Attachment, mentalization and borderline personality disorder

Abstract

In this paper we aim to show how psychoanalytic and attachment theory ideas can usefully be integrated to address the clinical problems encountered in the treatment of borderline personality disorder. The mentalization-based theory and treatment of borderline personality disorder summarised here has been discussed at greater length in a number of recent publications (Bateman & Fonagy, 2004, 2006; Fonagy & Bateman, 2006a; Fonagy, Target, Gergely, Allen, & Bateman, 2003).

Individuals with borderline personality disorder have schematic, rigid and sometimes extreme views, which make them vulnerable to powerful emotional storms and apparently impulsive actions, and which can create profound problems of behavioural regulation, including affect regulation. We consider that the classic symptoms of unstable interpersonal relationships, affect dysregulation, impulsive acts of violence, suicide, self-harm, and transient psychotic symptoms can best be understood in terms of a failure of mentalization and consequent re-emergence of more primitive modes of subjectivity. We consider the failure of mentalization within the attachment context to be the core pathology of BPD (Bateman & Fonagy, 2004) and our treatment package aims to assist in its recovery (Bateman & Fonagy, 2006).

Mentalization and attachment

Mentalizing is a form of imaginative mental activity that involves perceiving and interpreting human behaviour in terms of potential mental states, such as needs, desires, feelings, beliefs, goals, purposes and reasons. When we are talking or interacting with another person, we don’t really know what they are thinking or feeling; we have to imagine it.

How do we acquire this capacity? How do we learn what’s on other people’s minds? We believe that the capacity for mentalization is acquired in the context of the primary attachment relationship between the attachment figure and the infant. The infant’s expressions and
behaviour indicate to the caregiver what is on its mind, and the caregiver reflects this back to
the infant: she creates a representation in her mind of the infant’s mental state, which she then
mirrors back to the infant through her facial expressions, words, tones and gestures. The infant
then internalizes this representation to form the core of her or his psychological self.

In this way, the infant begins to acquire an understanding of his or her feelings. But in this
process, it is crucial that the caregiver metabolizes the affects that she mirrors, that is, gives
them back to the infant in a more manageable form which the infant is then able to internalize
as a representation of how she or he feels. This metabolizing process is important because
simply mirroring back what the infant feels is more likely to be upsetting than soothing; it will
make it seem as though the feeling is spreading into others rather than being contained. The
caregiver needs to ‘mark’ the affect, to show the infant that it is not her own feeling that she
is expressing, but rather something to do with the infant. This may be achieved, for example,
by combining different affects: perhaps combining the mirroring of sadness with a trace of
irony. This ‘marked’ representation is then bound to the constitutional state; this is symbolic
binding, which we often talk about as a second order representation of the internal state. These
interactions allow the infant to integrate earlier, more primitive modes of representing subjec-
tivity.

How is subjectivity experienced before the capacity for mentalization emerges? One of the key
concepts here is psychic equivalence. For a toddler, the mind and the world outside are
isomorphic. What exists in the mind is felt to exist out there and what exists out there must
also exist in the mind. Imagine reassuring a two-and-a-half year old child who is afraid that
there is a tiger under the bed, “look, there’s no tiger under the bed.” The child replies, “but
when you switch the light off there is a tiger under the bed.” This is psychic equivalence. He
believes that there is a tiger, therefore there is a tiger.

The acquisition of a sense that there is a difference between mental states and objective reality
is therefore essential. The child needs to be able to decouple his feelings from physical reality.
At first when this decoupling takes place it is complete, and the child enters what we have
called ‘the pretend mode’. When a three-year-old child plays that a chair is a tank, at that
moment he is in a little world of his own that cannot be touched from the outside. When the
first author asked his three year old son while he played that a chair was a tank, “is that a chair,
or is that a tank?” he looked up at him, and stopped playing. The game was spoiled. He could
pretend, he could have an imaginary world, only as long as it was decoupled from external
reality.

Gradually, by engaging in playful interaction with a concerned adult who both helps the child
to represent and thus manage his overwhelming feelings, and takes his pretend world
seriously, the pretend and psychically equivalent modes are integrated to form genuine subjectivity. However, if the caregiver fails to perform this crucial mirroring function, the child will internalise a representation of the caregiver in lieu of a self-representation. In place of the self will be an alien other that nevertheless feels like the self. The result of this discontinuity or splitting within the self is often that the incoherent, alien part of the self is externalized onto the attachment figure. Aged four or five, a child with this kind of disorganized attachment history may use coercive, controlling behaviour to elicit from the caregiver the aspects of themselves that they find intolerable within, for example, an angry or frustrated part of the self. This is also what Bion, we believe, called normal projective identification. Given that no mother can be 100% attentive to us, we all have alien parts within our self-structure. But once we have acquired the capacity to mentalize, this creates a necessary illusion of continuity within the self. Under normal circumstances our sense of self is relatively continuous and we are unaware of the presence of breaks and fissures within our self structure.

More serious problems start when there is a trauma, particularly when there is trauma in the attachment context. The alien self, the non-contingent part of the self, becomes host to the traumatizing figure through an internalization, which is carried out in order to achieve control over that traumatizing figure. This leads to the very well known clinical observation in working with traumatized individuals that they experience themselves as evil and hateful, because they have internalized the torturing figure in order to achieve some kind of control over him or her. This creates a self-harm state, where the attack from within is often concretized in attacks against the body, and then three things happen.

First, there is a repudiation of mentalization in an attachment context. Often it feels better for an abused child not to think about the adult’s mind than to try to conceive of why they are trying to harm them. But as soon as we stop thinking about our mental states, we remove our illusory sense of self-coherence. This leads to a feeling of being fragmented from within, and at that point the externalization of the alien parts of the self becomes a matter of life or death. If the self within is a torturing figure intent on destroying us, we urgently need to get it outside of ourselves. We then find, through externalization, a container, who will actually also end up torturing us, but from the outside rather than from within. So to the self comes to feel hated and attacked, and at the same time is linked to this container, the person who is able to contain those feelings, through a kind of addictive attachment bond. This is pathological projective identification.

Second, when mentalization ceases, thinking regresses to the modes of subjectivity that antedate the development of mentalization. Psychic equivalence is manifest in the experience of traumatic flashbacks, and a fantasy, even if it is projected, becomes terrifyingly real. In patients with a history of trauma there may be an intolerance of alternative perspectives, so
what I think is going on, is what’s going on: “You looked at your watch, you must be bored with me!” “No, I’m not bored. I just wanted to know how much time we had together.” “You were bored, and I’m not staying here, and I’m not putting up with this any more!” Even more painfully, self-related cognitions, the individual’s ways of seeing him- or herself, are experienced as having all the truth of physical reality.

The re-emergence of the pretend mode can be detected in the sense of emptiness that can follow trauma, a meaninglessness of internal states. In our borderline patients, at the extreme, this is a dissociative state: when they are in this state we can talk to them about feelings, but none of that discourse means anything. And the risk of this, of course, is that we then carry on talking to them for years and years and years without them getting anywhere. We think we are talking together, but they are not experiencing any meaning.

A third complication is introduced by the fact that whilst mentalization has its roots in the sense of being understood by an attachment figure, it is also more challenging to sustain mentalization in the context of an attachment relationship (GUNDERSON, 1996). Recent intriguing neuroscientific findings have highlighted how the activation of the attachment system tends temporarily to inhibit or decouple the normal adult’s capacity to mentalize. 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 as well as regions associated with mentalizing and social judgements. This suggests that strong emotional ties to an other (infant or partner) inhibit not just negative feelings but impede the functioning of neural networks that might assist in generating social judgements about the attachment figure. (BARTELS & ZEKI, 2004).

At first glance this apparently reciprocal relationship between mentalization and attachment may appear to contradict some of the to assumptions made above concerning the facilitative relationship between the two systems. There are a number of complexities to this negative association. First, the neural association between attachment and mentalization confirms the link we have identified between the two systems at a behavioural level (FONAGY, GERGELY, JURIST, & TARGET, 2002). Second, we have demonstrated how the capacity to mentalize in the context of an attachment relationship on the part of the parent creates the potential for secure attachment to develop in the infant (FONAGY, STEELE, MORAN, STEELE, & HIGGITT, 1991). It is possible, taking a sociobiological perpective, that the parent’s capacity to mentalize the infant or child serves to reduce the child’s experienced need to monitor the parent for trustworthiness. The relaxation of the interpersonal barrier serves to facilitate the emergence of the attachment bond. Third, we have seen that theory of mind emerges precociously in children securely attached in infancy (e.g. MEINS, 1997). While at first sight this finding may appear inconsistent with the inverse relationship between attachment and mentalization at brain level, if we think...
about the association developmentally, it is to be expected that over time in individuals whose attachment is secure, there are likely to be fewer calls for the activation of the attachment system. This in turn accounts for the precocious development of mentalization.

Fourth, we have consistently suggested that the capacity for mentalization in the context of attachment was in some respects independent from the capacity to mentalize about interpersonal experiences independent of the attachment context (Fonagy & Target, 1997). We have found that our specific measure of mentalization in the attachment context, reflective function (Fonagy, Target, Steele, & Steele, 1998), predicts behavioural outcomes that other measures of mentalization 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 mentalization in attachment contexts mediated outcome measured as the quality of adult romantic relationships. However, this was not the case for mentalization measured independently of the attachment context using the Reading the Mind in the Eyes test (Fonagy, Stein, Allen, & Fultz, 2003). It seems that measuring mentalization in the context of attachment captures a unique aspect of social behaviour.

The failure of mentalization in borderline personality disorder
In our model of the failure of mentalization in BPD, the role of the attachment environment is considered alongside constitutional vulnerabilities. The vulnerability reflected in the hereditability of BPD (TorgerSEN, 2000), may be directly linked to the capacity for mentalization or may represent the fragility of this capacity in situations of environmental deficiency as exemplified by severe neglect, psychological or physical abuse, childhood molestation or other forms of maltreatment.

As we considered above, mentalization may be temporarily inhibited by intense emotional arousal, by the intensification of attachment needs or by a defensive turning away from the world of hostile and malevolent minds in the context of severe maltreatment. Our current model stresses that minor experiences of loss or relatively small emotional upsets without expectation of comforting may be enough to cause the intense activation of the attachment system in these individuals. Their attachment system is hyper-activated, probably due to interpersonal experiences associated with childhood trauma. This state of arousal inhibits mentalization and combined with an unstable capacity for affect regulation triggers the typical symptoms of the disorder.
The re-emergence of prementalistic representation of internal states

We assume that the absence of fully functioning mentalization becomes evident through the re-emergence of the prementalistic modes of representing subjectivity discussed above. The clearest of these is the tendency to assume that mental states are direct representations of psychical reality. Mentalization gives way to a kind of ‘psychic equivalence’, (TARGET & FONAGY, 1996) which clinicians often consider under the heading of ‘concreteness of thought’. The individual with BPD, at these times, has an overriding sense of certainty in relation to their subjective experience. There is a suspension of the experience of ‘as if’. Everything, sometimes frighteningly, appears to be ‘for real’. The exaggerated reactions of patients are justified by the seriousness with which they suddenly experience their own and others’ thoughts and feelings. The vivid and bizarre quality of subjective experience can appear as ‘quasipsychotic’ symptoms (ZANARINI, GUNDERSON, & FRANKENBURG, 1990) and is reminiscent of the physically compelling memories associated with PTSD (MORRISON, FRAME, & LARKIN, 2003).

Conversely, thoughts and feelings can come to be dissociated to the point of near meaninglessness. Dissociation is a component of one of the diagnostic criteria of BPD; and is identified as a long-term outcome of disorganized attachment (LIOTTI, 2006; LYONS RUTH & SPIELMAN, 2004). It recalls the young child’s ‘pretend mode’ of thought. The young child can create and maintain pretend worlds only so long as these achieve complete separateness from the world of physical reality (GOPNIK, 1993). In these states patients may discuss experiences without contextualising these in any physical or material reality. Because mental representations are not constrained by reality, excessively complex representations of mental states in self and others are often elaborated in the pretend mode. Achieving such complexity of course occurs at the cost of a loss of balance and accuracy. Several studies using Rorschach, TAT and other narrative methods have provided evidence of idiosyncratic, hypercomplex representations of mental states of others which are often seen as malevolent (STUART ET AL., 1990; WESTEN, LOHR, SILK, GOLD, & KERBER, 1990; WESTEN, LUDOLPH, LERNER, RUFFINS, & WISS, 1990). Therapy in the “pretend mode” may appear to illuminate past or current relationships but will have no impact on the patient’s life.

The failure of mentalization reveals the disorganization of the self

The most socially disruptive feature of BPD is the apparently unstoppable tendency to create unacceptable experiences within the other. Our model links this tendency to the disorganization of the self that might be the consequence of disorganized attachment. In a nutshell, we assume that the legacy of disorganised attachment, a disorganized self structure, is brought into relief by the inhibition of mentalization at moments of intense emotion. An unstable sense of self is commonly noted in clinical descriptions of BPD (see e.g. JANIS, VEAGUE, & DRIVER-LINN, 2006; PARKER, BOLDERO, & BELL, 2006).
We have speculated that when the child is not given the opportunity to develop a representation of his own experience through mirroring (the self), he internalises the image of the caregiver as part of his self-representation (WINNICOTT, 1956). As we mentioned earlier we have called this discontinuity within the self ‘the alien self’. The caregiver’s threatening or dissociated mental states when experienced within the self may become so unbearable that the child’s attachment behaviour becomes focused on re-externalising these parts of the self onto attachment figures. The child becomes preoccupied by the need to control others’ actions.

Externalisation of the split-off parts of a disorganised self is desirable for the child with a disorganised attachment but may be a matter of life and death for a traumatised individual who has internalised the abuser as part of the self. We assume that to achieve control over malevolence a strategy available to those with disorganized self structures is to internalize the hostile state of mind into the alien part of the self where it becomes an abusive internal representation. Patients may view themselves as permanently damaged, or rotten to the core (ZITTEL CONKLIN & WESTEN, 2005). The extreme levels of dysphoric affect reported by individuals with BPD may be a reflection of this sense of being tormented from within (BRADLEY & WESTEN, 2005; ZANARINI, FRANKENBURG, HENNEN, & SILK, 2003).

Externalizations of these internal states are widely recognized in the common counter-transference reactions of therapists working with borderline patients – anger, helplessness, fear, worry, resentment, and urges to rescue the patient (GABBARD & WILKINSON, 1994). A fascinating study from Drew Westen’s laboratory showed that BPD patients triggered countertransference reactions in even senior clinicians and that these responses were most often associated with feeling criticised, mistreated and overwhelmed/disorganized (BETAN, HEIM, ZITTEL CONKLIN, & WESTEN, 2005). Complications arise when therapists respond to such externalisations by emotionally distancing themselves from individuals with BPD (AVIRAM, BRODSKY, & STANLEY, 2006). As AVIRAM ET AL. (2006) point out, the stigmatizing attitude – engendered through the externalisation of an alien part of the self – may exacerbate the behaviours in the patient that originally created the stigmatizing attitude. The result is often a self-fulfilling prophecy and a cycle of stigmatization to which both patient and therapist contribute.

**Hyperactivation of attachment**

The alternative to projective identification is obtaining relief from experiences of overwhelming and intolerable emotion through self-harm and suicide (KULLGREN, 1988; YEN ET AL., 2002). These and other actions can also serve to create a terrified alien self in the other – therapist, friend, parent – who thus becomes the vehicle for what is emotionally unbearable. The need for this other who “uniquely understands” (and thereby suffers) the patient’s dysregulated
affect not surprisingly, can become overwhelming and an adhesive, addictive pseudo-attachment to this individual may develop.

We see the capacity to mentalize as particularly helpful when people have been traumatized. Mentalization of adversity experiences can moderate their negative sequelae (Fonagy et al.). The capacity to mentalize enables those who are subjected to traumatic experiences to hold back modes of primitive mental functioning. It makes conceptual sense, therefore, for mentalizing to be a focus for therapeutic intervention if we are to help borderline patients bring primitive modes of mental functioning under better regulation and control.

**Mentalizing as the key to successful treatment**

Traditional psychotherapeutic approaches depend for their effectiveness on the individual’s capacity to consider their experience of their own mental state alongside its re-presentation by the psychotherapist. The integration of one’s current experience of mind with the alternative view presented by the psychotherapist is key to a change process. However, patients with BPD who at times have a very poor appreciation of their own and others’ perception of mind are unlikely to benefit from traditional psychological therapies. The weaker their sense of their own subjectivity, the harder they will find it to compare their own perceptions of their mind with their therapist’s. They might respond by accepting uncritically or rejecting totally what is being suggested. Neither response is likely to be helpful. It often leads to bewilderment, which in turn leads to instability as the patient attempts to integrate the different views and experiences. Unsurprisingly, this may result in more rather than less mental and behavioural disturbance.

With BPD patients the problem is compounded by the fact that, as we have seen, attachment and mentalization are loosely coupled systems existing in a state of partial exclusivity (Bartels & Zeki, 2004; Fonagy & Bateman, 2006b). Although the capacity for mentalization is rooted in the attachment figure’s marked mirroring, it is also true that mentalization is more challenging to maintain in the context of an attachment relationship (e.g. the relationship with the therapist) for individuals with dysfunctional attachment systems (Gunderson, 1996; Levy, 2005; Levy, Meehan, Weber, Reynoso, & Clarkin, 2005). Yet unless they are required to mentalize in the context of attachment relationships, BPD patients will never develop a capacity to adequately function psychologically in the context of intimate interpersonal relationships. Therapy will only be effective to the extent that it is able to enhance the patient’s mentalizing capacities without generating too many negative iatrogenic effects as it stimulates the attachment system. Following this principle we developed a treatment focusing on mentalizing and subjected it to research scrutiny. The treatment is lengthy (18 months) and involves both individual and group therapy. The understanding of behaviour in terms of underlying mental states forms a common thread running across both aspects of treatment. The focus of therapy
is on the patient’s moment-to-moment state of mind. Patient and therapist collaboratively try to generate alternative perspectives to the patient’s subjective experience of themselves and others by moving from validating and supportive interventions to exploring the therapy relationship itself as it suggests the possibility of alternative understandings of what is going on between the participants. Whilst the latter aspect makes it a psychodynamic treatment, using a psychoanalytic model of the mind, the aim is not to achieve ‘insight’, but rather to recover the capacity for mentalization. To achieve this particular attention is paid to the generic stance of the therapist rather than to specific types of intervention. The therapist’s stance, defined as a ‘not knowing or mentalizing stance’ is more interactive than traditionally used in long term psychodynamic treatments. The emphasis is on exploration and on elaborating a multi-faceted representation based on current experience particularly with the therapist. So, validation of patient experience moves gradually towards exploration in the current therapeutic relationship but first the therapist must demonstrate his understanding of the patient’s experience as real and justified. Only once that is established can alternative perspectives be brought in to the dialogue. Even then, in keeping with the ‘not-knowing’ stance of the therapist, this process is understood as impressionistic and the therapist’s contribution is considered as having no more or less validity than that of the patient – together they should arrive at an understanding but it is likely to be the therapist who teases out an alternative perspective. Once an alternative perspective about an interaction is identified the therapist must monitor not only his own reaction but that of the patient. The joint reaction then becomes the focus of the session and so the process moves on. It is especially important to note again that the aim of this process is not to increase insight and understanding, for example about the contribution of the past to the present, but rather to repair a current break in the self-structure and to facilitate mentalizing within the context of an emotional interaction when the attachment relationship is activated. The process of therapy becomes more important than the content.

The treatment was manualized and tested in two RCTs (BATEMAN & FONAGY, 1999, 2001) and long term follow-up indicates that positive changes identified at the end of treatment show further improvement over a period of 5 years after all treatment has ended (BATEMAN & FONAGY, 2008).

There is further evidence consistent with the MBT model. A recent trial of transference focused psychotherapy, supportive therapy and dialectical behaviour therapy for BPD has shown significant symptomatic improvements in all three groups (CLARKIN, LEVY, LENZENWEBER, & KERNBERG, 2007). There were more significant symptomatic changes in the TFP group than in either of the other therapies. The findings from the AAIs administered at the beginning and end of this 12 month treatment show major change in reflective function indicating a substantial shift in the patient’s capacity to mentalize the thoughts, feelings, intentions, and desires of self and others (LEVY ET AL., 2006). Notably, in this study only transference focused therapy was...
associated with a statistically significant improvement in mentalization (pre-post effect size $d=.98$).

**Conclusion**

Mentalization-based treatment is rooted in attachment theory and is consistent with many findings from attachment research, while its clinical technique leans heavily on psychoanalytic practice, including a relatively free therapeutic discourse about current events, particularly in an interpersonal attachment context. The focus of the approach, however, is provided by attachment theory inspired developmental research into the growth of understanding of mental states in self and other.

We have advanced a model of the development of social cognition that pays adequate regard to the role of attachment relationships as the primary teaching context within which understanding of minds in self and other occurs. This approach may also be helpful in explaining why individuals in a situation of psychological distress, when experiences within their own mind are confusing and anxiety-provoking, tend to seek an understanding person with whom they recreate the pedagogic configuration of one mind teaching another mind about aspects of subjectivity. We have tried to show how the dual components of attachment and interpersonal understanding are the key ingredients of recovering an equilibrium, a kind of recalibration of understanding of mental phenomena. Psychoanalysis, as the discipline which has made the intensive study of human subjectivity its specific concern, and attachment theory, which has made the dyadic human relationship its particular focus, are thus inseparably linked in providing a model for psychological therapy.

Attachment theory cannot and does not aspire to specify the full richness of subjective contents that preoccupy the ordinary mind, let alone the mind in distress. This is the ambition of psychoanalysis. Attachment theory, however, offers the best theoretical and research framework for understanding and elaborating the interpersonal process within which the human mind is best able to explore the subtleties of subjectivity. Attachment theory provides an understanding of the frame as well as some key facets of the contents; psychoanalysis guides, systematizes and inspires the mind’s understanding of mind. It is the achievement of such understandings, regardless of specific contents, that brings about change. In seeking to describe the relationship between attachment theory and psychoanalysis, we could say that the former offers a narrative framework whilst the latter provides the thickness of description required to create a complete story. Attempting to reduce one to the other will inevitably generate partial accounts and paradoxes: neither is complete without the other.
References


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

Anthony Bateman MA FRCPsych
Consultant Psychiatrist in Psychotherapy, Halliwick Unit, St. Ann’s Hospital, Barnet, Enfield, and Haringey Mental Health Trust London; Visiting Professor, University College, London

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