Translation: Mentalizing as treatment target in Borderline Personality Disorder

Peter Fonagy

Research Department of Clinical, Educational and Health Psychology, University College

London, UK

Patrick Luyten

Faculty of Psychology and Educational Sciences, University of Leuven, Belgium

Research Department of Clinical, Educational and Health Psychology, University College

London, UK

Anthony Bateman

Halliwick Personality Disorder Service, St Ann's Hospital, London

Correspondence concerning this chapter should be addressed to Peter Fonagy,

Research Department of Clinical, Educational and Health Psychology, University College

London, Gower Street, London WC1E 6BT, UK. E-mail: p.fonagy@ucl.ac.uk

Translation: Mentalizing as Treatment Target in Borderline Personality Disorder

This paper focuses on the clinical application of mentalizing ideas to the treatment of BPD. Over the past decades a number of evidence based treatment approaches for this severe condition have been developed, MBT being one of these (Stoffers et al., 2012). Here we examine the role of mentalizing in relation to BPD with a view to achieving improved levels of mentalizing in BPD patients as a therapeutic target. We seek to explain why mentalizing works as a treatment target for BPD, and argue that a mentalizing approach to BPD is at the core of any successful intervention. Recent developments in our understanding of mentalizing continue to influence approaches to treatment and this paper presents an update of the general theoretical and clinical approach, and specific treatment interventions and principles.

Mentalizing is the ability to understand others in terms of their thoughts, feelings, wishes, and desires; it is a very human capability that underpins everyday interactions. Without mentalizing there can be no robust sense of self, no constructive social interaction, no mutuality in relationships and no sense of personal security (Fonagy, Gergely, Jurist, & Target, 2002). Mentalizing is therefore a fundamental psychological process and so interfaces with all major mental disorders and has generic applicability in psychiatric care (Allen, Bleiberg, & Haslam-Hopwood, 2003; Choi-Kain & Gunderson, 2008).

Although mentalizing techniques are now applied to a wide range of psychological disorders (Bateman & Fonagy, 2012), it is for borderline personality disorder (BPD) that mentalization based treatment (MBT) was first developed and for which it has the most substantial evidence base (Bateman & Fonagy, 2003; Bateman & Fonagy, 2004; Bateman & Tyrer, 2004). The first section of this paper is an introduction to mentalizing in the context of BPD: it gives an overview of the development of the components that comprise full mentalization and how these are typically manifest in BPD patients, with a view to

appreciating how this contributes to our understanding of treatment targets for BPD. The second section explores the theory of social learning and epistemic trust in order to explore systems for therapeutic change in BPD and how this links to the importance of mentalizing as a theoretically cross-cutting factor across any successful therapeutic intervention. The third section focuses on translation of theory into practice in the consulting room: it goes into further detail about how therapy can and should address the typical mentalizing features and profiles of BPD patients overviewed in the first section, and factors contributing to the failure to develop full mentalizing as explored in the second section. The implications of the content covered in the previous sections are discussed in terms of the principles, structure, protocols and technique of MBT, and illustrated with clinical examples.

Attachment, mentalizing, and BPD

There has been much fruitful discussion of the role of mentalizing in the origins of BPD and the ways in which infants and children learn about managing relationships and emotional states (e.g. Fonagy & Bateman, 2008). Mentalizing skills are acquired in the context of early attachment relationships, which have typically been found to be preoccupied and disorganized in BPD patients (Choi-Kain, Fitzmaurice, Zanarini, Laverdiere, & Gunderson, 2009; Levy, Beeney, & Temes, 2011). Yet, attachment problems alone cannot explain the typical clinical picture of BPD. Problems in affect regulation, attentional control, and self-control stemming from dysfunctional attachment relationships (Aaronson, Bender, Skodol, & Gunderson, 2006; Agrawal, Gunderson, Holmes, & Lyons-Ruth, 2004; Barone, Fossati, & Guiducci, 2011; Lyons-Ruth, Yellin, Melnick, & Atwood, 2005; Scott et al., 2013; Sroufe, Egeland, Carlson, & Collins, 2005) are thought to be mediated through a failure to develop a robust mentalizing capacity (Fonagy & Bateman, 2008). In line with these assumptions, we see the defining characteristics of BPD – emotional dysregulation, impulsivity, interpersonal dysfunction – as rooted in an instability of the reflective, regulatory

capacities that mentalizing affords. From a wider mentalizing perspective, mental disorders in general can be seen as the mind misinterpreting its own experience of itself and of others (Bateman & Fonagy, 2010).

The potential for mentalizing seems to be an innate human characteristic (Kovacs, Teglas, & Endress, 2010) but the full acquisition of this ability is a developmental achievement, one that is likely to be highly responsive to environmental influences (Fonagy, Target, Gergely, Allen, & Bateman, 2003). Mentalizing capacities seem dependent on the quality of the early social learning environment and early attachment experiences; specifically, the attachment figures' ability to respond with contingent and marked affective displays in response to the infant's subjective experience. Mentalizing is thus seen as a fundamentally bi-directional or transactional social process (Fonagy & Target, 1997): it is influenced by the capacity of our attachment figures to mentalize, but their capacity to mentalize is also influenced by the child's characteristics (e.g., temperament), reflecting so-called evocative person-environment interactions (Fonagy, 2003; P. Fonagy, Luyten, & Strathearn, 2011).

The bi-directional nature of mentalizing is clearly fundamental to understanding its developmental origins, and has considerable bearing on how we conceptualize mental disorders in relation to their characteristic interpersonal difficulties. It is also central to how we formulate the treatment and treatment targets for psychotherapy in BPD. For individuals who have not benefitted from a stable and secure early environment in which they experienced consistent validation of their thoughts and feelings (Linehan, 1993), an effective therapeutic environment will be a theater for an introduction to mentalizing skills. The objective of reaching a state of improved mentalizing in the patient is reached through an interactional process whereby the therapist models their own mentalizing capacities and demonstrates their ability to mentalize the patient. In other words, mentalizing as an end

target is achieved through the experience of being effectively mentalized; it is an implicitly processed experience as well as a target of treatment.

The Multiple Mentalizing Competencies

We have conceptualized mentalizing as comprised of four attributes or polarities, each with two opposing values, or poles (Luyten & Fonagy, this issue). We assume that the limitations in mentalizing shown by BPD patients are explained by lack of balance across these polarities (see Figure 1). The aim of MBT is to address this imbalance in a manner that is closely attentive to moment-to-moment changes in current functioning. Consideration of the polarities of mentalizing is helpful in locating the loci where therapists should work (and probably normally find themselves working whether they are doing MBT or not) to restore an equilibrium into mentalizing capacity. Here, we consider how a BPD patient typically presents in relation to each polarity, and the resulting implications for therapy.

The automatic-controlled polarity: Controlled mentalizing is conscious, verbal, and reflective; automatic mentalizing is nonconscious, nonverbal, and unreflective. Individuals with BPD often tend toward this latter form of mentalizing. BPD patients are particularly likely to fall back on automatic mentalizing at moments of intense emotional arousal (e.g., in attachment contexts, challenging interpersonal situations, feelings of shame, guilt, anger, or inadequacy), often with severe impairments in social cognition as a consequence (e.g., being overly distrustful or trustful, being idealizing or overly denigrating).

The dramatically reduced capacity of BPD patients for reflective functioning (Fonagy et al. (1996)) that is necessary for controlled mentalizing has been shown to be reversible by psychotherapy (Levy et al., 2006). The task of the therapist, then, is to help slow down the patient's thinking and move them to a more reflective, explicit form of mentalizing which requires more conscious, verbal attention; all without generating a process of pseudo reflection or hypermentalizing (Sharp et al., 2013; Sharp et al., 2011).

The external–internal polarity: BPD patients tend to interpret the mental states of others on the basis of exterior cues (physical and visible features such as gestures or actions, or even their own actions) at the expense of internal ones. Indeed, the heightened sensitivity to emotional experiences in BPD seems to be intrinsically related to an increased sensitivity to external features of self and others as a source of knowledge about mental states (Fonagy & Luyten, in press).

The hypersensitivity of BPD patients to others' emotions (Gunderson & Lyons-Ruth, 2008), including those of the therapist, means that they often fail to develop plausible scenarios concerning others' states of mind based on these feelings, and are unable or unwilling to consider alternative explanations. Interventions that target mentalizing often need to start by examining interpretations based on external features and then generate possible plausible scenarios about internal states of mind, particularly the subtleties and complexities of people's internal worlds. The task of the therapist is to attempt to move the patient's focus to mental interiors: the thoughts, feelings, history and experiences that might offer further insight into someone's behavior. This shift involves not only slowing patients down and encouraging them to recognize their assumptions (e.g. relying on controlled or explicit mentalizing), but also to start to consider other people's subjective experiences using inference alongside mental state judgments based on appearance.

The affective-cognitive polarity: Full mentalizing involves the integration of *cognitive* belief-states (dominated by a cognitive awareness of how other people's attitudes and behavior are shaped by their own mental states and beliefs) and *affective* knowledge (dominated by inferences drawn from one's own feelings, self-affective state propositions). BPD patients often over-rely on the logic of emotion in preference to the logic of cognition when assessing subjective states, and indiscriminatingly apply the very particular logic of emotion to wider thoughts and beliefs (Fonagy & Luyten, 2009). These individuals thus will

show a bias toward attributing their own self-states to others (Baron-Cohen, Golan, Chakrabarti, & Belmonte, 2008; Baron-Cohen, Tager-Flusberg, & Cohen, 2000). At extremes, mental reality and physical/outer reality become equated and the patient's genuine sense of conviction generates a so-called "concrete understanding" of psychological experience and an intolerance of alternative perspectives.

The therapeutic target would be to encourage the patient to step back from this collapsing of appearance and reality, drawing attention to the patient's vulnerability to emotional contagion and their difficulty in inserting doubt or uncertainty into a process of inference. The BPD patient will frequently experience moments in which they are overwhelmed by affect; the response of the therapist should be to help the patient integrate this intense sense of affective knowledge or awareness (whether about the self or others) with more reflective and cognitive knowledge.

The self–other polarity: The capacity to differentiate between the self and other is often severely impaired in BPD patients (e.g., Barnow, Ruge, Spitzer, & Freyberger, 2005; Bender & Skodol, 2007; Blatt & Auerbach, 1988; Fuchs, 2007; Kernberg, 1984). From a mentalizing perspective, proneness to self-other confusion is explained by these patients' frequent inability to inhibit their own reactions when they are thinking about the mind of someone else. This means that the shared world and individual minds are not clearly demarcated for them; they fully expect others to know what they are thinking and feeling and to see situations in the same way they do. This may also explain the common experience of therapists who find that BPD patients are determined to control the thoughts and feelings of those around them. Underpinning their apparent need for control may be an inability to inhibit the impingement of others' mental states (Rigoni, Brass, Roger, Vidal, & Sartori, 2013; Spengler, von Cramon, & Brass, 2010).

When a patient starts to talk about their own state in a non-mentalizing (hypermentalizing or overly concrete) way, the strategy adopted in MBT is to shift the patient's attention to the mental state of others. This "contrary move", one of the typical interventions in MBT, aims to surprise the patient and force him/her to mentalize the states of others. Similarly, when the patient becomes intensely concerned with someone else's mental state, it may be helpful for them to consider how this preoccupation is affecting their own functioning. Working to shift the patient's thoughts and feelings between the self and the other in a consistent, flexibly responsive way targets the patient's tendency to become rigidly stuck in an unproductive mode of mentalizing by enabling him/her to be cognizant of the differences between himself/herself and others.

In working across the four mentalizing polarities, the basic and consistent aim is to re-establish mentalizing when it is lost – the typical starting point of mentalization-based interventions. This is clearest when a balanced form of mentalizing is no longer obvious. The task is not to seek structural or personality change in the patient with BPD, or explicitly aim to alter their representations, cognitions or schemas. Rather, the emphasis is on improving the patient's capacity for mentalizing and to make their mentalizing skills more stable and robust so that they are more able to manage affect and think about problems, particularly within interpersonal relationships.

Prementalizing Modes

Modes of prementalizing tend to re-emerge whenever we lose the ability to mentalize, as typically happens in individuals with BPD, particularly in high arousal contexts (Fonagy & Target, 1997). We now understand the emergence of these non-mentalizing modes as indications of imbalance in mentalizing, where it comes to be dominated by one end of a mentalizing polarity, presumably as a function of weakening of the other pole.

Conceptualizing this in terms of characteristic modes of subjectivity provides a further way

of appreciating the phenomenology of individuals with BPD features and thus provides the

basis for a helpful formulation for clinical work.

In the *psychic equivalence mode*, thoughts and feelings become "too real", and there

are no conceivable alternative perspectives. In children under five years old, this inability to

separate thought from reality is omnipresent (Gopnik, 1993). There is a suspension of doubt;

the individual increasingly believes that their own perspective is the only one possible. As we

have seen, this state reflects the domination of the propositional logic of emotion states

(where reality is defined by self-experience) over the propositional logic of cognition (where

agent:attitude propositions permit knowledge that belief is independent of reality; Baron-

Cohen et al., 2008). Psychic equivalence makes all subjective experience excessively real.

The overwhelming mental pain reported by BPD patients (Zanarini & Frankenburg, 2007) is,

in our view, rooted in this aspect of mentalizing dysfunction.

In the *teleological mode,* there is only recognition of real, observable goal-directed

behavior and objectively discernible events. The individual can recognize the existence and

potential role of mental states, but this recognition is limited to very concrete, observable

goals. Congruent with this position, the individual cannot accept anything other than a

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

reflects an extreme exterior focus in the mentalizing profile. The action proneness of

individuals with BPD is often, rightly or wrongly, considered under the heading of

impulsivity (Lawrence, Allen, & Chanen, 2010; Sebastian, Jacob, Lieb, & Tuscher, 2013).

From a mentalizing perspective, impulsivity is associated with the externally focused

subjectivity of the individual with BPD. If physical and visible features are prioritized,

others' or one's own actions also become an inevitable focus. If a focus on "mental interiors

– on thoughts, feelings and experiences – is inaccessible, then subjective reality may be

possible to create only by making an impact on the physical world, whether that outer reality

is that of the social world (creating an impact through getting others to act) or the more constrained world of one's own body (Lemma, 2012).

In *pretend mode,* thoughts and feelings become severed from reality (which we term "hypermentalizing" or "pseudomentalizing"). Again, a mode of pretending, where reality is suspended and internal states are all that matter, is universal in young children, yet such states of mind are readily punctured when reality (or an adult) intrudes into the game. Retreating to this mode in adulthood is associated with a sense of meaninglessness which may be defensive but is no longer experienced as pleasurable. It is linked to an experience of emptiness, which may result in the individual becoming dependent on strong guiding voices that seem to promise a sense of meaning. These features of BPD are neglected in research and theoretical conceptualizations of this condition – quite inappropriately, in our opinion, given their prevalence and significance (Luyten, Fonagy, Lowyck, & Vermote, 2012; Sharp et al., 2011). In these states, ideas form no bridge between inner and outer reality; the mental world is no longer fully coupled with external reality. To make things feel real, a person in desperation may turn to self-injury or other dramatic acts to break out of the feeling of emptiness. The subjective experience of meaninglessness is a manifestation of what happens when explicit mentalizing is overridden by implicit mentalizing, so that there is excessive internal focus unchecked by reflection. A common consequence is hypermentalizing, where groundless inferences are made about mental states, sometimes reminiscent of confabulation. The ultimate meaninglessness of hypermentalizing is assured by other failures in the mentalizing process including poor belief-desire reasoning, a vulnerability to fusion with others' identity, and a tendency to become lost in the complexity of the world of beliefs and desires with which physical reality is only loosely coupled.

These three prementalizing modes are particularly important as they are often accompanied by a more acute experience of disorganization within the experience of the self.

Most modern psychology assumes that self-coherence (the sense that one has continuity and consistency in thought and behavior) is somewhat an illusion (Bargh, 2011, 2014). It is an illusion that is maintained by the creation of a mentalizing narrative around one's thoughts and feelings, which in turn helps to create a coherent self-structure. If weakness of mentalizing capacity undermines this integrative process, incoherence in self-representation is likely to arise. Torturous feelings of badness, possibly linked to experiences of abuse which are felt to be part of the self but are not integrated with it (so-called "alien-self" parts), come to dominate self-experience. We assume that these discontinuities in self-experience (when the person feels aspects of their self-experience to be of themselves or their own, and yet also alien to their self-experience) generate a sense of incongruence, which is dealt with through externalizing – behaving toward others as though the others own the unmentalized self-experiences and on occasions even being successful in generating these experiences in them (Fonagy & Target, 2000). This brings about relief, even if the immediate impact of externalizing a torturing part of the self in this way is to manipulate another person into punitive persecutory behavior towards oneself. Because of their intensity, the person with BPD experiences no option but to rid themselves of these feelings and attempt to dominate the mind of others by "manipulativeness", self-injury, or other types of behavior that in the teleological mode are expected to relieve tension and arousal (Fonagy & Target, 2000).

Mentalizing Profiles: Complexities and Paradoxes

Understanding the different components and spectrums of mentalizing, as well as propensities towards certain modes of prementalization, is central to understanding mentalizing in BPD patients and how it should be targeted therapeutically (Fonagy & Luyten, 2009). Improved mentalizing can only genuinely be reached by appreciating the particular strengths and weaknesses in mentalizing that any patient may show, and how these abilities may be undermined or reinforced in different situations.

The therapist should recognize that an individual's capacity to mentalize varies across time according to their environment and stress arousal level. In particular, gaps in mentalizing in BPD patients are more likely to occur when the attachment system has been triggered (Chiesa & Fonagy, 2014; Fonagy et al., 1996). This inextricable link between mentalizing and attachment arousal means that mentalizing impairments in BPD may be partly relationship-specific. Assessment and evaluation of mentalization of BPD patients is therefore potentially unhelpful without regard to context. Any anomalies in relation to mentalizing are unlikely to be manifest in BPD patients unless the relationship in which mentalizing is being observed "pulls" for controlled mentalizing. The higher the level of attachment arousal in a particular relationship at a particular moment, the more likely that anomalies in mentalizing will emerge in these patients. Evidence strongly implies that as the attachment bond between therapist and client intensifies, the quality of BPD patients' mentalizing will tend to deteriorate (Diamond, Stovall-McClough, Clarkin, & Levy, 2003). Thus, initial assessment of clients can leave therapists with the impression that they are working with an individual with relatively high psychological mindedness (Bateman & Fonagy, 2012) and someone highly suitable for insight oriented psychotherapy. As trauma is typically associated with attachment insecurity, and anxious and disorganized attachment in particular, more traditional insight-oriented treatments might be especially risky (Fonagy & Bateman, 2006). Furthermore, transference typically intensifies as treatment progresses, activating the patient's internal working models of particular child–parent relationships and their attachment system in general; the quality of psychological mindedness is likely to deteriorate significantly and the patient's capacity to perceive the therapist's mind as different from his or her own mental state will be quite limited at times (Allen, Fonagy, & Bateman, 2008). Treatment should therefore aim to find the optimal balance between attachment

activation and mentalizing (Fonagy & Bateman, 2006) as well as applying an especially sensitive approach to mentalizing in the context of remembering trauma (Allen, 2012, 2013).

       With regards to the multi-dimensional nature of mentalizing, it has long been observed that BPD patients are not necessarily unable to mentalize; they may in some respects and in certain situations show normal or even superior mentalizing skills. Individuals with BPD commonly excel in one particular form of mentalizing: intuitive empathy. This contrast between eminent capability in intuitive empathy and drastic impairment in other areas of mentalizing is commonly referred to as the "empathy paradox". Dinsdale and Crespi (2013) reviewed 28 studies of empathic functioning in BPD in an attempt to resolve the so-called "empathy paradox". They report that in about half the studies, mentalizing – assessed in terms of levels of empathy – is enhanced in BPD. They particularly note the superiority of BPD patients in tasks that call for inference of others' intentions (as in Ladisich & Feil, 1988), or perceiving and responding strategically to small social cues indicating fairness (Franzen et al., 2011). Some studies have found that individuals with BPD are superior to normal controls in the accuracy with which they attribute mental states to others on the basis of external features (e.g., Domes et al., 2008; Lynch et al., 2006; Schulze, Domes, Koppen, & Herpertz, 2013); (yet see Matzke, Herpertz, Berger, Fleischer, & Domes, 2014; Mier et al., 2013), indicating the reliance of BPD patient on the use of external cues in their assessment of mental states. Furthermore, Ladisich and Feil (1988) assessed the perception of other people's feelings and personality by comparing individuals' self-ratings with personality ratings made by others (BPD patients, non-BPD patients, and psychiatrist expert raters). Ratings made by BPD patients matched self-ratings better than those of non-BPD patients, indicating higher empathy among the BPD patients, although they did not quite manage to outperform the expert raters. The experience of "borderline empathy" is familiar to most therapists suddenly struck by a remarkably insightful comment by their BPD patient.

However, the superior intuition often displayed by individuals with BPD will often be conspicuously absent in times of interpersonal stress and activation of the attachment system.

The dangers of not recognizing the unevenness and complexity of an individual's mentalizing performance can lead to the apparently iatrogenic effects often described in unsuccessful therapeutic encounters for BPD (Higgitt & Fonagy, 1992). Therapists working with BPD patients usually learn to recognize that exceptional interpersonal sensitivity should not be taken as an indication of psychological mindedness (Bateman & Fonagy, 2006). However, some therapists may misjudge this and address issues at a level of complexity which is simply beyond the patient's capacity to process (Fonagy & Bateman, 2006). Such unwarranted sophistication will result in failure to overcome the rigidity of thinking that characterizes patients with personality disorders in general. In the next section we will discuss the theory of epistemic trust and how this might account for the rigid, 'hard-to-reach' quality observed in individuals with BPD.

Mentalizing and Communication: A Common Factor in the Treatment of BPD?

We argue that the reasons for the importance of using mentalizing as a target for therapy in BPD partly lie in the particular attachment history and impairments in mentalizing, and related epistemic hypervigilance (Fonagy, Luyten, & Allison, 2014), of patients with BPD. The special significance of emotional neglect and abuse to BPD is particularly accounted for by the theory of epistemic trust (Sperber et al., 2010). Sperber suggests that evolution has prepared us to acquire culturally relevant new knowledge either because of its content (e.g., its deductive relations with other beliefs which we arrive at though logical inference) or, more commonly, we accept knowledge about our world and about ourselves from "teachers" on account of their authority as a source of knowledge, which overcomes natural and appropriate epistemic vigilance (Sperber et al., 2010). Such information may be considered to be "deferentially" transmitted, (Recanati, 1997) by someone whose

communication has the quality of epistemic trust: its source is known, remembered and

judged to be reliable (or epistemically trustworthy). This information can then be assumed to

be common knowledge shared by members of one's community.

But what of an individual whose social experiences have led them to a state of chronic

epistemic mistrust, where (perhaps because of hypermentalizing) they imagine the motives of

the communicator to be malign? In this context the individual will appear to be resistant to

new information, might be considered to be rigid or even bloody-minded, because they treat

new knowledge from the communicator with deep suspicion and will not internalize it (i.e.,

modify internal structures to accommodate it). Their epistemic trust has been undermined,

and an evolutionarily prepared channel for the acquisition of personally relevant information

is blocked. We suspect that it is less likely to be the frank brutality of abuse that undermines

epistemic trust (although of course it can do) and that neglect and emotional abuse will play a

larger role in making an individual excessively vulnerable to disturbances in trusting

information from others. Taking the perspective of epistemic trust as the mediator of culture,

and its key underlying engine for progression, we now consider *the destruction of trust in*

social knowledge as the key mechanism in pathological personality development.

It is the disturbance of epistemic trust in our opinion that generates the apparent

rigidity typical of BPD patients. The rigidity, however, is in the eyes of the communicator

who, in accordance with the principles of theoretical rationality, expects the recipient to

modify their behavior on the basis of the information they received and apparently

understood. But, in the absence of trust, the capacity for change is absent. The information

presented is not used to update the individual's social understanding. In terms of the theory of

natural pedagogy (Csibra & Gergely, 2009), the person has (temporarily) lost the capacity for

social learning. From a therapist's standpoint, he/she has become hard to reach and

interpersonally inaccessible. Research suggests that there may be different routes to

problems with epistemic distrust in BPD (Fonagy & Luyten, in press). In most patients, problems trusting others as a source of knowledge about the world reflect a combination of cognitive problems, impulsivity, and disruptions in attachment relationships, and trauma in particular. The typical affect-driven, automatic mentalizing of BPD patients further disrupts epistemic trust.

The Learning Systems Involved in Improved Mentalizing

It is an empirical fact that a number of psychotherapies appear to work roughly equally well in the treatment of BPD. Based on the mentalizing model of BPD and the model of communication outlined above, we argue that three systems underpin the process of therapeutic change in BPD. In our opinion, these three systems relate to each other in order, cumulatively, to make change possible. Yet, current views tend to favor and prioritize primarily the first of these systems, leading to a relative neglect of the two other systems.

Communication System 1: The teaching and learning of content

All evidence-based treatments of BPD provide a coherent, consistent, and continuous framework enabling the patient to examine, in a safe and low-arousal context, the issues that are deemed central according to the theoretical approach concerned (e.g., early schemas, invalidating experiences, object relations, attachment experiences). Across the course of therapeutic treatment – if the treatment is sufficiently coherent, reliable and predictable in its delivery – the patient begins to feel safe enough to allow a relaxation of epistemic hypervigilance. This enables the patient to reach a point where they can digest and contemplate the therapeutic "wisdom" being proposed to them. The relative importance of System 1 needs to be understood: therapies without a coherent body of knowledge based on systematically established principles have been observed to fail – to that extent, the content of this "wisdom" matters (Fonagy & Bateman, 2006). All evidence-based effective models of

therapy present the patient with models of mind, disorder, and processes of change that are convincing and accurate enough to let the patient relax their hypervigilance and see the model's relevance to their state of mind. This process of teaching and learning between therapist and patient – of showing the relevance of the wisdom (or the theory that underpins the orientation) – is made possible only by mentalizing and a collaborative working approach that models a mentalizing stance. However, the fact that there are so many different therapies, using so many different theoretical models that have been found to have some beneficial effect, indicates that the significance of System 1 lies not so much in the essential truth of its wisdom as in the fact that it allows the patient to put to apply this new learning in a more or less concrete way. This brings us to System 2.

Communication System 2: The re-emergence of social learning

Where System 1 concerns the content of *what* is learned in treatment, System 2 is concerned with *how learning is made possible again*. In the process of effectively passing on knowledge about the patient's condition (in System 1), the therapist uses ostensive cues (Russell, 1940; Wilson & Sperber, 2012), which signal meaningful communication of relevance. As Csibra and Gergely (2011) have established, this signaling works through conveying that the communicator is specifically and uniquely concerned with the individual being communicated to and is seeking to understand that individual's perspective. Ostensive communication with the patient is in essence modeling mentalizing. By creating an open and trustworthy social situation, a better understanding of the other is made possible. This in turn allows for a more trusting, less paranoid interpersonal relationship between therapist and patient, which facilitates not just the transmission of model-specific knowledge (as in System 1) but, through improving the patient's capacity to understand others, regenerates their potential to receive and absorb social information again. Ideally, the patient's feeling of having been sensitively responded to opens a virtuous cycle in interpersonal communications

in which mentalizing has been learned to be possible; this, in turn, facilitates learning from others about one's own mind as well as about the mind of others. By the therapist showing that their mind has been changed by the patient, they are giving agency to the patient and increasing their faith in the value of social understanding.

Communication System 3: Learning beyond therapy.

The erosion of epistemic hypervigilance and the improvements in mentalizing that the patient experiences in the context of the therapeutic relationship is likely to lead to meaningful change only if the virtuous circle of communication is maintained beyond therapy. Improvements in mentalizing lead to improvements in social relations, and a higher level of epistemic trust enables the individual to learn from their social experiences in a positive way. This goes beyond applying insights gained in treatment to relationships outside treatment, as the generalization of knowledge belongs to System 1 (and in part also to System 2). Learning beyond therapy involves much more, and is also not that easy. The advantages of these engagements with the wider social world are available to the patient only if their social environment is benign enough to generate such benefits.

The common feature in successful therapies, despite their often wildly different approaches, is thus, in our opinion, that they involve mentalizing. Mentalizing is a generic way of establishing epistemic trust, and therefore achieving change by being open to different kinds of social experience. Having one's subjectivity understood – that is, being mentalized by someone else – is the necessary key to reopening the ability to learn about and from the social world. The experience of being safely and appropriately thought about then makes the patient feel safe enough to go on to think about the wider social world. Mentalizing opens up communication – the epistemic superhighway for the transfer of personally relevant information, the fundamental biological route to information transmission – and, therefore, the possibility of changing perceptions, expectations and information about others, about

social processes, and the myriad knowledge elements that sustain all of us and support our adaptation to the social world.

Implications for Treatment

The Mentalizing Focus in Mentalization-Based Treatment

While we have already discussed general treatment principles following from our considerations concerning the nature and origins of BPD, the mentalizing approach has also led to the development and manualization of specific mentalization-based treatments, including MBT for BPD (Allen et al., 2008; Bateman & Fonagy, 2006). Here, we briefly summarize the core principles, interventions and treatment strategies of MBT.

Our theoretical model implies that in order to maximize the impact on the patient's ability to think about thoughts and feelings in relationship contexts, especially in the early phases of treatment, the therapist is probably most helpful when interventions (a) are simple and easy to understand, (b) are affect focused, (c) actively engage the patient, (d) focus on the patient's mind rather than their behavior, (e) relate to a current event or activity – whatever is the patient's currently felt mental reality (in working memory), (f) make use of the therapist's mind as a model (by therapists disclosing their anticipated reaction in response to the event being discussed, i.e., talking about how the therapist anticipates that he or she might react in that situation), (g) flexibly adjust complexity and emotional intensity in response to the intensity of the patient's emotional arousal (i.e., withdrawing when arousal and attachment are strongly activated).

The key task of therapy is thus to promote curiosity about the way mental states motivate and explain the actions of self and others. Therapists achieve this through the judicious use of the "inquisitive stance", highlighting their own interest in the mental states underpinning behavior, qualifying their own understanding and inferences (showing respect

for the opaqueness in mental states), and demonstrating how such information can help the patient to make sense of their experiences. Pseudomentalization and other fillers to replace genuine mentalization must be explicitly identified by therapists, and the lack of practical success associated with them clearly highlighted. In this way therapists can help their patients to learn about how they think and feel about themselves and others, how that shapes their responses to others, and how "errors" in understanding self and others may lead to inappropriate actions. Put simply, it is not for the therapist to "tell" patients about how they feel, what they think, or how they should behave, or what the underlying reasons may be for their difficulties. Any therapy approach that moves towards claiming to "know" how patients "are", how they should behave and think, and "why they are the way they are", is likely to be harmful to patients with a vulnerable capacity to mentalize.

This principle applies to CBT as much as to psychodynamic psychotherapy. For example, Davidson and colleagues demonstrated that high levels of therapist integrative complexity (an indication of the number of ideas being combined in a single statement) was associated with relatively poor outcome in CBT, while the patients' increase in integrative complexity marked improvement in social functioning (Davidson, Livingstone, McArthur, Dickson, & Gumley, 2007).

Individuals with BPD may perform experimental mentalizing tasks relatively well under low arousal (Arntz, Bernstein, Oorschot, & Schobre, 2009), but cannot explain the states of mind they experience under high arousal. Unfortunately, psychotherapists of many orientations often attempt to provide mentalistic understandings for issues that trigger intense emotional reactions (challenging interpersonal situations, issues of shame, guilt, feelings of inadequacy, etc.) at a time when the patient's capacity for effective explicit mentalization is practically inaccessible (Fonagy & Bateman, 2006). Particularly in severely disturbed BPD patients, treatments that strongly rely on reflective capacities might actually become

iatrogenic (Lemma, Target, & Fonagy, 2011; Luyten, Fonagy, Lemma, & Target, 2012). Such patients may benefit more from *a mental process* or *mentalizing* approach, where the focus is on distortions in processes related to the metacognitive ability to reflect on the self and others (Fonagy, Moran, & Target, 1993; Luyten, Blatt, & Fonagy, 2013). A titrated but more or less exclusive focus on the BPD patient's current mental state while activating the attachment relationship is expected to enhance the patient's mentalizing capacities without generating iatrogenic effects. Hence, treatment should avoid situations where patients are expected to talk of mental states that they cannot link to subjectively felt reality. In the case of BPD, we argue that the therapeutic aim needs to be reconfigured away from the traditional psychodynamic pursuit of insight, to an emphasis on the recovery of balanced mentalizing: the achievement of representational coherence and integration. Thus, with regard to dynamic therapies, there should be (a) a de-emphasis of "deep" unconscious interpretations in favor of conscious or near conscious content, (b) careful eschewing of descriptions of complex mental states (e.g., conflict, ambivalence, unconscious) that are incomprehensible to a person whose mentalizing is vulnerable, (c) avoidance of extensive discussion of past trauma, except in the context of reflecting on current perceptions of mental states of maltreating figures and changes in mental state from being a victim in the past versus one's experiences now. Careful attention to the BPD patient's current mental state, while activating the attachment relationship, serves to enhance the patient's mentalizing capacities without generating the iatrogenic effects that can arise from the activation of a disorganized attachment system.

MBT Structure and Protocol

MBT is organized around the development of an attachment relationship with the patient, offering a careful focus on the patient's internal mental processes as they are experienced moment by moment and emphasizing the therapeutic alliance with the active repair of ruptures in the patient–therapist relationship (Bateman & Fonagy, 2009).

Importantly, the treatment is delivered according to a carefully constructed protocol, which informs the therapist about how to manage common clinical situations using basic principles. In sum, these are:

1. A collaborative approach to and formulation of patient problems

2. Identification of non-mentalizing processes

3. General stance

   a) Interventions consistent with the patient's mentalizing capacity

   b) Monitoring of the state of affective arousal

   c) Focus on maintaining therapist mentalizing

   d) Openness of therapist's mind states and explicit identification of therapist's feelings related to patient's mental states

   e) Alert to breaks in mentalizing.

4. "Not-knowing" stance of curiosity

5. Identification of mentalizing poles

6. Trajectory of sessions: implementation and marking interventions structured from empathic validation to exploration, clarification, and challenge through affect identification and affect focus to mentalizing the relationship and counter-relationship.

1. Collaborative approach and formulation

The therapist generates a collaborative process in which both patient and therapist become inquisitive about agreed difficulties. Problems are initially organized in a mentalizing formulation, which is jointly agreed and reviewed every three months. Overall, the collaborative approach is embedded in the therapist's general stance and attitude.

2. Identification of non-mentalizing process

Clinicians carefully monitor for non-mentalizing processes. This is indicated for example by over-generalization, fixed beliefs of self and other, motives understood in terms of behaviors rather than mental states. Some non-mentalizing states are indicated by the therapist's behavior. Therapists who lose concentration or only grunt as the patient talks are often affected by pretend mode functioning in the patient; therapists suggest how to solve problems or tell the patient what to do without exploration are likely to be involved in teleological processes; the confused therapist who nods wisely is often struggling with understanding psychic equivalence. However, it should be noted that the three primary non-mentalizing modes are not mutually exclusive and are more likely to interweave than to manifest in pure form.

3. General stance

There are certain general principles for the therapist to follow in the clinical practice of MBT (Bateman, 2012).

a) Any intervention needs to *take into account the patient's mentalizing ability*. If their state of mind is held in psychic equivalence, complex interpretation or cognitive appraisal of the validity of the belief will be outside their comprehension. Such interventions need a higher level of mentalizing if they are to be effective. Fundamentally, the patient has to be able to think about his/her current state and appraise it without "living" it before such interventions are useful.

b) The therapist takes into account the patient's *current affective state*. If the patient is aroused, mentalizing will be compromised; the therapist needs to deal with the emotional dysregulation first. Once the patient begins to mentalize, the therapist can increase focus on affect.

c) The *therapist* ensures that he/she maintains his/her *own mentalizing*. This is a priority in MBT. If the therapist's mentalizing is lost, retrieving it takes precedence over all other processes; if necessary, the therapist temporarily stops the session to facilitate this. Openly retrieving one's own mentalizing is a useful model for the patient: "I am sorry I lost my ability to concentrate there and became too reactive when I said that. Let me go back…"

d) The *openness of the therapist's mind states in relation to the patient's mind states* is central to MBT. The therapist will talk about what is in his/her mind in relation to the patient's mind. In this regard there is a certain level of self-disclosure, but this is carefully "marked". Marking is explicitly identifying whether what is said is the therapist's state of mind, his/her subjective experience, or the therapist's representation of the patient's state of mind. For example, the therapist's statement "I am confused" is about the therapist's state, but "It comes across to me that you are confused" is about the patient.

e) The therapist is sensitive to *breaks in mentalizing* as evidenced by the dialogue, and is alert to changes in the patient's emotional arousal. This ensures that the patient focuses on actual mental processing as it happens and begins to recognize that emotions disrupt thoughts and interpersonal process.

A patient was talking about her children, who had been taken into care through child protection procedures, and her determination to have them back. She became angry about how she had been treated, then abruptly said that what was really a problem for her was that the gearbox had gone on her car and the drive shaft had broken. The therapist captured this sudden change in topic by saying "What happened there? How come you suddenly shifted to your broken car away from the children? What happened in your mind?"

The therapist also seeks to see whether he/she contributed to any break in mentalizing that occurs in the therapy. This will indicate how sensitive the patient is to aspects of interpersonal interaction. For example, when the therapist said "Calm down" to a patient in a concerned manner, the patient reacted strongly by saying "Don't tell me what to do. I am not taking orders from you". The therapist's remark had triggered a sensitive area, which could be recognized and explored to stimulate the patient's capacities to monitor how he/she processes his/her thoughts and feelings and what impairs these capacities.

4. Not-knowing stance

The not-knowing stance requires the therapist to work authentically from the perspective of equality and collaboration – the therapist can never know what is really going on in others' mind states. The therapist has to accept the validity of the patient's experience even if he/she does not understand it. The therapist does not have to understand the patient or to make sense of what seems incomprehensible. If the therapist does not know what the patient is talking about, the therapist does not try to piece it together but says: "You know, I am having a real problem here, I can't follow this, I can't put it together, can we try again?"

The aim of the not-knowing stance is to rekindle mentalizing in the session, to reflect on non-reflection as manifested in non-mentalizing. It is a key therapeutic attitude to enhance curiosity about mental process and experience. Curiosity is modelled by the therapist through reflecting on his/her own mind states without judgment and with empathic acceptance of experience.

5. Identification of Mentalizing poles

The therapist becomes attuned to indicators of non-mentalizing in the dialogue such as overuse of absolutes, simplistic over-determined explanations, and mental rigidity, which arises when the patient becomes stuck at one of the poles of mentalizing; the therapist tries to

rebalance this. For example, if the patient is highly other-externally focused, watching the therapist's movements carefully, the therapist directs the dialogue toward an internal state to see whether this instils more reflection.

A patient stated that the therapist was not to turn around to look at him as they walked down the corridor to the consulting room. The therapist asked what this backward glance did to the patient – an intervention to focus the patient on "self-internal" to balance his "other-external" focus. The patient kept the focus on the therapist and so the therapist accepted the patient's "other-external" focus by saying he turned around simply as a social gesture and that he was not aware of any wish to offend. Having done this, the therapist again tried to rebalance the focus by asking the patient to describe what he had experienced from the backward glance.

The same maneuver can be used if a person is excessively self-focused. The therapist intervenes to pull them out and get them to consider the "other" in some way, to balance their fixed focus on self.

The self–other dimension can be focused upon in two ways. The first is by getting the patient to shift internally toward thinking about the other – for example, if they are blaming themselves for a relationship problem, they may be supported in focusing on the role of the partner. The second way is by harnessing the intersubjective, interpersonal experience between the patient and therapist. MBT involves both ways. So the therapist could say to a patient, on the one hand, "How are you so sure that you are useless and that the relationship problem is all you? What about your boyfriend's role in this?" or, on the other hand, "I have never seen you like that. We seem to have a big difference here".

The same principle applies if a patient is excessively cognitive. The therapist balances this by harnessing the use of affective experience. This move to the affective pole can be difficult without becoming formulaic, for example, by continually asking someone how they

feel. This can be irritating for the patient, who may not know how they feel, and is often an unproductive intervention in terms of stimulating further mentalizing. For the therapist, the important factor is the quality of mentalizing – that is, whether it has become fixed and rigid rather than whether the mental processing is cognitive or affective. MBT recommends that the therapist increases interpersonal affectivity when the patient is fixed in a rigid cognitive rational process, or increases cognitive processing when the patient is trapped in affective dysregulation. To move from the cognitive pole, the therapist uses relational interventions in the dialogue (e.g., asking how the patient is *feeling* about someone); to move away from the affective pole, the therapist reduces the relational component and becomes more practical. For example a patient was talking about how she wanted her baby back. She stated 'no one else is having my baby'. 'They are not entitled to my baby. She is mine.' She had worked out that if she behaved herself in a particular way when she was with the social workers they would think that she was now emotionally controlled and release her baby from child protection. She did not want her child fostered so she had worked out a strategy for not telling people about problems. This rationale was presented in the session. The therapist considered this to be a cognitively determined level of mentalizing of both self and other so he increased the relational and affective component of the dialogue. He said that he could see that she had worked out what to do to meet her goal of not letting someone else have her baby. But he had always had a sense that it was her *love* for her baby that meant she wanted her back rather than the entitlement. This was an initial intervention to increase the affective component of the dialogue and to place the therapist's perspective as a relational aspect of the dialogue. From here the discussion moved to the relationship the patient had with her baby and if she could meet the baby's emotional needs. The aim of these interventions is to make a mentalizing process more flexible, more responsive to context, and increasingly implicit.

6. Trajectory of sessions

Empathy and support. MBT not only has an overall structure to the treatment programme (Bateman & Fonagy, 2006), but also identifies a trajectory for each session. In each session there is a recommended stepwise move from a supportive position toward a more relational, subjective, experiential process. MBT requires the therapist to start from an *empathic and supportive* position before moving toward a more relational focus. The therapist seeks to demonstrate an empathic understanding by using *validation* as the starting point, finding out the subjective truth of the patient's experience, and demonstrating that he/she has understood it from the patient's perspective. Only then can the therapist "sit alongside the patient" so that they both look at subjective experience from a shared vantage point.

A patient asked, "Do you know what I have gone and done?" The patient stated that what she had done was utterly crazy. "I've got myself a job." From the not-knowing stance, the therapist asked what was so crazy about it. The patient explained why it was crazy from her perspective. She had thought it was time she went to work and so got a part-time job, but explained that now she was working, paradoxically, she was earning less than when she had received unemployment benefit. So the therapist said "That is crazy – working and having less money and increased costs!" In this sense the therapist is empathic about her experience of being crazy. Taking an initial attitude that the decision was not crazy would have undermined the alliance and potentially left the patient feeling misunderstood. The complexity of the balance between the "craziness" and "about time I went to work" can be considered and explored later. The patient talked about getting less money, having to get out of bed, having to interact with people, and why the hell was she doing this? At this point the therapist was empathic, saying "That's a great question. Why do that?" It is from there that

the therapist and patient together work out what made her take the decision and what it had given her in terms of self-esteem.

Exploration and clarification. As soon as the therapist senses that he/she and the patient have a shared affective platform, exploration and elaboration takes place with the clarification of mental states. In the example above, the therapist helps the patient elaborate and clarify the processes that she went through to make her decision, and how she came to a final conclusion that she wanted employment despite it being "crazy". In addition, the therapist brings in some of his/her own thoughts about it. Clarification requires a reconstruction of events, but with an emphasis on the changing mental states, a tracing of process over time, and a recognition that decisions may be capricious and yet also of value.

Challenge. Challenge as an intervention has certain defined characteristics: it is nearly always outside the normal therapy dialogue, out of line with the current dialogue, and comes as a surprise to the patient. The aim is for the patient to be suddenly derailed in their non-mentalizing process. If the intervention is successful, the therapist "stops and stands" the moment to prevent the patient continuing in the same mode. Once a stop and stand challenge has been effective in halting non-mentalizing, it is important to rewind to the point at which either the patient or therapist was mentalizing.

A female patient was engaged in a diatribe about the prison service and its ill-treatment of prisoners. She was highly aroused, shouting, ranting, and "reliving" her anger and rage. Any interruption by the therapist resulted in a dismissive comment. The therapist looked out of the window, wondering how to intervene and thinking that a challenge was necessary. As he looked out of the window, the patient said, "Don't look out the window, you listen to me". So the therapist retorted that he could look and listen at the same time, stating that he could multi-task. Before the patient could respond, he said, "Do you know why I can multi-task? Because I am a man." At this point the patient stopped, not knowing whether to

laugh or react contemptuously. The therapist said that as a man he could only multi-task for a short time, so exhorted her to rest for a moment so that they could both collect their thoughts. This was a challenge: it was unexpected, and it stopped the ranting. The therapist was able to say that he thought it was better to sit back for a few minutes and rewind the session to start reflecting about what had happened that had led her to be sent to prison.

Affect focus. Once therapist and patient are able to maintain a mentalizing interaction, MBT suggests increasing focus on affect and the interpersonal domain. The purpose of this is to recreate the core sensitivity of BPD patients in the session itself. People with BPD are highly sensitive to interpersonal process; arousal in the interpersonal domain triggers emotional dysregulation, which in turn disrupts mental processing further. MBT for BPD focuses on this area of sensitivity to generate more robust mentalizing around interpersonal processing.

Affect focus is not simply labelling feelings – it is a way of increasing affective experience within the interpersonal relationship in the session by identifying implicit mentalizing and making it more explicit when the therapist and patient *share* some implicit process. It requires the therapist to recognize that both he/she and the patient are making unquestioned, jointly held, unspoken, assumptions. So the therapist names the experience as something that is shared between them.

A patient was anxious in a session and managed his arousal by turning away from the therapist, falling silent and then saying "Yeah, yeah, I don't know". Implicit in this interactive process was the patient's assumption that the therapist wanted him to talk more; there was some truth in this, for the therapist probed further at such times. But it was also apparent that the patient struggled with fears of becoming emotionally dysregulated to the extent of having to leave the session. The therapist was in a similar position with his assumption that the patient wanted to say more and his wariness that probing further could

increase the patient's anxiety. MBT suggests that the therapist identifies the shared dilemma, making the implicit anxiety more explicit. In this example, the therapist said "We are both uncertain at the moment. With me, I am concerned that if I probe more it will make things worse for you and yet this is an area that we have to explore more. It looks to me like you are saying 'don't go further because it might not be safe to continue'. Where are you in this?"

Having made this shared dilemma explicit, the therapist develops the mentalizing process around this interpersonal affectively charged area. To some extent this is a rehearsal *in vivo* of an affectively salient interpersonal interaction that may derail the patient in their close relationships. Accurate identification of the current affectively salient focus allows the therapist to segue to mentalizing the relationship without clumsily disrupting the interpersonal process.

Mentalizing the relationship. The aim of mentalizing the relationship is to increase the affective interpersonal experience with the patient whilst maintaining mentalizing. If an attempt to mentalize the relationship triggers non-mentalizing, the process is abandoned and the therapist returns to empathy and supportive work before trying to move down the relational trajectory again. If mentalizing continues as the focus on the relationship progresses, MBT suggests a number of steps for the therapist to take.

If the patient says something within the patient–therapist relationship that is of significance in the patient's external relationships, the first task for the therapist is to validate the patient's experience. Next, the therapist has to explore this sensitive area to get to an alternative perspective, or at least, a more complex understanding of what has happened.

A patient told her therapist that he was too modest. To validate this experience, the therapist asked the patient what he does that is "too" modest. Importantly, he did not question it as a distortion; it is a valid experience contributed to by the therapist's attitude. After the patient explained her reasons for believing that the therapist was too modest, and

identifying a recent example, the therapist's task is to identify what it is like for the patient to be working with a therapist who is too modest – what does it matter that she has a therapist who she sees as too modest? What actually had happened was that the patient felt that if the therapist could not be proud of his achievements, it reduced her own achievements in life to futile meaningless events, because she saw them as being minimal compared to those of the therapist. She experienced this in psychic equivalence, so her experience of her achievements as useless meant that she as a human being was useless.

Mentalizing the relationship is an attempt within the relationship to generate meaningful complexity about what has happened by engaging in a slowly unfolding relational process. At all times the therapist monitors the patient's reaction to the alternative perspective.

MBT adds caution in mentalizing the relationship. Side effects stimulated by the therapist are common: for example, if the patient's experience is seen by the therapist as a distortion and the patient is alienated, or if the process becomes a jointly elaborated pretend mode in which both patient and therapist believe that they are working at depth when in fact they are engaged in clever cognitive work that is out of contact with affective reality.

Mentalizing the counter-relationship. Mentalizing the counter-relationship, by definition, links to the therapist's self-awareness and often relies on the affective components of mentalizing. Some therapists tend to default to a state of self-reference whereby they consider most of what they experience in therapy as relevant to the patient. This default mode needs to be resisted; therapists need to be mindful that their mental states might unduly color their understanding of the patient's mental states and that therapists tend to equate them without adequate foundation. The therapist has to "quarantine" his/her feelings. How the therapist does so informs the MBT technical approach to countertransference – defined as those experiences, both affective and cognitive, that the therapist has in sessions which he/she

thinks might further develop an understanding of mental processes. Feelings in the therapist are not considered initially as a result of projective processes and the therapist must identify experiences clearly as his/her own. That is, they are "marked", and interventions using the counter-relationship are stated as the therapist's experience.

The simplest way to release countertransference experience from quarantine without equating the therapist's feeling with that of the patient is for the therapist to state "I" at the beginning of an intervention. Intriguingly, this seems to be difficult for therapists, who understandably worry about violating therapeutic boundaries. Yet MBT does not suggest that therapists start to express their personal problems or talk about any feeling they might have in a session, whether relevant to the process or not. Rather, the therapist's current experience of the process of therapy with the patient is to be shared openly, to ensure that the complexity of the interactional process can be considered. Patients need to be aware that their own mental processes have an effect on others' mental states and that these, in turn, will influence the interaction.

A patient with antisocial personality disorder sat in sessions leaning forward, pointing his finger at the therapist as he talked, and often using threatening language. Naturally this unnerved the therapist. The therapist was aware that the more unnerved he became the less likely it was for him to be able to maintain his mentalizing. At an appropriate moment in the discussion the therapist presented his current feelings in the therapy. 'Now you mention the way you intimidate people perhaps I can bring something up here related to that. I don't want to divert our discussion but many times in this session I have felt intimidated even if you don't mean to intimidate me. It is a problem because I find I can then not easily concentrate on what you are talking about.' At this point the patient interrupted to say that it was the therapist's problem if he felt intimidated. So the therapist agreed and said that indeed it was his problem but because of the effect it had on him it then became a problem for treatment.

This allowed the patient and therapist to consider intimidation in interpersonal interactions and it became apparent that the patient was unaware of the effect he was having on others by his attitude and discourse.

Common countertransference experiences associated with particular modes of psychological functioning include boredom with pretend mode, and anxiety to do something with psychic equivalence. Therapists need to become comfortable with managing these states of mind and be able to express them constructively in the service of extending the patient–therapist collaboration.

In all circumstances the therapist, once alerted by a change in his/her own feelings or behavior, should focus more carefully on his/her feeling and identify it openly – a move from implicit functioning to explicit process allowing a shared scrutiny.

It is highly likely that our formulations concerning the role of mentalizing in BPD also have implications for other evidence-based treatments of this condition. In fact, we believe that these formulations clarify the role of common factors in these treatments, and suggest the need to develop more comprehensive and integrative treatments that focus on restoring the capacity for social learning in patients with BPD and allied conditions.

Conclusion

This paper emphasizes the simultaneous consideration of disruptions in three closely linked domains in individuals with BPD: (a) in attachment relationships, (b) in different polarities of mentalizing, and (c) in the quality of epistemic vigilance and trust. Such a focus not only provides a comprehensive understanding of patients with BPD, rendering seemingly paradoxical features of patients with BPD more comprehensible. In so doing, this approach also provides a clear therapeutic focus, enabling the therapist to monitor the therapeutic

process in terms of (impending) mentalizing impairments and epistemic mistrust as a result of the activation of the attachment system. We believe that the effectiveness of the mentalizing approach in helping patients with BPD might be particularly explained by the fact that it enables the therapist to maintain and foster a mentalizing stance, even – and perhaps particularly – under high arousal conditions, so typical of work with these patients. As a result, MBT may lead to a relaxation of epistemic hypervigilance in these patients, opening them up for what fundamentally characterizes human beings: an openness to learn from others about oneself, others, and oneself in relation to others.

References

Aaronson, C. J., Bender, D. S., Skodol, A. E., & Gunderson, J. G. (2006). Comparison of attachment styles in borderline personality disorder and obsessive-compulsive personality disorder. *Psychiatric Quarterly, 77*, 69-80.

Agrawal, H. R., Gunderson, J., Holmes, B. M., & Lyons-Ruth, K. (2004). Attachment studies with borderline patients: A review. *Harvard Review of Psychiatry, 12*, 94-104.

Allen, J. G. (2012). *Restoring mentalizing in attachment relationships: Treating trauma with plain old therapy*. Washington, DC: American Psychiatric Press.

Allen, J. G. (2013). *Mentalizing in the development and treatment of attachment trauma*. London, UK: Karnac Books.

Allen, J. G., Bleiberg, E., & Haslam-Hopwood, G. T. G. (2003). *Mentalizing as a compass for treatment*. Houston, TX: The Menninger Clinic.

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

Arntz, A., Bernstein, D., Oorschot, M., & Schobre, P. (2009). Theory of mind in borderline and cluster-C personality disorder. *Journal of Nervous and Mental Disease, 197*, 801-807.

Bargh, J. A. (2011). Unconscious thought theory and its discontents: A critique of the critiques. *Social Cognition, 29*, 629-647.

Bargh, J. A. (2014). Our unconscious mind. *Scientific American, 310*, 30-37.

Barnow, S., Ruge, J., Spitzer, C., & Freyberger, H. J. (2005). Temperament und Charakter bei Personen mit Borderline-Personlichkeitsstorung [Temperament and character in persons with borderline personality disorder]. *Nervenarzt, 76*, 839-840, 842-834, 846-838.

Baron-Cohen, S., Golan, O., Chakrabarti, B., & Belmonte, M. K. (2008). Social cognition and autism spectrum conditions. In C. Sharp, P. Fonagy & I. Goodyer (Eds.), *Social cognition and developmental psychopathology*. Oxford, UK: Oxford University Press.

Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. J. (Eds.). (2000). *Understanding other minds: Perspectives from developmental cognitive neuroscience*. Oxford: Oxford University Press.

Barone, L., Fossati, A., & Guiducci, V. (2011). Attachment mental states and inferred pathways of development in borderline personality disorder: A study using the Adult Attachment Interview. *Attachment and Human Development, 13*, 451-469.

Bateman, A., & Fonagy, P. (2003). Health service utilization costs for borderline personality disorder patients treated with psychoanalytically oriented partial hospitalization versus general psychiatric care. *American Journal of Psychiatry, 160*, 169-171.

Bateman, A., & Fonagy, P. (2009). Randomized controlled trial of outpatient mentalization-based treatment versus structured clinical management for borderline personality disorder. *American Journal of Psychiatry, 166*, 1355-1364.

Bateman, A., & Fonagy, P. (2010). Mentalization based treatment for borderline personality disorder. *World Psychiatry, 9*, 11-15.

Bateman, A. (2012). Treating borderline personality disorder in clinical practice. *American Journal of Psychiatry, 169*, 560-563.

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

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

Bateman, A., & Fonagy, P. (Eds.). (2012). *Handbook of mentalizing in mental health practice*. Washington, DC: American Psychiatric Publishing.

Bateman, A., & Tyrer, P. (2004). Psychological treatment for personality disorders. *Advances in Psychiatric Treatment, 10*, 378-388.

Bender, D. S., & Skodol, A. E. (2007). Borderline personality as a self-other representational disturbance. *Journal of Personality Disorders, 21*, 500-517.

Blatt, S. J., & Auerbach, J. S. (1988). Differential cogntive disturbances in three types of borderline patients. *Journal of Personality Disorders, 2*, 198-211.

Chiesa, M., & Fonagy, P. (2014). Reflective function as a mediator between childhood adversity, personality disorder and symptom distress. *Personality and Mental Health, 8*, 52-66.

Choi-Kain, L. W., Fitzmaurice, G. M., Zanarini, M. C., Laverdiere, O., & Gunderson, J. G. (2009). The relationship between self-reported attachment styles, interpersonal dysfunction, and borderline personality disorder. *Journal of Nervous and Mental Disease, 197*, 816-821.

Choi-Kain, L. W., & Gunderson, J. G. (2008). Mentalization: Ontogeny, assessment, and application in the treatment of borderline personality disorder. *American Journal of Psychiatry, 165*, 1127-1135.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*, 148-153.

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

Davidson, K., Livingstone, S., McArthur, K., Dickson, L., & Gumley, A. (2007). An integrative complexity analysis of cognitive behaviour therapy sessions for borderline personality disorder. *Psychology and Psychotherapy, 80*, 513-523.

Diamond, D., Stovall-McClough, C., Clarkin, J. F., & Levy, K. N. (2003). Patient-therapist attachment in the treatment of borderline personality disorder. *Bulletin of the Menninger Clinic, 67*, 227-259.

Dinsdale, N., & Crespi, B. J. (2013). The borderline empathy paradox: Evidence and conceptual models for empathic enhancements in borderline personality disorder. *Journal of Personality Disorders, 27*, 172-195.

Domes, G., Czieschnek, D., Weidler, F., Berger, C., Fast, K., & Herpertz, S. C. (2008). Recognition of facial affect in borderline personality disorder. *Journal of Personality Disorders, 22*, 135-147.

Fonagy, P. (2003). The development of psychopathology from infancy to adulthood: The mysterious unfolding of disturbance in time. *Infant Mental Health Journal, 24*, 212-239.

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

Fonagy, P., & Bateman, A. (2008). The development of borderline personality disorder: A mentalizing model. *Journal of Personality Disorders, 22*, 4-21.

Fonagy, P., & Bateman, A. (2006). Mechanisms of change in mentalization-based treatment of BPD. *Journal of Clinical Psychology, 62*, 411-430.

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

Fonagy, P., Leigh, T., Steele, M., Steele, H., Kennedy, R., Mattoon, G., . . . Gerber, A. (1996). The relation of attachment status, psychiatric classification, and response to psychotherapy. *Journal of Consulting and Clinical Psychology, 64*, 22-31.

Fonagy, P., & Luyten, P. (2009). A developmental, mentalization-based approach to the understanding and treatment of borderline personality disorder. *Development and Psychopathology, 21*, 1355-1381.

Fonagy, P., & Luyten, P. (in press). A multilevel perspective on the development of borderline personality disorder. In D. Cicchetti (Ed.), *Development and psychopathology* (3rd ed.). New York, NY: John Wiley & Sons.

Fonagy, P., Luyten, P., & Allison, E. (2014). Teaching to learn from experience: Epistemic mistrust at the heart of BPD and its psychosocial treatment. *Manuscript in preparation*.

Fonagy, P., Luyten, P., & Strathearn, L. (2011). Borderline personality disorder, mentalization, and the neurobiology of attachment. *Infant Mental Health Journal, 32*, 47-69.

Fonagy, P., Moran, G. S., & Target, M. (1993). Aggression and the psychological self. *International Journal of Psycho-Analysis, 74 (Pt 3)*, 471-485.

Fonagy, P., & Target, M. (1997). Attachment and reflective function: Their role in self-organization. *Development and Psychopathology, 9*, 679-700.

Fonagy, P., & Target, M. (2000). Playing with reality: III. The persistence of dual psychic reality in borderline patients. *International Journal of Psychoanalysis, 81*, 853-874.

Fonagy, P., Target, M., Gergely, G., Allen, J. G., & Bateman, A. (2003). The developmental roots of borderline personality disorder in early attachment relationships: A theory and some evidence. *Psychoanalytic Inquiry, 23*, 412-459.

Franzen, N., Hagenhoff, M., Baer, N., Schmidt, A., Mier, D., Sammer, G., . . . Lis, S. (2011). Superior 'theory of mind' in borderline personality disorder: An analysis of interaction behavior in a virtual trust game. *Psychiatry Research, 187*, 224-233.

Fuchs, T. (2007). Fragmented selves: Temporality and identity in borderline personality disorder. *Psychopathology, 40*, 379-387.

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.

Gunderson, J. G., & Lyons-Ruth, K. (2008). BPD's interpersonal hypersensitivity phenotype: A gene-environment-developmental model. *Journal of Personality Disorders, 22*, 22-41.

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

Kernberg, O. F. (1984). *Severe personality disorders: Psychotherapeutic strategies*. New Haven, CT: Yale University Press.

Kovacs, A. M., Teglas, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science, 330*, 1830-1834.

Ladisich, W., & Feil, W. B. (1988). Empathy in psychiatric patients. *British Journal of Medical Psychology, 61 (Pt 2)*, 155-162.

Lawrence, K. A., Allen, J. S., & Chanen, A. M. (2010). Impulsivity in borderline personality disorder: Reward-based decision-making and its relationship to emotional distress. *Journal of Personality Disorders, 24*, 786-799.

Lemma, A. (2012). *Under the skin: A psychoanalytic study of body modification* London, UK: Routledge.

Lemma, A., Target, M., & Fonagy, P. (2011). *Brief dynamic interpersonal therapy: A clinician's guide*. Oxford, UK: Oxford University Press.

Levy, K. N., Beeney, J. E., & Temes, C. M. (2011). Attachment and its vicissitudes in borderline personality disorder. *Current Psychiatry Reports, 13*, 50-59.

Levy, K. N., Meehan, K. B., Kelly, K. M., Reynoso, J. S., Weber, M., Clarkin, J. F., & Kernberg, O. F. (2006). Change in attachment patterns and reflective function in a randomized control trial of transference-focused psychotherapy for borderline personality disorder. *Journal of Consulting and Clinical Psychology, 74*, 1027-1040.

Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder*. New York, NY: Guilford Press.

Luyten, P., Blatt, S. J., & Fonagy, P. (2013). Impairments in self structures in depression and suicide in psychodynamic and cognitive behavioral approaches: Implications for clinical practice and research. *International Journal of Cognitive Therapy, 6*, 265-279.

Luyten, P., Fonagy, P., Lemma, A., & Target, M. (2012). Depression. In A. Bateman & P. Fonagy (Eds.), *Handbook of mentalizing in mental health practice* (pp. 385-417). Washington, DC: American Psychiatric Association.

Luyten, P., Fonagy, P., Lowyck, B., & Vermote, R. (2012). Assessment of mentalization. In A. W. Bateman & P. Fonagy (Eds.), *Handbook of mentalizing in mental health practice* (pp. 43-65). Washington, DC: American Psychiatric Publishing.

Lynch, T. R., Rosenthal, M. Z., Kosson, D. S., Cheavens, J. S., Lejuez, C. W., & Blair, R. J. (2006). Heightened sensitivity to facial expressions of emotion in borderline personality disorder. *Emotion, 6*, 647-655.

Lyons-Ruth, K., Yellin, C., Melnick, S., & Atwood, G. (2005). Expanding the concept of unresolved mental states: Hostile/helpless states of mind on the Adult Attachment Interview are associated with disrupted mother-infant communication and infant disorganization. *Development and Psychopathology, 17*, 1-23.

Matzke, B., Herpertz, S. C., Berger, C., Fleischer, M., & Domes, G. (2014). Facial reactions during emotion recognition in borderline personality disorder: A facial electromyography study. *Psychopathology, 47*, 101-110.

Mier, D., Lis, S., Esslinger, C., Sauer, C., Hagenhoff, M., Ulferts, J., . . . Kirsch, P. (2013). Neuronal correlates of social cognition in borderline personality disorder. *Social Cognitive and Affective Neuroscience, 8*, 531-537.

Recanati, F. (1997). Can we believe what we do not understand? *Mind & Language, 12*, 84-100.

Rigoni, D., Brass, M., Roger, C., Vidal, F., & Sartori, G. (2013). Top-down modulation of brain activity underlying intentional action and its relationship with awareness of intention: An ERP/Laplacian analysis. *Experimental Brain Research, 229*, 347-357.

Russell, B. (1940). *An inquiry into meaning and truth*. London, UK: Allen & Unwin.

Schulze, L., Domes, G., Koppen, D., & Herpertz, S. C. (2013). Enhanced detection of emotional facial expressions in borderline personality disorder. *Psychopathology, 46*, 217-224.

Scott, L. N., Kim, Y., Nolf, K. A., Hallquist, M. N., Wright, A. G., Stepp, S. D., . . . Pilkonis, P. A. (2013). Preoccupied attachment and emotional dysregulation: Specific aspects of borderline personality disorder or general dimensions of personality pathology? *Journal of Personality Disorders, 27*, 473-495.

Sebastian, A., Jacob, G., Lieb, K., & Tuscher, O. (2013). Impulsivity in borderline personality disorder: A matter of disturbed impulse control or a facet of emotional dysregulation? *Current Psychiatry Reports, 15*, 339.

Sharp, C., Ha, C., Carbone, C., Kim, S., Perry, K., Williams, L., & Fonagy, P. (2013). Hypermentalizing in adolescent inpatients: Treatment effects and association with borderline traits. *Journal of Personality Disorders, 27*, 3-18.

Sharp, C., Pane, H., Ha, C., Venta, A., Patel, A. B., Sturek, J., & Fonagy, P. (2011). Theory of mind and emotion regulation difficulties in adolescents with borderline traits. *Journal of the American Academy of Child and Adolescent Psychiatry, 50*, 563-573.

Spengler, S., von Cramon, D., & Brass, M. (2010). Resisting motor mimicry: Control of imitation involves processes central to social cognition in patients with frontal and temporo-parietal lesions. *Social Neuroscience, 5*, 401-416.

Sperber, D., Clement, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language, 25*, 359-393.

Sroufe, L. A., Egeland, B., Carlson, E. A., & Collins, W. A. (2005). *The development of the person: The Minnesota Study of Risk and Adaptation from Birth to Adulthood*. New York, NY: Guilford Press.

Stoffers, J. M., Vollm, B. A., Rucker, G., Timmer, A., Huband, N., & Lieb, K. (2012). Psychological therapies for people with borderline personality disorder. *Cochrane Database of Systematic Reviews, 8*, CD005652.

Wilson, D. B., & Sperber, D. (2012). *Meaning and Relevance*. Cambridge, UK: Cambridge University Press.

Zanarini, M. C., & Frankenburg, F. R. (2007). The essential nature of borderline psychopathology. *Journal of Personality Disorders, 21*, 518-535.