair bonds after mating but montane and meadow voles do not form selective attachments. Thus it is of substantial interest that in prairie but not montane voles dopamine release in Nac is associated with mating (Gingrich, Liu, Cascio, Wang, & Insel, 2000) and that D2 receptor agonists infused into the Nac induce and antagonists prevent partner preferences (Wang et al., 1999). These, and numerous other, studies

support the notion that mesolimbic dopamine activation of D2 receptors is necessary and sufficient for the development of a partner preference (in prairie voles). Voles research also suggests the hypothesis that mating releases OT and AVP which amplifies the dopamine signal in the Nac shell (Insel, 2003). To this we might add more recent evidence that the ventral pallidum, located within the ventral forebrain and part of the mesolimbic dopamine reward pathway, shows high density of vasopressin V1a receptors (V1aRs) in the monogamous prairie and pine voles, but not in the promiscuous meadow or montane voles (Lim, Murphy, & Young, 2004). Site-specific infusion of a selective V1aR antagonist into the ventral pallidum blocks pair bond formation in prairie voles (Lim & Young, 2004). Thus, V1aRs in the ventral forebrain appear to be crucial for pair bond formation, and this V1aR pattern seems to be correlated with monogamous social organization. Experimental evidence supporting this hypothesis comes from a single gene manipulation study (Lim, Wang et al., 2004). Partner preference formation (measured as time spent huddling with partner as opposed to stranger vole) in the socially promiscuous meadow vole could be increased by using viral vector V1aR gene transfer into the ventral forebrain. The overexpression of V1aR in the ventral pallidum leads to the development of partner preference in this interpersonally indiscriminating rodent. What we see here is a potential molecular mechanism for the rapid evolution of complex social behaviour and the possible recreation of a singular critical evolutionary event in the laboratory.

Knowledge from preclinical (animal) models is confirmed by neuroimaging studies which demonstrate an association between functional

brain activity related to attachment and cortical and subcortical sites in the human brain that contain a high density of the neurohormones OT and AVP. Imaging shows that the mesocorticolimbic dopaminergic pathway is activated while processing attachment related stimuli. Strathearn (2002) studied the brain activation in mothers while viewing pictures of their own infants and compared this to the pattern of activation when they looked at infants who were familiar but not their own in three affective states (crying, smiling and neutral expressions). He reported significant differences in the activation of the right globus pallidus/ventral striatum, the left putamen, the bed nucleus of the stria terminalis, the nucleus accumbens, the amygdala, bilateral hippocampi, and the fusiform face area (FFA). A very similar study (Nitschke et al., 2004) also reported activation in the orbitofrontal areas (OFC) in the mothers watching their own infants, which correlated with their hedonic mood state which was also measured in the study. OFC activation is of course also observed in other studies in response to pleasant experiences: taste, smell, money, winning, positive feedback, nicotine, cocaine (Kawabata & Zeki, 2004). The most compelling recent study comes from Bartels and Zeki (2004) who, using the contrast of own vs. other child controlled for age and familiarity, were able to demonstrate activity in the substantia nigra, globus pallidus, sub-thalamic nuclei, bed nucleus of the stria terminalis and the ventral tegmental area comprising almost all of the regions critical for the attachment-mediating neuropeptides in the human brain. As these workers had already reported an fMRI study of romantic love using a similar contrast design (Bartels & Zeki, 2000), they were able to confirm that most of the regions activated by maternal love were the same as those that they had



found to be associated with romantic love, i.e. those in the striatum (the putamen, globus pallidus, caudate nucleus), the middle insula and the dorsal part of the cingulate cortex.

In summary it has been established that mesocorticolimbic dopamine is an important candidate in the mediation of reward, the capacity for deferred gratification, and addiction, but is also critical for maternal behavior in rats and pair bonding in voles. A circuit linking a vasopressin sensitive mechanism within the anterior hypothalamus (MPOA) to the VTA and the nucleus accumbens shell may be especially important for mediating the rewarding properties of social interaction which are dysfunctional in BPD in line with their difficulties in forming normal attachments. The neuropeptides OT and AVP are released by socio-sexual experience in rodents and humans. When we activate this reward circuit a change in attachment behaviour follows (at least in voles). fMRI studies indicate activation of same pathways in response to stimuli relating to the participant's own infant and partner. We may conclude that a major neural system underpinning attachment has been identified.

#### Dysfunctional interpersonal relatedness and a deficit in mentalisation

Disturbed interpersonal relatedness has also been identified as a key aspect of BPD relative to other PDs (e.g. Skodol, Gunderson, McGlashan et al., 2002). This can refer to a range of difficulties including dramatic shifts from idealization of to disillusionment with others, frantic efforts to avoid perceived abandonment, and inappropriate interpersonal aggression. However, an emerging literature suggests that all of these may share common mechanistic and etiologic features. Specifically, disorganized attachment in the relationships of individuals with BPD may mediate the

expression of interpersonal problems among individuals. Individuals with BPD have a model of relationships characterized by insecurity of attachment relative to other groups of Axis I or Axis II patients (e.g. Barone, 2003). On numerous measures of adult attachment, patients with BPD are identified as insecure, preoccupied and fearful in their relationships (e.g. Patrick, Hobson, Castle, Howard, & Maughan, 1994). Patients with BPD have been characterized as having a specific type of disorganized, anxious-preoccupied attachment focused around an approach-avoidance dilemma where the attachment figure is simultaneously perceived as a source of threat and a secure base (Crandell et al., 2003). A variety of lines of research suggest that the disordered attachment that is characteristic of BPD results from psychosocial experiences of maltreatment (e.g. Trull, 2001) and premorbid temperamental attributes of negative affectivity and aggressive impulsivity (e.g. Silk, 2000). Surprisingly, while there is strong evidence for both the psychosocial mechanism (abnormal attachment) and psychosocial etiology (maltreatment) of interpersonal difficulties in BPD (e.g. Skodol, Siever et al., 2002), a neurobiological account of disrupted interpersonal interactions remains elusive. Explorations of the subscales of a self report measure of interpersonal relatedness identified two types of problems experienced by BPD patients: one group have difficulty in achieving closeness to others and another group feel extremely submissive, unable to state needs and avoiding conflict (Leihener et al., 2003). At least the problems of the first group may be related to a deficit of social cognition that makes accurate perception of the respective mental states of self and other and self-other differentiation difficult (Fonagy et al., 2000). Deficits of this aspect of interpersonal perception have

been demonstrated in analogue studies using film clips (e.g. Arntz & Veen, 2001), affect recognition and alexithymic symptoms (e.g. Sayar, Ebrinc, & Ak, 2001) and narratives of childhood experience (Fonagy et al., 1996).

A deficit of interpersonal awareness implies an underlying failure to distinguish clearly between one's own and others' mental states. Some of the brain abnormalities identified in borderline patients are consistent with the suggestion that a failure of representation of self-states is a key dysfunction. Some evidence suggests that the anterior cingulate cortex plays a key role in mentalizing the self, at least in the domain of emotional states (Frith & Frith, 2003). Lane has proposed more specifically that implicit self-representations (i.e., phenomenal self-awareness) can be localized to the dorsal anterior cingulate, whereas explicit self-representations (i.e., reflection) can be localized to the rostral anterior cingulate (Lane, 2000). Activation of the medial prefrontal cortex has been demonstrated in a series of neuroimaging studies in conjunction with a wide range of mentalization inferences, in both visual and verbal domains (Gallagher et al., 2000). It appears that the prefrontal cortex is involved when mentalizing interactively in a way that requires implicitly representing the mental states of others. The mesial prefrontal cortex, the parieto-temporal junction and the temporal poles constitute a network of areas that are invariably active when mentalizing activity is taking place (Gallagher & Frith, 2003). The same area of the brain is involved in other tasks which have been clinically described as challenging to patients with borderline problems, including assessing social trustworthiness (Winston, Strange, O'Doherty, & Dolan, 2002), interpreting the meaning of facial expressions (Critchley et al., 2000), making moral judgements (Greene &

Haidt, 2002) and tasks that entail attending to one's own emotions (Gusnard, Akbudak, Shulman, & Raichle, 2001). It has been argued that exposure to stress impairs prefrontal cortical function and the impairment may be catecholamine mediated (Arnsten, 1998). In line with this suggestion is the observation that N-acetyl-aspartate (NAA), a marker of neural integrity, is lowered in the anterior cingulate region of the medial prefrontal cortex of maltreated children and adolescents (De Bellis, Keshavan, Spencer, & Hall, 2000).

### The loose coupling of attachment and mentalisation

So far we have described two systems where borderline functional deficits are evident, certainly at the behavioural level and, on preliminary neurobiological evidence, at the biological level. First, the mesolimbic dopaminergic reward system underpinning attachment may be dysfunctional in BPD. This is suggested both by the poor ability of individuals with BPD to delay gratification and the evident disorganisation of their attachment system. The overlapping neural underpinnings of these deficits might suggest that therapeutically addressing either set of issues may indirectly benefit the other. Thus behavioural approaches, such as DBT, which begin by addressing the reward systems in general and impulsivity in particular, might, through such interventions also have an impact on the quality of functioning within the attachment system (although this was not part of the original aim of Dialectical Behavior Therapy). Equally the possibility of a common deficit underpinning these two aspects of BPD dysfunction might explain the beneficial effects that improvements in attachment relationships might generate for problems of impulsivity or decision making.

Second, the mesial prefrontal cortex, the parieto-temporal junction and the temporal poles are related to the deficits in mentalisation in interpersonal interactions. The suggestion here is that a focus on mentalisation addresses some of the social cognitive dysfunctions that generate inappropriate behaviour in BPD, particularly suspiciousness, aggression, insensitivity to social situations and adequate capacity to focus attention on self states to a degree pertinent to the interaction.

Attachment and mentalisation systems are probably not independent of each other but in subtle ways might be “loosely coupled”. This suggestion is based on the finding that securely attached children are relatively precocious in the development of mentalisation (Fonagy, Steele, Steele, & Holder, 1997; Meins, Fernyhough, Russel, & Clark-Carter, 1998) and the quality of parental mentalisation of a child facilitates the development of secure attachment (Fonagy, Steele, Moran, Steele, & Higgitt, 1991). More recently, the intersection of mentalisation capacity and social experience has been more generally noted (Carpendale & Lewis, 2004), although the specific nature of the interface remains controversial.

Increased understanding of the neural mechanisms underpinning attachment may provide a further important clue. As previously mentioned, Bartels and Zeki (2000, 2004), in two separate studies, reported that maternal and romantic attachment appeared systematically to suppress brain activity in regions associated with emotionally charged memories, negative emotions, and those associated with mentalising and social judgements. This suggests that strong emotional ties to an other (infant or partner) not only inhibit negative feelings but also impede the functioning of neural networks that

might assist in generating social judgements about the attachment figure. These ideas are extremely important and deserve consideration in some detail.

### The function of networks deactivated by the attachment system

While maternal and romantic love clearly serve different functions, they share a number of subjective and objective qualities such as preoccupation, deep concern, high level of commitment etc. Underpinning this may be a common set of brain mechanisms that are activated when attachment feelings are powerfully triggered but also deactivation of a characteristic set of other neural functions. 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 specialised 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 Maddock, 1999). In addition, studies of individuals with lesions in these areas suggest a role in judgements involving negative emotions (Adolphs, Damasio, Tranel, Cooper, & Damasio, 2000). It is possible that as projections from the affect oriented limbic/paralimbic regions modulate the activity of these areas, they could sub-serve the ways mood can inhibit or enhance cognitive processing (Mayberg et al., 1999). Further, these areas may play a role in recalling emotion-related material and generating

emotion-related imagery (Maddock, 1999) which may be relevant in relation to understanding the typology of attachment.

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, judgements of social trustworthiness, moral judgements, '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 et al., 2001). Making judgements that involve mental states has been shown to be associated with activation of the same system. Thus intuitive judgements of moral appropriateness (rather than moral reasoning) are linked (Greene & Haidt, 2002) as is assessment of social trustworthiness based on facial expression (Winston et al., 2002).

If confirmed by further studies, the pattern of activation of these three systems (the attachment system, system A and system B) has important implications for our understanding of the nature of individual differences in attachment, the relationship of attachment and mentalisation and consequently our understanding of dysfunctions associated with BPD and the mechanisms underpinning its psychological treatment. 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 (system A). This is to be expected since a key function of the attachment system is to moderate negative affects in the infant and presumably later in development (Sroufe, 1996). The overwhelming negative affect associated with the loss of attachment figures, the need for attachment figures at times of sadness, and the hedonic effect of “finding love” are obvious common observations in line with these findings. Not only is the loss of attachment likely to be aversive because of the loss of ‘reward’ (addiction) but also the prior inhibition of systems associated with the generation of negative affect is removed.

Equally consistent with expectations, is the suppression of social and moral judgements (system B) associated with the activation of the attachment system. Judgements of social trustworthiness and morality serve to distance us from others but become less relevant and may indeed interfere with our relationships with those to whom we are strongly attached (Belsky, 1999).

#### Some implications of the reciprocal activation of mentalisation and attachment

The apparently reciprocal relationship of mentalisation and attachment is puzzling as, at first sight, it appears to contradict some of the assumptions concerning the facilitative relationship between the two systems outlined above (viz. the assumption that mentalisation and attachment are positively correlated). Further scrutiny suggests a more complex relationship. First, the neural association between attachment and mentalisation confirms the link we have identified between the two systems at a behavioural level (Fonagy et al., 2002). Second, we have demonstrated how the parent’s capacity to mentalise in the context of an attachment relationship facilitates



the development of secure attachment in the infant (Fonagy et al., 1991). It is possible, taking a sociobiological perspective, that the parent's capacity to mentalise the infant or child serves to reduce the child's experienced need to monitor the parent for trustworthiness. The relaxation of the interpersonal barrier serves to facilitate the emergence of the attachment bond. Third, we have seen that mentalization emerges precociously in children who were securely attached in infancy (e.g. Fonagy et al., 1997). While at first sight this finding may seem inconsistent with the inverse relationship between attachment and mentalisation at brain level, if we consider the association developmentally, 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 accounts for the precocious development of mentalisation.

Fourth, we have consistently suggested that the capacity for mentalisation in the context of attachment was in some respects independent of the capacity to mentalise about interpersonal experiences independently of the attachment context (Fonagy et al., 2002). We have found that our specific measure of mentalisation in the attachment context, reflective function (Fonagy, Target, Steele, & Steele, 1998) is predictive of behavioural outcomes that other measures of mentalisation did 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 mentalisation in attachment contexts mediated outcome measured as the quality of adult romantic relationships but mentalisation measured independently of the attachment context using the Reading the

Mind in the Eyes test did not (Fonagy, Stein, Allen, & Fultz, 2003). It seems that measuring mentalisation in the context of attachment measures a unique aspect of social behaviour.

### Mentalisation in the context of attachment and the classification of attachment types

While, from an evolutionary perspective, mentalisation may be generally less relevant in an attachment context than in other social contexts, nevertheless it is quite likely that the ability to mentalize in the context of attachment relationships points to a highly desirable capacity. Individuals who are able to mentalize while thinking about romantic partners or infants are likely to manage these relationships better and may, for example, have less turbulent interpersonal relationships, or be perhaps particularly effective at times of inevitable conflict and argument. This could be why secure attachment is marked by a relatively good capacity to mentalise and generate coherent narratives of even turbulent interpersonal episodes (Main, 2000). The simple empirical prediction that follows from these speculations is that individuals who are able to retain a relatively high activation of the temporal lobes, the parieto-temporal junction together with the mesial prefrontal cortex (system B) in the presence of the activation of the dopaminergic mesolimbic pathways (attachment and reward system) will be those most likely to be classified as secure in their attachment.

### A new understanding of the deficit in mentalisation in BPD

We have arrived at a new understanding of the deficit of mentalisation entailed in borderline personality disorder. We have observed that

mentalisation in individuals with BPD frequently represented a challenge described (inappropriately) in the psychoanalytic literature as a failure of symbolisation or concreteness of thinking (e.g. Grotstein, 1983). Yet mentalisation in general does not appear to be a problem for most individuals with BPD; it is only in the context of intimate relationships that the patient's capacity to depict mental states in others accurately appears to falter (Fonagy et al., 2002).

One simple model that may account for this somewhat paradoxical observation is that the attachment system is "hypersensitive", triggered too readily, and consequently the reciprocally deactivated systems are often inefficient (relatively deactivated) in their functioning. Two of the core symptoms of BPD, frantic efforts to avoid real or imagined abandonment and the characteristic pattern of unstable and intense interpersonal relationships characterized by alternating between extremes of idealization and devaluation are perhaps directly linked to such hyperactivity or hypersensitivity (Gunderson, 2001). A number of frequently noted observations made about individuals with BPD are consistent with this simple assumption. (1) Relationships of BPD individuals have a rapidly escalating tempo moving from acquaintance to great intimacy far faster than one might expect. (2) Hyperactivity of the attachment system removes the system responsible for maintaining a normal emotional barrier between self and others and generates an impression of entangled and preoccupied relationships (Agrawal, Gunderson, Holmes, & Lyons-Ruth, 2004) and frequently, somewhat unwisely, removes the need to assess the social validity of the social partner. (3) The excessively positive character of the initial phase

relationships that individuals with BPD form (often labelled 'idealization') may reflect the suppression of negative relationship specific affects and the inability to integrate emotion and cognition. (4) Affective instability, particularly the characteristic intense brief episodes of dysphoria may be the result of some form of rebound phenomena related to the hypersensitive or hyperactive attachment system. Such rebound effects might also account for outbursts of violent anger and interpersonal suspiciousness (paranoia) which might reflect overactivity in system B (Pickup & Frith, 2001). (5) The reduction of the influence of affectively laden episodic memory may relate to the chronic feelings of emptiness often encountered by individuals with this diagnosis.

Of central concern to our theoretical and clinical propositions is the mentalisation deficit we have reported in these patients (Fonagy et al., 1996). While standing by our description of the mental processes that emerge as characteristic of the intentional stance of the individual with BPD, we would now like to specify the cause of this deficit. Previously, we have argued that the deficit was a self-induced one - a defensive reaction in vulnerable individuals when confronted with hostile states of mind in the context of interpersonal trauma. While much of the extant data still appears to us to fit with this model (e.g. the association of early neglect with BPD, the undermining of the development of symbolisation in the families of maltreated youngsters, the high prevalence of attachment trauma, the high prevalence of individuals who show no BPD symptoms following trauma), it seems highly likely that mentalisation deficit can be secondary to the abnormal functioning of the attachment system (i.e. its hyperactivity). The latter of course is likely to

be the consequence of developmentally early dysfunctions of the attachment system in combination with later traumatic experiences in an attachment context. We speculate that deficits of mentalisation might sometimes occur because the capacity to mentalise is taken off line by the activation of the attachment system. Of course, this is not an 'all or nothing affair'. We assume that in these individuals, who are insecurely attached and therefore would not be able to maintain mentalisation normally in any case in the context of attachment relationships, the hyper-responsiveness of the attachment system, perhaps related to traumatic or other early experiences or genetic predisposition, has an unusually negative impact upon mentalising with the expected effects already described above. It remains probable that activating the capacity to mentalize in the context of attachment relationships generates substantial anxiety for traumatised individuals which in turn increases the activation of the attachment system in even those with an avoidant-dismissing pattern of attachment. Of course here we have a potentially extremely vicious cycle of heightened attachment, increasingly decoupled mentalisation, and increased vulnerability to further interpersonal trauma. A number of testable hypotheses follow from this model. Predictions from this model include: (1) mentalisation dysfunctions should be observable only when the attachment system is active; (2) mentalisation dysfunction should be associated with negative affect; (3) problems of accurate social and moral judgements should correlate with mentalisation capacity; (4) the degree of disorganisation of attachment relationships should correlate with the likelihood of mentalisation problems; (5) there are likely to be deficits associated with the retrieval of emotion laden memories when the attachment

system is active; (6) ambiguous stimuli (e.g. polysemous words with attachment and non-attachment meanings) will be more likely to trigger the attachment system of BPD individuals.

### Implications for therapy and the mechanisms of change

How can this model of the nature of deficit in BPD serve to focus our work with borderline patients, or rather how do we understand the changes we observe given the current focus of our work which is to assist in the recovery of the capacity to accurately think about thoughts or feelings? There is an important proviso to this aim. MBT in a range of contexts attempts to enhance mentalisation but always in the context of an attachment relationship. Both in individual and in group therapy the therapist through a range of largely unconscious techniques activates the attachment system. This occurs through a) the discussion of current attachment relationships, b) the discussion of past attachment relationships, c) through creating a safe and sensitive interpersonal environment that assists with the patient's regulation of affect, the therapist encourages but regulates the patient's attachment bond to the therapist, d) in the context of group therapy the therapist attempts to engender attachment bonds between members of the group. At the same time, paradoxically, the therapist attempts to enhance mentalisation not just in the techniques defined in the therapy manual but perhaps more importantly and generically simply by taking an interest in the mental world of the patient. This creates what we now understand as a somewhat paradoxical situation in terms of brain activity insofar as psychological therapy simultaneously activates what may be two mutually inhibitory sets of systems. There are two other ways in which this somewhat paradoxical pattern of activation is

maintained: a) the titrated activation of negative emotions as the therapist encourages the confrontation of adverse or traumatic experiences and b) the encouragement to retrieve affect-laden episodic memories. In these ways the individual with BPD is encouraged to counteract the normal pattern of attachment related deactivation of mentalising negative emotions and social and moral judgements.

We can see that overall MBT encourages the patient not to relinquish mentalisation at the slightest suggestion of attachment related brain activation. This is likely to have an impact on the attachment system as well, since we have seen that mentalisation strengthens the development of secure patterns of attachment. In other words, at the same time as strengthening mentalisation we speculate that MBT moves the pattern of arousal within these systems closer to that characteristic of a secure attachment. Evidently, to achieve this the therapist must be careful to balance the intensity of attachment relationships and the complexity of mentalisation required of the patient. The technical recommendation in MBT is to focus the patient's mentalisation on relationships with relatively low levels of involvement and only gradually to focus the patient's thinking on relationships closer to the patient's core self. Similarly the tasks of mentalisation vary in demand characteristics with clarification at the most superficial end and exploring of repudiated intense emotions in relation to the attachment figure at the more complex end. The therapist's aim is to reduce the likelihood of an anxious catastrophic response to the introduction of needing to think about the states of mind of attachment figures.

Thus using this model may help us to differentiate it from alternative therapeutic packages. While such distinctions are never more than caricatures, nevertheless they do point to important differences in technique between therapeutic approaches which are probably broadly equivalent in terms of achieving substantial therapeutic change. Thus aspects of DBT clearly also focus on the enhancing of mentalisation, e.g. the encouragement of mindfulness. However, this rarely occurs in the context of requiring individuals to conceptualise mental states in the context of attachment relationships. By contrast, supportive psychotherapy is less specifically focused on mental states but probably attempts to reduce the hypersensitivity of the attachment system of an individual who suffered severe attachment trauma. Transference focused psychotherapy is perhaps closest to MBT in its orientation to mental states in attachment contexts, the difference being a far more limited emphasis on the titration of interventions. The observations reported by Levy, Clarkin and Kernberg (Levy & Clarkin, **in press**) are consistent with the propositions here: attachment related mentalisation (reflective function) improves only in TFP and not in DBT or supportive psychotherapy.

#### Implications for the general process of change in therapy

The therapeutic process we have described in relation to MBT with BPD is probably not specific to either this approach or to this patient group. A strong case can be made that all forms of psychotherapy take some advantage of the simultaneous activation of these normally mutually inhibitory systems. It is possible that psychotherapy in general works because it arouses the attachment system at the same time as applying interpersonal



demands (psychotherapy technique) which require the patient to mentalise, to confront and experience negative affect and confront and review issues of morality (superego). Why might this be helpful? We speculate that thinking about feelings, thoughts and beliefs in the context of attachment is helpful because in this 'paradoxical' brain state there may be more access to modifying preset ways of conceptualising the contents of one's own and other's minds, as well as issues of morality and social judgement. Activating the attachment system harnesses brain biology partially to remove the dominance of constraints on the present from the past (long-term memory) and creates the possibility of re-thinking, re-configuring intersubjective relationship networks. The specific advantage of MBT in this process may be its focus on the simultaneous activation of the attachment system whilst encouraging development of psychological processes that are normally inhibited as a result. To this extent MBT represents a confluence of biology and psychology and goes some way towards meeting Kazdin's edict that an understanding of the processes underpinning psychopathology should inform treatment innovation.

## References

- Adolphs, R. (2003). Cognitive neuroscience of human social behaviour. *Nature Reviews*, *4*, 165-178.
- Adolphs, R., Damasio, H., Tranel, D., Cooper, G., & Damasio, A. R. (2000). A role for somatosensory cortices in the visual recognition of emotion as revealed by three-dimensional lesion mapping. *J Neurosci*, *20*(7), 2683-2690.
- Agrawal, H. R., Gunderson, J., Holmes, B. M., & Lyons-Ruth, K. (2004). Attachment studies with borderline patients: a review. *Harv Rev Psychiatry*, *12*(2), 94-104.
- Arnsten, A. F. T. (1998). The biology of being frazzled. *Science*, *280*, 1711-1712.
- Arntz, A., & Veen, G. (2001). Evaluations of others by borderline patients. *J Nerv Ment Dis*, *189*(8), 513-521.
- Barone, L. (2003). Developmental protective and risk factors in borderline personality disorder: a study using the Adult Attachment Interview. *Attach Hum Dev*, *5*(1), 64-77.
- Bartels, A., & Zeki, S. (2000). The neural basis of romantic love. *Neuroreport*, *11*(17), 3829-3834.
- Bartels, A., & Zeki, S. (2004). The neural correlates of maternal and romantic love. *Neuroimage*, *21*(3), 1155-1166.
- Bateman, A. W., & Fonagy, P. (1999). The effectiveness of partial hospitalization in the treatment of borderline personality disorder - a randomised controlled trial. *American Journal of Psychiatry*, *156*, 1563-1569.

- Bateman, A. W., & Fonagy, P. (2001). Treatment of borderline personality disorder with psychoanalytically oriented partial hospitalization: an 18-month follow-up. *American Journal of Psychiatry*, *158*(1), 36-42.
- Bateman, A. W., & Fonagy, P. (2003). Health service utilization costs for borderline personality disorder patients treated with psychoanalytically oriented partial hospitalization versus general psychiatric care. *Am J Psychiatry*, *160*(1), 169-171.
- Bateman, A. W., & Fonagy, P. (2004). *Psychotherapy for Borderline Personality Disorder: Mentalization Based Treatment*. Oxford: Oxford University Press.
- Battle, C. L., Shea, M. T., Johnson, D. M., Yen, S., Zlotnick, C., Zanarini, M. C., et al. (2004). Childhood maltreatment associated with adult personality disorders: findings from the Collaborative Longitudinal Personality Disorders Study. *J Personal Disord*, *18*(2), 193-211.
- Bazanis, E., Rogers, R. D., Dowson, J. H., Taylor, P., Meux, C., Staley, C., et al. (2002). Neurocognitive deficits in decision-making and planning of patients with DSM-III-R borderline personality disorder. *Psychol Med*, *32*(8), 1395-1405.
- Belsky, J. (1999). 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.
- Bleiberg, E., Fonagy, P., & Target, M. (1997). Child psychoanalysis: Critical overview and a proposed reconsideration. *Psychiatric Clinics of North America*, *6*, 1-38.

- Cabeza, R., & Nyberg, L. (2000). Neural bases of learning and memory: functional neuroimaging evidence. *Curr Opin Neurol*, 13(4), 415-421.
- Cardinal, R. N., Pennicott, D. R., Sugathapala, C. L., Robbins, T. W., & Everitt, B. J. (2001). Impulsive choice induced in rats by lesions of the nucleus accumbens core. *Science*, 292(5526), 2499-2501.
- Carpendale, J. I. M., & Lewis, C. (2004). Constructing an understanding of mind: The development of children's social understanding within social interaction. *Behavioral and Brain Sciences*, 27, 79-151.
- Crandell, L. E., Patrick, M. P. H., & Hobson, R. P. (2003). 'Still-face' interactions between mothers with borderline personality disorder and their 2-month-old infants. *British Journal of Psychiatry*, 183, 239-247.
- Critchley, H., Daly, E., Phillips, M., Brammer, M., Bullmore, E., Williams, S., et al. (2000). Explicit and implicit neural mechanisms for processing of social information from facial expressions: a functional magnetic resonance imaging study. *Hum Brain Mapp*, 9(2), 93-105.
- Crittenden, P. M. (1997). Toward an integrative theory of trauma: A dynamic-maturation approach. In D. Cicchetti & S. L. Toth (Eds.), *Rochester Symposium on Developmental Psychopathology: Developmental Perspectives on Trauma* (Vol. 8, pp. 33-84). Rochester, NY: University of Rochester Press.
- Daruna, J. H., & Barnes, P. A. (1993). A neurodevelopmental view of impulsivity. In W. G. McCann, J. L. Johnson & M. B. Shire (Eds.), *The impulsive client: Theory, research and treatment* (pp. 76-89). Washington, DC: American Psychological Association.

- De Bellis, M. D., Keshavan, M. S., Spencer, S., & Hall, J. (2000). N-Acetylaspartate concentration in the anterior cingulate of maltreated children and adolescents with PTSD. *Am J Psychiatry*, *157*, 1175-1177.
- Dougherty, D. M., Bjork, J. M., Huckabee, H. C., Moeller, F. G., & Swann, A. C. (1999). Laboratory measures of aggression and impulsivity in women with borderline personality disorder. *Psychiatry Res*, *85*(3), 315-326.
- Everitt, B. J., & Wolf, M. E. (2002). Psychomotor stimulant addiction: a neural systems perspective. *J Neurosci*, *22*(9), 3312-3320.
- Flavell, J. H., & Miller, P. H. (1998). Social cognition. In W. Damon, D. Kuhn & R. S. Siegler (Eds.), *Handbook of child psychology, 5th ed.* (pp. 851–898). New York: Wiley.
- 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., 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, H., Steele, M., & Holder, J. (1997). Attachment and theory of mind: Overlapping constructs? *Association for Child Psychology and Psychiatry Occasional Papers, 14*, 31-40.
- Fonagy, P., Stein, H., Allen, J., & Fultz, J. (2003). *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., & 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., Steele, H., & Steele, M. (1998). *Reflective-Functioning Manual, version 5.0, for Application to Adult Attachment Interviews*. London: University College London.
- Freud, A. (1936). *The Ego and the Mechanisms of Defence*. New York: International Universities Press, 1946.
- 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 Cogn Sci, 7*(2), 77-83.

- Gallagher, H. L., Happe, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: an fMRI study of 'theory of mind' in verbal and nonverbal tasks. *Neuropsychologia*, 38(1), 11-21.
- Gingrich, B., Liu, Y., Cascio, C., Wang, Z., & Insel, T. R. (2000). Dopamine D2 receptors in the nucleus accumbens are important for social attachment in female prairie voles (*Microtus ochrogaster*). *Behav Neurosci*, 114(1), 173-183.
- 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 Cogn Sci*, 6(12), 517-523.
- Grotstein, J. (1983). A proposed revision of the psychoanalytic concept of primitive mental states, Part 2. The borderline syndrome - section 1. Disorders of autistic safety and symbiotic relatedness. *Contemporary Psychoanalysis*, 19, 570-604.
- Gunderson, J. G. (1996). The borderline patient's intolerance of aloneness: Insecure attachments and therapist availability. *American Journal of Psychiatry*, 153(6), 752-758.
- Gunderson, J. G. (2001). *Borderline Personality Disorder: A clinical Guide*. Washington DC: American Psychiatric Publishing.
- Gunderson, J. G., Bender, D., Sanislow, C., Yen, S., Rettew, J. B., Dolan-Sewell, R., et al. (2003). Plausibility and possible determinants of sudden "remissions" in borderline patients. *Psychiatry*, 66(2), 111-119.

- Gusnard, D. A., Akbudak, E., Shulman, G. L., & Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proc Natl Acad Sci U S A*, 98(7), 4259-4264.
- Holmes, J. (2003). Borderline personality disorder and the search for meaning: an attachment perspective. *Aust N Z J Psychiatry*, 37(5), 524-531.
- Insel, T. R. (2003). Is social attachment an addictive disorder? *Physiol Behav*, 79(3), 351-357.
- Insel, T. R., & Young, L. J. (2000). Neuropeptides and the evolution of social behavior. *Curr Opin Neurobiol*, 10(6), 784-789.
- Kawabata, H., & Zeki, S. (2004). Neural correlates of beauty. *J Neurophysiol*, 91(4), 1699-1705.
- Kazdin, A. E. (2004). Psychotherapy for Children and Adolescents. In M. Lambert (Ed.), *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change, 5th Edition* (pp. 543-589). New York: Wiley.
- Koob, G. F., & Le Moal, M. (1997). Drug abuse: hedonic homeostatic dysregulation. *Science*, 278(5335), 52-58.
- Lane, R. D. (2000). Neural correlates of conscious emotional experience. In R. D. Lane & L. Nadel (Eds.), *Cognitive neuroscience of emotion* (pp. 345-370). New York: Oxford University Press.
- Lee, A., Li, M., Watchus, J., & Fleming, A. S. (1999). Neuroanatomical basis of maternal memory in postpartum rats: selective role for the nucleus accumbens. *Behav Neurosci*, 113(3), 523-538.
- Leihener, F., Wagner, A., Haaf, B., Schmidt, C., Lieb, K., Stieglitz, R., et al. (2003). Subtype differentiation of patients with borderline personality



- disorder using a circumplex model of interpersonal behavior. *J Nerv Ment Dis*, 191(4), 248-254.
- Levy, K. N., & Clarkin, J. F. (in press). To examine the mechanisms of change in the treatment of borderline personality disorder. *Journal of Clinical Psychology*.
- Lim, M. M., Murphy, A. Z., & Young, L. J. (2004). Ventral striatopallidal oxytocin and vasopressin V1a receptors in the monogamous prairie vole (*Microtus ochrogaster*). *J Comp Neurol*, 468(4), 555-570.
- Lim, M. M., Wang, Z., Olazabal, D. E., Ren, X., Terwilliger, E. F., & Young, L. J. (2004). Enhanced partner preference in a promiscuous species by manipulating the expression of a single gene. *Nature*, 429(6993), 754-757.
- Lim, M. M., & Young, L. J. (2004). Vasopressin-dependent neural circuits underlying pair bond formation in the monogamous prairie vole. *Neuroscience*, 125(1), 35-45.
- Linehan, M. M., Dimeff, L. A., Reynolds, S. K., Comtois, K. A., Welch, S. S., Heagerty, P., et al. (2002). Dialectical behavior therapy versus comprehensive validation therapy plus 12-step for the treatment of opioid dependent women meeting criteria for borderline personality disorder. *Drug Alcohol Depend*, 67(1), 13-26.
- Liotti, G. (2002). The inner schema of borderline states and its correction during psychotherapy: a cognitive-evolutionary approach. *Journal of Cognitive Psychotherapy*, 16, 349-366.
- Lyons-Ruth, K., & Jacobovitz, D. (1999). Attachment disorganization: Unresolved loss, relational violence and lapses in behavioral and

- attentional strategies. In J. Cassidy & P. R. Shaver (Eds.), *Handbook of attachment theory and research* (pp. 520-554). New York: Guilford.
- MacLean, P. (1990). *The triune brain in evolution: role in paleocerebral functions*. New York: Plenum.
- Maddock, R. J. (1999). The retrosplenial cortex and emotion: new insights from functional neuroimaging of the human brain. *Trends Neurosci*, 22(7), 310-316.
- Main, M. (2000). The organized categories of infant, child and adult attachment: Flexible vs. inflexible attention under attachment-related stress. *Journal of the American Psychoanalytic Association*, 48(4), 1055-1096.
- Mayberg, H. S., Liotti, M., Brannan, S. K., McGinnis, S., Mahurin, R. K., Jerabek, P. A., et al. (1999). Reciprocal limbic-cortical function and negative mood: converging PET findings in depression and normal sadness. *Am J Psychiatry*, 156(5), 675-682.
- Meins, E., Fernyhough, C., Russel, J., & Clark-Carter, D. (1998). Security of attachment as a predictor of symbolic and mentalising abilities: a longitudinal study. *Social Development*, 7, 1-24.
- Nitschke, J. B., Nelson, E. E., Rusch, B. D., Fox, A. S., Oakes, T. R., & Davidson, R. J. (2004). Orbitofrontal cortex tracks positive mood in mothers viewing pictures of their newborn infants. *Neuroimage*, 21(2), 583-592.
- Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. Oxford: Oxford University Press.

- Patrick, M., Hobson, R. P., Castle, D., Howard, R., & Maughan, B. (1994). Personality disorder and the mental representation of early social experience. *Developmental Psychopathology, 6*, 375-388.
- Pickup, G. J., & Frith, C. D. (2001). Theory of mind impairments in schizophrenia: symptomatology, severity and specificity. *Psychol Med, 31*(2), 207-220.
- Posner, M. I., Rothbart, M. K., Vizueta, N., Levy, K. N., Evans, D. E., Thomas, K. M., et al. (2002). Attentional mechanisms of borderline personality disorder. *Proc Natl Acad Sci U S A, 99*(25), 16366-16370.
- Rogers, R. D., Everitt, B. J., Baldacchino, A., Blackshaw, A. J., Swainson, R., Wynne, K., et al. (1999). Dissociable deficits in the decision-making cognition of chronic amphetamine abusers, opiate abusers, patients with focal damage to prefrontal cortex, and tryptophan-depleted normal volunteers: evidence for monoaminergic mechanisms. *Neuropsychopharmacology, 20*(4), 322-339.
- Roth, A., & Fonagy, P. (2005). *What works for whom? A critical review of psychotherapy research. Second edition.* New York: Guilford Press.
- Sayar, K., Ebrinc, S., & Ak, I. (2001). Alexithymia in patients with antisocial personality disorder in a military hospital setting. *Isr J Psychiatry Relat Sci, 38*(2), 81-87.
- Silk, K. R. (2000). Borderline personality disorder. Overview of biologic factors. *Psychiatr Clin North Am, 23*(1), 61-75.
- Skodol, A. E., Gunderson, J. G., McGlashan, T. H., Dyck, I. R., Stout, R. L., Bender, D. S., et al. (2002). Functional impairment in patients with

- schizotypal, borderline, avoidant, or obsessive-compulsive personality disorder. *Am J Psychiatry*, 159(2), 276-283.
- Skodol, A. E., Gunderson, J. G., Pfohl, B., Widiger, T. A., Livesley, W. J., & Siever, L. J. (2002). The borderline diagnosis I: psychopathology, comorbidity and personality and personality structure. *Biological Psychiatry*, 51(12), 936-950.
- Skodol, A. E., Siever, L. J., Livesley, W. J., Gunderson, J. G., Pfohl, B., & Widiger, T. A. (2002). The borderline diagnosis II: biology, genetics, and clinical course. *Biol Psychiatry*, 51(12), 951-963.
- Sroufe, L. A. (1996). *Emotional development: The organization of emotional life in the early years*. New York: Cambridge University Press.
- Stack, E. C., Balakrishnan, R., Numan, M. J., & Numan, M. (2002). A functional neuroanatomical investigation of the role of the medial preoptic area in neural circuits regulating maternal behavior. *Behav Brain Res*, 131(1-2), 17-36.
- Strathearn, L., & McClure, S. M. (2002). *A functional MRI study of maternal responses of infant facial cues*. Paper presented at the Annual Scientific Meeting of the Society for Neuroscience. 2002. Abstract Viewer/Itinerary Planner, Program No. 517.5. Online., Washington, DC.
- Torgersen, S., Lygren, S., Oien, P. A., Skre, I., Onstad, S., Edvardsen, J., et al. (2000). A twin study of personality disorders. *Compr Psychiatry*, 41(6), 416-425.

- Trull, T. J. (2001). Structural relations between borderline personality disorder features and putative etiological correlates. *J Abnorm Psychol*, *110*(3), 471-481.
- Vernotica, E. M., Rosenblatt, J. S., & Morrell, J. I. (1999). Microinfusion of cocaine into the medial preoptic area or nucleus accumbens transiently impairs maternal behavior in the rat. *Behav Neurosci*, *113*(2), 377-390.
- Wang, Z., Yu, G., Cascio, C., Liu, Y., Gingrich, B., & Insel, T. R. (1999). Dopamine D2 receptor-mediated regulation of partner preferences in female prairie voles (*Microtus ochrogaster*): a mechanism for pair bonding? *Behav Neurosci*, *113*(3), 602-611.
- 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.
- Winston, J. S., Strange, B. A., O'Doherty, J., & Dolan, R. J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nat Neurosci*, *5*(3), 277-283.
- Young, L. J., Lim, M. M., Gingrich, B., & Insel, T. R. (2001). Cellular mechanisms of social attachment. *Horm Behav*, *40*(2), 133-138.
- Zanarini, M. C., Frankenburg, F. R., Hennen, J., & Silk, K. R. (2003). The longitudinal course of borderline psychopathology: 6-year prospective follow-up of the phenomenology of borderline personality disorder. *Am J Psychiatry*, *160*(2), 274-283.