Development and Psychopathology **21** (2009), 1355–1381
Copyright © Cambridge University Press, 2009
doi:10.1017/S0954579409990198

A developmental, mentalization-based approach to the understanding and treatment of borderline personality disorder

PETER FONAGY[a] AND PATRICK LUYTEN[b]
[a]*University College London; and* [b]*University of Leuven*

Abstract

The precise nature and etiopathogenesis of borderline personality disorder (BPD) continues to elude researchers and clinicians. Yet, increasing evidence from various strands of research converges to suggest that affect dysregulation, impulsivity, and unstable relationships constitute the core features of BPD. Over the last two decades, the mentalization-based approach to BPD has attempted to provide a theoretically consistent way of conceptualizing the interrelationship between these core features of BPD, with the aim of providing clinicians with a conceptually sound and empirically supported approach to BPD and its treatment. This paper presents an extended version of this approach to BPD based on recently accumulated data. In particular, we suggest that the core features of BPD reflect impairments in different facets of mentalization, each related to impairments in relatively distinct neural circuits underlying these facets. Hence, we provide a comprehensive account of BPD by showing how its core features are related to each other in theoretically meaningful ways. More specifically, we argue that BPD is primarily associated with a low threshold for the activation of the attachment system and deactivation of controlled mentalization, linked to impairments in the ability to differentiate mental states of self and other, which lead to hypersensitivity and increased susceptibility to contagion by other people's mental states, and poor integration of cognitive and affective aspects of mentalization. The combination of these impairments may explain BPD patients' propensity for vicious interpersonal cycles, and their high levels of affect dysregulation and impulsivity. Finally, the implications of this expanded mentalization-based approach to BPD for mentalization-based treatment and treatment of BPD more generally are discussed.

Borderline personality disorder (BPD) is a severe condition with a lifetime prevalence that has been estimated to be as much as 6% (Grant et al., 2008). BPD is often comorbid with mood and anxiety disorders, bipolar disorder, and schizotypal and narcissistic personality disorder (Grant et al., 2008). It may be particularly common in outpatient (Korzekwa, Dell, Links, Thabane, & Webb, 2008) and forensic populations (Black et al., 2007) where between 25 and 33% may be expected to meet criteria for the diagnosis. In recent years, the developmental emergence of BPD has become a topic of significant current interest (Cohen, 2008; Rogosch & Cicchetti, 2005), particularly as there is increasing evidence that the disorder has strong roots in early development. This paper presents a coherent, clinically relevant developmental model of BPD based on recently accumulated data. We summarize the key features of the disorder and link these features to impairments in specific facets of mentalization, each related to relatively specific neural systems, and discuss moderators of these relationships. Finally, we outline the implications of these findings for future research and clinical practice.

The ideas presented in this paper draw on the work of an active group of clinicians and scientists. We thank Anthony Bateman, Linda Mayes, Mary Target, George Gergely, Jon Allen, Efrain Bleiberg, Rudi Vermote, Benedicte Lowyck, and Elizabeth Allison.

Address correspondence and reprint requests 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.

Core Features of BPD

Although the precise nature of BPD remains elusive, there is growing consensus that the core features of the disorder consist of emotional dysregulation (e.g., Linehan, 1993; Reisch, Ebner-Priemer, Tschacher, Bohus, & Linehan, 2008), high levels of impulsivity (e.g., Grootens et al., 2008), leading to self-harm and suicidality (Black, Blum, Pfohl, & Hale, 2004), and disturbed interpersonal functioning (Hill et al., 2008), which is also expressed in high levels of preoccupied and disorganized attachment patterns in BPD patients (for a review, see Levy, Meehan, Weber, Reynoso, & Clarkin, 2005). In borderline patients we and others have noted a characteristic pattern of fearful attachment (attachment–anxiety and relational avoidance), painful intolerance of aloneness, hypersensitivity to social environment, expectation of hostility from others, and greatly reduced positive memories of dyadic interactions (e.g. Critchfield, Levy, Clarkin, & Kernberg, 2008; Gunderson & Lyons-Ruth, 2008).

We underscore three further facets of BPD that emerge as aspects of disturbed social relatedness from both factor analytic and clinical studies (e.g., Sanislow et al., 2002). First, dissociative comorbidity is more frequently reported by patients with BPD than other groups (e.g. Sar, Akyuz, Kugu, Ozturk, & Ertem-Vehid, 2006; Zanarini, Frankenburg, Jager-Hyman, Reich, & Fitzmaurice, 2008). It appears to be a response to increased stress (Stiglmayr et al., 2008), to be associated with emotional neglect (Sar et al., 2006) and to be linked to suicidal ideation (Klonsky, 2008). Second, BPD is also characterized by a disturbed sense of identity (e.g., Blatt & Auerbach, 1988; Bradley & Westen, 2005), rooted in a dysfunction or deficit of a sense of agency or self-directedness (e.g., Barnow, Ruge, Spitzer, & Freyberger, 2005; Bender & Skodol, 2007). Third, the experience of intense inner pain out of keeping with the significance of the event may be a core feature of BPD (Zanarini & Frankenburg, 2007) frequently associated with disturbed relationships (Gunderson & Lyons-Ruth, 2008), and childhood abuse and neglect (Holm & Severinsson, 2008). It is commonly prompted by social situations that involve rejection, abandonment, and/or isolation (e.g., Stiglmayr et al., 2008), includes an intense sense of shame (Levy, Edell, & McGlashan, 2007; Rusch et al., 2007), and can prompt dissociation (Philipsen et al., 2004) and self-harm (Coid, 1993).

If the unequivocal description of BPD in adulthood is problematic, its reliable description in adolescence is an even greater challenge and remains controversial in terms of the homogeneity of the category, comorbidity between and within axes, arbitrary cutoffs, and poor test–retest reliability (Miller, Muehlenkamp, & Jacobson, 2008). In addition, current dimensional models such as the five-factor model of personality may not necessarily be the most suitable for a detailed clinical description of maladaptive features of personality pathology at different ages (Blatt & Luyten, in press; De Clercq, De Fruyt, Van Leeuwen, & Mervielde, 2006; Westen, Dutra, & Shedler, 2005), given the potentially different phenotypic expression of trait symptoms at younger ages (Sharp & Romero, 2007). A classic 2-year prospective questionnaire study (Crick, Murray-Close, & Woods, 2005) with a representative sample of 400 fourth to sixth graders, for instance, found that the developmental precursors of BPD features in adolescence showed moderate stability, although "emotion sensitivity" (intense unstable inappropriate emotion) and "physical aggression" were no longer significant predictors once level of depression had been controlled for. The best predictors of BPD features from fourth to sixth grade were indicators of social dysfunction such as friend exclusivity (overly close relationship with friend), relational aggression (impulsivity), and cognitive sensitivity (hostile, untrusting paranoid world view). The adaptive functioning and psychiatric symptoms of adolescents who meet criteria for BPD are significantly worse than those from clinic samples who meet criteria for other PDs or no PD criteria (Chanen, Jovev, & Jackson, 2007). Specific features of BPD such as self-harm or traits of impulsivity and affective instability present during childhood and adolescence are predictive of receiving an adult BPD diagnosis mainly in retrospective studies (Zanarini et al., 2006). Both clinical and community based studies have found only moderate stability of the diagnosis in adolescence (Chabrol & Leichsenring, 2006) and the stability of these

diagnoses into adulthood is suggestive rather than well-established (Cohen, 2008; Deschamps & Vreugdenhil, 2008). On balance, the best evidence for developmental antecedents of BPD comes from examinations of the particular type of social dysfunction also characteristic of the adult form of the disorder comprising lack of trust and relational aggression. BPD probably consists of both stable and unstable features, basic temperamental characteristics that show stability over time, and more specific, potentially transient symptoms that are related to these personality traits in some yet to be specified way. BPD may consist of both symptoms that are manifestations of an acute illness and symptoms that represent more enduring aspects of the disorder, with many patients experiencing long term disability. However, it would be too simplistic to consider early established traits as tantamount to inherited vulnerabilities. The interaction of unfolding constitutional givens and the psychosocial environment within which they manifest and which they create is likely to be the leitmotif of both this and other personality disorders.

The Mentalization-Based Approach to BPD

It is our fundamental assumption that the core features of BPD are not isolated characteristics, but are related on various levels. Over the last two decades, we have attempted to provide a theoretically consistent way of conceptualizing the development and interrelationship of these features, aiming to provide clinicians with a conceptually sound and empirically supported approach to BPD.

Mentalization is a form of social cognition. It is the imaginative mental activity that enables us to perceive and interpret human behavior in terms of intentional mental states (e.g., needs, desires, feelings, beliefs, and goals; Allen, Fonagy, & Bateman, 2008; Bateman & Fonagy, 2006). Previously we have argued that impairments in this capacity play an important role in the development of various psychiatric disorders that involve pathology of the self (Bateman & Fonagy, 2004; Fonagy & Bateman, 2006b, 2008; Fonagy & Target, 2006; Luyten, Fonagy, Mayes, & Van Houdenhove, 2009; Sharp, Fonagy, & Goodyer, 2008). We have consistently argued the following:

1. Understanding the behavior of others in terms of their likely thoughts, feelings, wishes, and desires is a not a constitutional given but a developmental achievement.
2. The acquisition of this capacity depends on the quality of attachment relationships, particularly but not exclusively, early attachments, as the latter reflect the extent to which our subjective experience was adequately mirrored by a trusted other.
3. The quality of affect mirroring impacts on the development of affect regulative processes and self-control (including attention mechanisms and effortful control) as well as the capacity for mentalization.
4. Disruptions of early attachment and later trauma can disrupt the capacity for mentalization and, linked to this, the development of a coherent self-structure.
5. The capacity to mentalize has both "trait" and "state" aspects that vary in quality in relation to emotional arousal and interpersonal context.
6. Mentalization and the associated capacities for affect representation, affect regulation, and attentional control normally obscure forms of subjectivity that developmentally antedate mentalization.
7. The failure of mentalizing, in combination with profound disorganization of self-structure, may account for the core features of borderline personality functioning as described above. In particular, we have argued that borderline personality functioning can be understood as the consequence of (a) the loss of mentalization in emotionally intense relationship contexts, (b) the reemergence at these times of modes of thinking about subjective experience that antedate full mentalization, and (c) the constant pressure for externalization of internal states (projective identification), which we conceive of as the reexternalization of disorganized intolerably painful self-states (the self-destructive alien self).
8. Finally, we argue that a therapeutic intervention that focuses on the patient's capacity to

mentalize in the context of attachment relationships can be helpful in improving both behavioral and affective aspects of the condition.

This paper presents an expanded version of the mentalization-based approach to BPD based on recently accumulated data. In particular, we reexamine the mentalization construct in the light of current neuroimaging findings and suggest a four-component developmental model (Luyten, Fonagy, et al., 2009) that might further increase our understanding of the relationships among the core features of BPD from a developmental perspective.

Mentalization as a Multidimensional Construct

The concept of mentalization has been appropriately criticized as a marker of a specific form of psychopathology such as BPD because in its original formulation the concept was too broad and multifaceted to be operationalized (Choi-Kain & Gunderson, 2008; Holmes, 2005). Elsewhere (Luyten, Fonagy, et al., 2009), based on extant neuroimaging, developmental social and cognitive research, as well as heuristic considerations, we have proposed that mentalization can be characterized as organized along four polarities: automatic/controlled, cognitive/affective, internal/external-based, and self/other focused. Each of these dimensions reflects the involvement of two relatively distinct neural systems. Below we consider the relevance of each of these polarities to the difficulties with mentalization experienced by patients with BPD, and describe how these difficulties are related to the core features of BPD discussed above.

Implicit-automatic versus explicit-controlled mentalization

The distinction between automatic and controlled social cognition has existed in both social and cognitive psychology for some time, but it has only been applied relatively recently in the field of mentalizing (Lieberman, 2007; Satpute & Lieberman, 2006; Uddin, Iacoboni, Lange, & Keenan, 2007). Explicit mentalization is typically interpreted, conscious, verbal, and reflective. It is a serial and slow process that requires attention, intention, awareness, and effort. By contrast, automatic or implicit mentalization is perceived, nonconscious, nonverbal, and unreflective. It is typified by mirroring. It presumes parallel and therefore much faster processing, is reflexive, and requires little effort, focused attention or intention (Satpute & Lieberman, 2006). Developmentally, nonverbally determined implicit mentalizing appears to be robust early in the second year (Onishi, Baillargeon, & Leslie, 2007), or perhaps earlier (Gergely, Bekkering, & Kiraly, 2002), whereas verbal recognition of another's perspective is reliable only in the fourth (Carpendale & Lewis, 2006). Different brain areas may be recruited for automatic and controlled mentalizing. Automatic social cognition involves the amygdala, basal ganglia, ventromedial prefrontal cortex (VMPFC), lateral temporal cortex, and the dorsal anterior cingulate cortex (ACC), in general, phylogenetically older brain circuits that rely heavily on sensory information. By contrast, controlled mentalization recruits the lateral prefrontal cortex (LPFC), medial prefrontal cortex (mPFC), lateral parietal cortex (LPAC), medial parietal cortex (mPAC), medial temporal lobe (mTL), and rostral ACC (rACC), which are phylogenetically newer brain circuits involved in processing of linguistic and symbolic material.

Arousal may facilitate automatic mentalization and inhibit neural systems associated with controlled mentalization (Lieberman, 2007; Mayes, 2006), and BPD patients probably have a lower threshold for activation of flight–fight systems (Jogems-Kosterman, de Knijff, Kusters, & van Hoof, 2007), although as we will see this may be a more complex association than it appears (Kahl et al., 2006). Abnormal stress regulation and the developmental and neurological separation of these two systems, however, may explain how emotional arousal in BPD may on the one hand cause the loss of explicit mentalizing yet, on the other hand patients often appear implicitly particularly attuned to the states of mind of individuals around them. By the same token they appear able to perform experimental mentalizing tasks relatively well under low arousal (Arntz, Bernstein, Oorschot, Robson, & Schobre, 2006), but cannot explain the states of mind they experience under high arousal and show confusion about mental states at times

when they are dominated by reflexive assumptions about the internal states of others. 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 capacity for effective explicit mentalization is practically inaccessible (Fonagy & Bateman, 2006a).

Mentalization based on internal versus external features of self and others

Different neural networks appear to be involved depending on whether the focus of attention for making imaginative inferences about both our own and others' states of mind is internal or external (Lieberman, 2007). A focus on "mental interiors" means directly considering thoughts, feelings, and experiences, and is accompanied by the activation of a medial frontoparietal network (Satpute & Lieberman, 2006). Focusing mentalization on exteriors, on physical and visible features, or actions of another's or our own actions, appears to recruit a more lateral frontotemporoparietal network. As the mental interior may be thought of as a second-order representation of the exterior (Fonagy, Gergely, & Target, 2007), the internal/external distinction applies as much to self as other-related mentalization.

BPD patients primarily show deficits in making mental state judgments when these call for a direct focus on mental interiors. King-Casas and colleagues (2008) asked BPD patients to engage in a neuroeconomic social exchange task and found that they had difficulties in anticipating the likely impact of their actions on a partner player they never met before. Results showed that they were poor at "coaxing" their partner (i.e., generating a good feeling in them about the collaboration). By contrast, the weight of evidence suggests that mentalizing tasks that focus attention on external features cause fewer problems for these patients, and BPD patients have actually been found to be hypersensitive to facial expressions (e.g., Domes et al., 2008; Lynch et al., 2006).

These findings give rise to a model for the development of affect regulation based on the "external" versus "internal" attention distinction. We have suggested that mental interiors are discovered by the infant through the contingent interaction between external signs of affect in the child and their processed or "metabolized" or "marked" reflection by the caregiver (Fonagy, Gergely, Jurist, & Target, 2002; Gergely & Watson, 1996). In mirroring the infant the caregiver must achieve more than contingency (in time, space, and emotional tone). The mirroring must be "marked" (e.g., exaggerated), in other words, slightly distorted, if the infant is to understand the caregiver's display as part of the infant's emotional experience rather than an expression of hers (Fonagy et al., 2007; Gergely, 2004). We have speculated that the origin of affect regulation difficulties in patients with BPD may be associated with an early psychosocial environment where their internal experiences were not adequately mirrored. Noncontingent affect expressions by the parent will undermine the appropriate "labeling" of the infant's internal states (i.e., the establishment of introspectively accessible second-order representations for them) that may, in turn, remain confusing, experienced as unsymbolized, and hard to regulate. One recent study has suggested that individuals recruit the same neural tissue for both mother-referential and self-referential processing, indicating a strongly integrated network for the two (Vanderwal, Hunyadi, Grupe, Connors, & Schultz, 2008), a finding that is strikingly consistent with the suggestion that affect mirroring plays a critical role in self-development and development of mentalizing capacity.

Congruent with these assumptions, there is evidence that the quality of the caregiver's mirroring in the first 6 months may influence the child's capacity to regulate affect, as indicated by the strange situation (Gergely, Koós, & Watson, 2002; Koós & Gergely, 2001). However, it is equally likely that temperamental characteristics make certain infants' affects hard to read and their experience of affect mirroring will be limited as a consequence. Such gene–environment effects could be one way that the known influence of heredity (Distel et al., 2008; Lyons-Ruth et al., 2007; Ni et al., 2007) is mediated in BPD. It may be of particular significance that the polymorphism that appears to mark the infant's openness to environmental influence

(maternal sensitivity) in developing secure attachment, the *5-HTTLPR* short allele (Barry, Kochanska, & Philibert, 2008), is also a marker for vulnerability for developing borderline and antisocial traits (Lyons-Ruth et al., 2007). However, we have suggested that a major part of the vulnerability that evidence suggests was caused by early neglect (Johnson, Cohen, Chen, Kasen, & Brook, 2006) is mediated by compromise of normal affect mirroring in neglectful psychosocial environments (e.g., O'Connor, 2006). The concept of mutually responsive orientation advanced by Kochanska (1997) seems to us the best depiction of the qualities of the "cooperative interpersonal set" that ultimately generate the child's mental interior, which in turn normally generates self-regulation and conscience. The developmental experience of borderline children may fall short for reasons of constitution, neglect and emotional abuse, or an interaction of these factors.

Cognitive versus affective mentalization

The most commonly made developmental and neuroscientific distinction in relation to mentalization is between cognitive and affective content. Baron-Cohen, Golan, Chakrabarti, and Belmonte (2008) identified two independent processing systems that might normally be charged with parsing these categories of content: (a) the theory of mind mechanism (TOMM) mediates agent–attitude–proposition (or M-representations) such as "Mother–believes Johnny–took the cookies," and (b) the empathizing system (TESS) that uses self-affective state-proposition (E-representations) such as "I am sorry–you feel hurt–by what I said." The latter are appropriately labeled self-affective propositions because TESS is normally constrained always to create representations where emotion in the other is consistent with the self-affective state. Thus, it will not create "I am pleased that you are in pain" because it has to be a state that the self can generate in relation to the presumed state in the other. This constraint may not be present in psychopaths (Blair, 2008). Emotion understanding and belief–desire reasoning normally interact and only in combination generate genuine social understanding in what we have termed mentalized affectivity or "the feeling of feeling" (Fonagy et al., 2002).

Agent–attitude propositions are probably subserved by several areas within the prefrontal cortex, particularly cortical midline structures, whereas self-affective propositions appear to be more likely to be processed by a more automatic, embodied, and lateralized system including parts of the VMPFC (Sabbagh, 2004; Shamay-Tsoory & Aharon-Peretz, 2007; Shamay-Tsoory, Aharon-Peretz, & Levkovitz, 2007). The latter structure may be particularly important as TESS and TOMM representations come to be integrated, but it is currently not known how these two systems might function together to generate what we experience as empathy. If there are two systems, a more basic "emotional contagion" system associated with the inferior prefrontal gyrus and a more advanced cognitive perspective-taking system associated with the prefrontal cortex as a whole (Shamay-Tsoory, Aharon-Peretz, & Perry, 2009), then we might also expect that dysfunction in one system might lead to overcompensation in the other. In relation to emergent BPD, we might expect to find an overly influential TESS system overcompensating for a dysfunctional TOMM (VMPFC) system. This might yield characteristics such as susceptibility to emotional contagion, oversensitivity to some emotional cues, a predisposition to become overwhelmed by affect, and inability to integrate affective knowledge with more reflective and cognitive knowledge in relation to both other and self (Blatt, 2008). This is likely to have a particular impact on the individual's capacity for genuine empathy if (as is usually the case) the actual match of attitudes between self and other is imperfect. The constraint Baron-Cohen suggested for the TESS system (requiring consistency with the self-affective state) is likely often to lead to a state of affairs when the emotional state of the self is inappropriately extended to the other.

Mentalization with regard to self versus others

Under this heading we are less concerned with a dichotomy. Rather, we consider the implications of the surprising commonality between the developmental and brain processes underpinning mentalization of self and other. Neuroimaging studies have consistently supported the

assumption that envisioning the mind of another is underpinned by the same brain systems as identify one's own thoughts and feelings (Dimaggio, Lysaker, Carcione, Nicolo, & Semerari, 2008; Lieberman, 2007; Uddin et al., 2007). The common circuitry used in mentalizing self and other may explain the struggle of the normally developing child to acquire a sense of selfhood and the self-other confusions in disorders such as BPD, which may be associated with the disruption of these neural systems. In BPD, the capacity to distinguish self and other is severely impaired (e.g., Bender & Skodol, 2007; Blatt & Auerbach, 1988; Fuchs, 2007). We have argued that "identity diffusion" can be understood as an absence of a sense of agency that reflects a more general failure of mentalization, because the continuity of our sense of self is normally ensured via our sense of a connection between intentional mental states and actions (Fonagy, 1991).

Reviews of the neuroimaging literature suggest that two distinct neural networks are shared by self-knowing and knowing others (Lieberman, 2007; Uddin et al., 2007). The first system involves a more body-based, frontoparietal mirror neuron system that is involved in understanding the multimodal embodied self (e.g., face and body recognition) and understanding others through motor-simulation mechanisms (Gallese, Keysers, & Rizzolatti, 2004; Rizzolatti & Craighero, 2004). This suggests that a fundamental process that allows us to appreciate the actions and emotions of others involves the activation of the mirror neuron system for actions and the activation of visceromotor centers for the understanding of affect. This is thought to be one of the key evolutionary mechanisms underpinning social empathy: knowing from the inside how someone else feels. This is an implicit, automatic system, providing physical other to self and self to other mapping, which is involved in the immediate understanding (or misunderstanding) of self and others.

The sharing of structures not only implies a model for acquisition but also highlights the potential for conflating the embodied simulation of another person's experience with one's own experience. For instance, studies have shown that observing an action has a powerful influence on movement execution, with congruent observation priming actions while noncorresponding observation interferes with movement execution (e.g., Kilner, Paulignan, & Blakemore, 2003). Such conflations often appear to occur in BPD (Allen et al., 2008). Developmentally we have tended to assume a state of egocentrism in children; that is, that they are limited to interpreting the world from their own point of view (Piaget & Inhelder, 1948/1956). However, motor neuron findings suggest the opposite; that is, that infants have to develop the capacity to learn what is "me" as distinguished from states of mind assumed to be shared by others. So how does the child learn to differentiate self from others?

A second, cortical midline system that consists of the mPFC, ACC, and the temporal parietal junction in the LPAC (Lieberman, 2007; Uddin et al., 2007), appears to play a central role in this process. This system is less bodily based, and processes information about the self and others in more abstract and symbolic ways (Frith, 2007; Uddin et al., 2007). Importantly, unlike the frontoparietal system, it appears to be mainly shaped across development by interpersonal relationships.

Recently the work of Marcel Brass and colleagues has offered a suggestion about how these two apparently independent systems, both of which process representations of both self and other, may interact to create an experience of not-me and therefore me (Brass, Derrfuss, Forstmann, & von Cramon, 2005; Brass & Haggard, 2008; Brass, Schmitt, Spengler, & Gergely, 2007). If perceiving an intention in another triggers the impulse to perform the same behavior, the question of how automatic imitation can be avoided becomes a central issue. Neuroimaging studies indicate that the neural regions that are most often recruited in the inhibition of imitative behavior are the anterior frontomedial cortex (aFMC) and the temporoparietal junction (TPJ) area (Brass et al., 2005), which are cortical areas that are also related to mentalizing, self-referential processing, and self-agency. The TPJ is involved in perspective taking (e.g., Aichhorn, Perner, Kronbichler, Staffen, & Ladurner, 2006), in sense of agency (e.g., Decety & Grezes, 2006), and mentalizing (Frith & Frith, 2006). The aFMC is involved in mentalizing (Frith & Frith,

2006; Gilbert et al., 2007) and self-referential processing (Northoff et al., 2006). In fact, recent work by Brass's group, using a within-subject experimental design with reflective mentalizing and imitation–inhibition tasks, suggests that there is a functional relationship between the extent an individual can inhibit imitative behavior and the capacity for belief–desire reasoning (Brass et al., 2007). Thus, it appears that the inhibition of imitative behavior involves cortical areas that are also related to mentalizing, self-referential processing, and determining self-agency. We assume that this overlap reflects common underlying processes such as self/other distinction and decoupling of self and other.

Hence, by reflecting on the frontoparietally identified intent of the other, we can create a distinction between self and other experience and rapidly decouple the direct activation of corresponding motor representations. The capacity to inhibit imitative behavior may be key to enabling us to generate a sense of "me"-ness through achieving a "not-other"-ness. In other words, each time we interpret the actions of another, there may be a sequence in which an initial imitative matching response with the other within a motor neuron self-other system interacts with the reflective mentalizing self-other system. This interaction inhibits the mirror system and reduces the extent of "primary identification" with the other. To anticipate later considerations concerning the emergence of BPD slightly, we might hypothesize that the failure of medial prefrontal and temporoparietal mentalizing function in the course of development might leave the individual with difficulties in decoupling their representations of another person's experience from their self-representations. This would leave patients with BPD vulnerable to losing a sense of self in interpersonal interchange because they cannot adequately inhibit the alternative state of mind that is imposed on them through social contagion. Perhaps, then, the evident determination to "manipulate and control" the mind of others that is considered so characteristic of BPD patients can be seen as a self-preservative defensive reaction, defending the integrity of the self within attachment contexts. Without such control, they might feel excessively vulnerable to losing their sense of separateness and individuality.

In summary, as noted earlier, knowing oneself and others are two capacities that are clearly connected developmentally and in terms of the brain structures subserving these processes. This applies to both symbolic and automatic ways of knowing. The shared ontogeny and anatomy of these functions means that an intervention that encourages self-reflection within an interpersonal setting is well placed to modify even deeply rooted dysfunctional processes. The overlap in ontogeny and neural structure points to the interpersonal origins of the self who experiences itself and other selves as motivated by mental states: a developmental psychopathology approach that has increasingly strong empirical support (Hobson, 2002; Rochat, 2009). The acute suffering that experiences of separation bring for BPD patients becomes more understandable if we assume that when mentalization fails the withdrawal or physical disappearance of the attachment figure takes on catastrophic proportions. This is because in the absence of the capacity to mentalize, the individual is exposed to a direct and unmediated experience of vulnerability to others' reactions, now necessarily fragmented and impossible to differentiate from phenomenological self-experience.

Aspects of the Phenomenology of BPD

Thus far we have suggested that the failure of mentalization in BPD is specific to the explicit, internal, cognitive aspects. This impairs self–other differentiation, leaving the individual at risk of being overwhelmed by others' mental states. We have identified both constitutional and social influences that might undermine the normal development of mentalization, particularly the absence of marked mirroring. Does this model help us appreciate the subjective experience of individuals with BPD?

We assume that the absence of fully functioning mentalization is most evident through the reemergence of prementalizing modes of representing subjectivity. The clearest of these is the tendency to assume that mental states are direct representations of psychical reality, normal in a 20-month-old child (Gopnik & Meltzoff, 1997). The modes of social cognition that are characteristic of the ways of thinking of BPD patients

can be understood as prementalistic ways of social reasoning that reemerge with the disappearance of controlled mentalizing. Mentalization gives way to a kind of "psychic equivalence" (Target & Fonagy, 1996), which clinicians often consider under the heading of "concreteness of thought." What is thought is real. The hypothesis that a situation is dangerous ("I am being victimized" or "Everyone hates me") demands extreme measures of avoidance because experienced in the mode of psychic equivalence, even a passing thought feels real. No alternative perspectives are possible. There is a suspension of the experience of "as if." Everything, sometimes frighteningly, appears to be "for real." This can add drama as well as risk to interpersonal experience and patients' exaggerated reactions are justified by the seriousness with which they suddenly experience their own and others' thoughts and feelings. The vividness and bizarreness of subjective experience can appear as "quasipsychotic" symptoms (Zanarini, Gunderson, & Frankenburg, 1990), and is reminiscent of the physically compelling memories associated with posttraumatic stress disorder (PTSD; Morrison, Frame, & Larkin, 2003). For example, women with BPD not only report higher levels of proneness to shame and guilt, but they also show greater shame proneness on implicit tests of self-concepts such as the Implicit Association Test (Rusch et al., 2007). Shame is felt as "more real" by these patients than anxious patients or normal controls, and hence, the stronger association with self-esteem and quality of life. The extent to which internal experiences are experienced as if they are real events relates to psychotic features identified in this group that have been shown to be mediators between histories of childhood sexual abuse and suicidality (Soloff, Feske, & Fabio, 2008). Similar findings are also emerging in relation to anxiety sensitivity in these patients (Gratz, Tull, & Gunderson, 2008).

Evidence for the continued influence on adults of developmentally earlier modes of thought is available from studies of reasoning "errors" (e.g., hindsight bias, "the curse of knowledge," "actions speak louder than words"; Birch & Bloom, 2007; Blank, Nestler, von Collani, & Fischer, 2008; Wertz & German, 2007), which have been used to illuminate the architecture of the belief-desire reasoning processes. We believe that it is clinically helpful to view the mental functioning in BPD patients as a reemergence of developmentally earlier modes of mental functioning, because it offers us a better perspective on their experience of the world. To take just one example, both the young child and, at times, the individual with BPD, can have an overriding sense of certainty in relation to their subjective experience. Considering this developmentally, we assume that for good evolutionary reasons (see Csibra & Gergely, 2006) all young children consider the things that they are taught to be shared cultural knowledge available to all others. The small child assumes that his knowledge is knowledge held by all. Toddlers readily assume that other children will know facts that they themselves have just learned (Birch & Bloom, 2003). They will therefore also assume that there is nothing unique about their own thoughts or feelings (Fonagy et al., 2007). Developmentally, as they come to realize that not all knowledge is shared by all, a key aspect of theory of mind (ToM; Bloom, 2004), children normally learn the conditions under which this assumption should be suspended. Naïve realism (Pronin, Gilovich, & Ross, 2004), the failure to perceive one's own biases and to see others as more susceptible to a host of cognitive and motivational distortions, and the "curse of knowledge bias'" (Birch & Bloom, 2004), thinking that if one knows something about the world then everyone else knows it too, stem from the same developmental source predating the emergence of the self from this sense of "oneness." Developmentally it is not the overvaluing of private knowledge, but rather the undifferentiated experience of shared knowledge that hinders perspective taking prior to the development of the PFC. It is social centrism rather than egocentrism that remains a challenge for the individual with BPD, because their unusually intense mirroring representations of the other are not balanced by reflective processes. It is important for the clinician to be aware that it is not the overvaluing of their own perspective that can make BPD patients adopt an egocentric point of view; rather, their experience of their perspective feels universal. They expect other people, including their clinician, to know what they are thinking and

feeling, and to see situations in the same way they do. Thus, thwarting their intentions seems malign or willfully obtuse, rather than the result of a different point of view, alternative priorities, and so forth. This makes such frustrations not merely hurtful but intolerable and maddening, a denial of what they believe to be a shared reality.

Although thoughts can feel too real for the individual who has unreliable access to thoughts and feelings as second order representations, they can come to be almost dissociated to the point of near meaninglessness. The young child creates mental models and pretend worlds, but can maintain these only so long as they achieve complete separateness from the world of physical reality. Although psychic equivalence makes subjective experience too real, the pretend mode severs its connection with reality and may even lead to dissociative experiences. Dissociation, which we have seen is sometimes a response to stress induced loss of mentalization (Stiglmayr et al., 2008), may be an extreme indication of such separateness (Ross, 2007; Zanarini et al., 2008). In our view, the chronic sense of emptiness that has been found to characterize BPD is closely linked to both dissociation and suicidal ideation (Klonsky, 2008). In these states patients can discuss experiences without contextualizing them in any kind of physical or material reality. This could explain why psychophysiological reactivity to abandonment themes can be normal in BPD patients prone to dissociation (Schmahl et al., 2004). Reality places fewer limits on the creation of mental representations. Several studies using Rorschach, the Thematic Apperception Test, and other narrative methods have provided evidence of hypercomplex representations of mental states of others that are often seen as malevolent and idiosyncratically elaborated (Stuart et al., 1990; Westen, Ludolph, Lerner, Ruffins, & Wiss, 1990). There is also neuroimaging evidence that points to hyperactive mentalizing (at least in terms of TPJ/temporal parietal sulcus activation) in BPD patients in response to attachment related stimuli (Buchheim et al., 2006). Attempting psychotherapy with patients who are in this mode can lead the therapist to lengthy but inconsequential discussions of internal states that have no link to genuine experience.

The teleological mode, finally, refers to a mode of thinking that equates others' desires and feelings with their observable behavior. Developmentally early modes of conceptualizing action solely in terms of that which is apparent can come to dominate motivation. Within this "teleological mode" there is a primacy of the physical and observable. Experience is only felt to be valid when its consequences are apparent to all. Affection, for example, is only felt to be genuine if accompanied by a physical expression (e.g., a touch or caress). For many patients with BPD, one can only be loved if one is also physically touched.

The most socially disruptive feature of borderline cognition is the apparently unstoppable tendency to create unacceptable experience within the other. Externalization of the split-off parts of a disorganized self is desirable for the child with a disorganized attachment but is a matter of life and death for the traumatized individual who has internalized an abuser as part of the self. This is evident from the extreme levels of dysphoric affect reported by individuals with BPD, what Zanarini and colleagues (1998) have referred to as "the pain of being borderline." The unbearable emotional experience can include feeling abandoned, evil, betrayed, like a helpless child, misunderstood, mistreated, victimized, inferior, monstrous, and can characterize these patients' experience on a day-to-day basis in a relatively stable way. Patients can view themselves as permanently damaged, or rotten to the core (Zittel Conklin & Westen, 2005).

The externalization of these internal states are widely recognized in the common countertransferential reactions of therapists working with borderline patients: anger and hatred, helplessness and worthlessness, fear and worry, resentment, and urges to save and rescue the patient (Gabbard & Wilkinson, 1994). Complications come when therapists identify with the projection and emotionally distance themselves from individuals with BPD (Aviram, Brodsky, & Stanley, 2006). In addition to being unusually sensitive to rejection and abandonment, BPD patients may react negatively (e.g., by harming themselves or withdrawing from treatment) when they (mis)perceive such distancing and

rejection. As Aviram et al. (2006) point out, the stigmatizing attitude, engendered through the externalization of an alien part of the self, can exacerbate the behaviors in the patient that create the stigmatizing attitude. The result is a self-fulfilling prophecy and a cycle of stigmatization to which both patient and therapist contribute.

The alternative to projective identification is obtaining relief from experiences of overwhelming and intolerable emotion through the destruction of the self in a teleological mode, that is, physically, by self-harm and suicide (Kullgren, 1988). 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 pseudoattachment to this individual may develop. Indeed, studies of adult attachment patterns of BPD patients repeatedly highlight a combination of preoccupied and unresolved attachment associated with BPD (Agrawal, Gunderson, Holmes, & Lyons-Ruth, 2004; Fonagy et al., 1996; Nakash-Eisikovits, Dutra, & Westen, 2002).

Moderators of Mentalization

In this section, we consider the impact of potential moderators of the robustness of mentalization in relation to etiological factors of BPD from a developmental perspective.

Childhood attachment environment and the emergence of mentalization

A relation between attachment in infancy and early social understanding was first reported by Bretherton, Bates, Benigni, Camaioni, and Volterra (1979), who found that children who were securely attached at age 12 months used more protodeclarative pointing at age 11 months than other infants. Subsequently, a number of studies have reported associations between the quality of children's primary attachment relationship and the passing of standard ToM tasks somewhat earlier (see Fonagy, Gergely, & Target, 2007b; Sharp & Fonagy, 2008b). It should be noted that not all studies have found a relationship between attachment classification and ToM tasks, and the association is somewhat more likely to be observed for emotion understanding than ToM (Meins et al., 2002; Raikes & Thompson, 2006). The pathway connecting the two is unlikely to be a direct one although recent evidence suggests that the oxytocinergic system that is involved in caregiving behavior (Bartels & Zeki, 2004; Champagne, Diorio, Sharma, & Meaney, 2001) also enhances social awareness. Intranasal oxytocin, for instance, has been found to facilitate social function, including improving social memory, memory of facial expressions and identity, enhancing the recognition of mental states revealed by facial expression, probably by causing selective fixation on the eye region when viewing faces, and increasing manifestations of trust (for a review, see Neumann, 2008). Oxytocin may link sensitivity to social cues, such as infant facial expressions, with dopamine-associated reinforcement pathways (Strathearn, Fonagy, Amico, & Montague, in press). Secure attachment and mentalization may both be facilitated by aspects of parenting. In particular, a psychological perspective adopted by mothers in relation to their own actions or in relation to their child, including maternal "mind-mindedness" and "reflective function" as they interact with or describe their infants, is associated with both secure attachment and mentalization (for a review, see Sharp & Fonagy, 2008a). There is also a somewhat less well-established association between childhood adversity, the potential to develop disorganized attachment, and delays and problems with mentalization (Cicchetti, Rogosch, Maughan, Toth, & Bruce, 2003; Frodi & Smetana, 1984; Pears & Fisher, 2005; Rogosch, Cicchetti, & Aber, 1995; Smith & Walden, 1999). Mind-mindedness is likely to be one of those parental attributes that is most adaptive in moderation. Although evidence on this issue is still lacking, on the basis of clinical observations we have proposed that maladaptive aspects of parental mentalizing of a child can be either deficient (concrete and stimulus bound) or excessive or hypermentalizing (necessarily going beyond the data, often quite distorted and sometimes paranoid; Fearon et al., 2006; Williams et al., 2006).

As we have already noted, retrospective evaluations have led to recognition that disorganization of the attachment system is a key aspect of the psychopathology of BPD (Gunderson & Lyons-Ruth, 2008). Moreover, two longitudinal studies following children from infancy to early adulthood have found associations between insecure attachment in early childhood and BPD symptoms on follow-up (Lyons-Ruth, Yellin, Melnick, & Atwood, 2005; Rogosch & Cicchetti, 2005; Sroufe, Egeland, Carlson, & Collins, 2005). There is some suggestive evidence that the extent of the delay in emotion-focused mentalization observed in maltreated children is of developmental significance relevant to BPD. For example, the quality of understanding of the possible situational determinants of sad and angry emotions at approximately 6 years of age was found to predict social competence at 8 years of age, and the experience of physical abuse was found to predict social isolation at 8 years of age to the extent that it had impacted on emotion understanding, even controlling for verbal ability (Rogosch et al., 1995). Earlier, we discussed how emotion dysregulation and social competence problems are among the early precursors of borderline PD.

Yet, as many patients with BPD do not experience sexual or physical abuse (Paris, 2004), and the majority of abuse victims will not develop personality disorder (Binder, McNiel, & Goldstone, 1996; Horwitz, Widom, McLaughlin, & White, 2001), a model of BPD must be able to accommodate these findings. In our view, it is less the fact of maltreatment than being in a family environment that discourages coherent discourse concerning mental states that is likely to predispose the child to BPD. In line with our emphasis on the importance of a mirroring relationship for the establishment of mentalization, a number of studies support the view that neglect, low parental involvement, and emotional maltreatment rather than the presence of physical and sexual abuse are the critical predictors of BPD (e.g. Johnson et al., 2001; Ludolph et al., 1990). Studies that have examined the family context of childhood trauma in BPD tend to see the unstable, nonnurturing family environment as the key social mediator of abuse (Bradley, Jenei, & Westen, 2005) and underinvolvement the best predictor of suicide (Johnson et al., 2002) and personality dysfunction (Zweig-Frank & Paris, 1991). In one small longitudinal study, early maltreatment and disrupted parent–infant communication predicted BPD symptoms (Lyons-Ruth et al., 2005), and in a larger study verbal (emotional) abuse and neglect even more than physical maltreatment marked out those who went on to develop BPD (Johnson et al., 2001, 2002, 2006). In addition, anomalies in parenting and anxious attachment have been suggested as possible mediating mechanisms between low socioeconomic status and BPD symptoms (Cohen et al., 2008). Early neglect may indeed be an underestimated risk factor (Kantojarvi et al., 2008; Watson, Chilton, Fairchild, & Whewell, 2006), as there is some evidence from adoption and other studies to suggest that early neglect interferes with emotion understanding (e.g. Shipman, Edwards, Brown, Swisher, & Jennings, 2005) and this plays a role in the emergence of emotional difficulties in preschool (Vorria et al., 2006) and even in adolescence (Colvert et al., 2008). We have suggested that one developmental path to impairments in mentalizing in BPD is a combination of early neglect, which might undermine the infant's developing capacity for affect regulation, with later maltreatment or other environmental circumstances, including adult experience of verbal, emotional, physical and sexual abuse, that are likely to activate the attachment system chronically (Fonagy & Bateman, 2008, and see further discussion below). The finding of elevated PTSD scores among those 8.5-year-old children exposed to violence who had been disorganized in their attachment with their mothers at 12 months of age is also consistent with this suggestion (MacDonald et al., 2008). There is a marked convergence between our formulation concerning emotional abuse and the compelling proposals advanced by Linehan concerning the assumption of invalidating family environments of the young person with prodromal BPD (Linehan, 1993).

Stress, the attachment system, and mentalization

Full mentalizing is likely to fail to dominate behavior in the context of intense emotional

arousal as the fight or flight response comes online. Earlier formulations (e.g., Fonagy & Bateman, 2006b) proposed that, at extreme levels, the activation of the attachment system is associated with a deactivation of the mentalization system along with other emotion-induced cognitive dysfunction. Our current formulation emphasizes that in understanding the relation between emotional arousal and mentalizing, it is essential to go beyond a unitary concept of arousal (Robbins, 1997). Key neuromodulators, for example, contribute to different forms of arousal: norepinephrine contributes to alerting, vigilance, and controlled attentional processing in the face of stress; dopamine energizes approach behavior in response to potentially rewarding incentives; and serotonin modulates arousal in the norepinephrine and dopamine systems (Pliszka, 2003). Furthermore, the effects of arousal in any of these systems vary not only with the extent of transmitter secretion but also with the receptor subtype activated (Arnsten, Mathew, Ubriani, Taylor, & Li, 1999; Mayes, 2000). Following the model outlined by Mayes (2000, 2006) we suggest that with increased arousal there is a switch from cortical to subcortical systems, from controlled to automatic mentalizing and subsequently to nonmentalizing modes (see Figure 1). Mayes (2000, 2006) proposed that stress regulation is a differential balance of excitation and inhibition involving multiple, interactive neural systems with different neurochemical substrates regulating specific and different aspects of prefrontal, posterior cortical, and subcortical functions.

Two points are critical for understanding impairments in mentalizing in this context. First, owing to what can be construed as this neurochemical switch associated with escalating levels of emotional stress, patterns of brain functioning can shift from flexibility to automaticity, that is, from relatively slow executive functions mediated by the prefrontal cortex to faster habitual and instinctual behaviors mediated by posterior cortical (e.g., parietal) and subcortical structures (e.g., amygdala, hippocampus, and striatum). Concomitantly, mentalizing appears to disappear as self-protective physical reactions (fight–flight–freeze) come to dominate behavior. This has the presumed evolutionary value of promoting immediate adaptive responses to danger. In situations of interpersonal stress, where complex cognitive–emotional functioning (i.e., mentalizing) may be helpful, however, the loss of mentalization may be, to say the least, a significant inconvenience. Thus, the degree of arousal generated by interpersonal situations is critical. More generally, there will be situational variations when social stress triggers the threshold for switching from executive (mentalizing) to automatic (fight or flight) responding. We also assume, following Arnsten and Mayes, that the threshold for

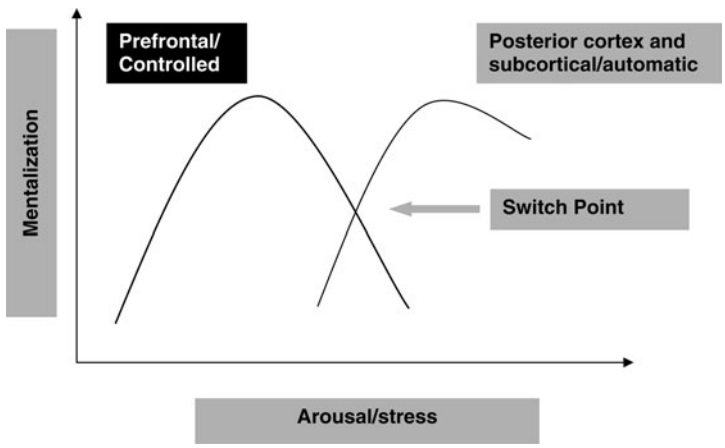

Figure 1. A biobehavioral switch model of the relationship between stress and controlled versus automatic mentalization (based on Luyten, Mayes, et al., 2009).

switching can be lowered as a result of exposure to early stress and trauma of the kind that has been documented for BPD patients, as noted earlier.

Second, both situational and more stable within-person variations may play a role in the switch from more controlled to automatic mentalization. This may also be a domain where genetic as well as developmental influences make themselves felt. In particular, research suggests that in BPD normal stress regulation, with a main role for cortisol, is disturbed (see Jogems-Kosterman et al., 2007). Functional magnetic imaging (fMRI) studies of BPD patients that manipulated the background level of stress and/or attachment system activation (e.g., Minzenberg, Fan, New, Tang, & Siever, 2007) confirm an abnormal pattern of frontal deactivation and associated hyperresponsiveness of the limbic system. For example, Silbersweig and colleagues (2007) reported that under conditions of negative emotion and behavioral inhibition, BPD patients showed relatively decreased ventromedial prefrontal activity (including medial orbitofrontal and subgenual anterior cingulated) and increased amygdalar–ventral striatal activity correlating with decreased constraint. Findings with implications for the hypothalamic–pituitary–adrenal axis function have confirmed that BPD patients, at least those with explicit trauma history, show a reduction in pituitary size (Garner et al., 2007), elevated cerebrospinal levels of corticotropin releasing hormone (Lee, Geracioti, Kasckow, & Coccaro, 2005), dysfunctions of cortisol responsivity (Jogems-Kosterman et al., 2007; Minzenberg et al., 2006; Walter et al., 2008), and disturbed dexamethasone suppression test response (Wingenfeld et al., 2007).

Although these studies require further replication in the light of sometimes quite small sample sizes, relatively disparate experimental paradigms, and considerable heterogeneity in sample selection (e.g., comorbidity with depression, childhood abuse, PTSD, and copying styles; Fertuck et al., 2006; Kahl et al., 2006), research provides considerable evidence that the activation and deactivation of the attachment system is closely linked to arousal and stress regulation (Heinrichs & Domes, 2008; Lieberman, 2007; Mayes, 2006). Studies on rodent models and neuroimaging studies have shown that activation of the attachment system underpinning both caregiving and partner preference is associated with the activation of two specific systems that have consistently been shown to play an important role in promoting and maintaining maternal behavior: (a) the dopaminergic reward processing system (Champagne et al., 2004; Strathearn, Li, Fonagy, & Montague, 2008) and (b) the oxytocinergic system (Bartels & Zeki, 2004; Champagne et al., 2001; Levine, Zagoory-Sharon, Feldman, & Weller, 2007). The functioning of the oxytocin mechanism provides an obvious way of accounting for the replicated association between attachment security and the developmental acquisition of mentalization competencies. Vasopressin may play an analogous role, perhaps especially in males (Caldwell, Lee, Macbeth, & Young, 2008; Lim et al., 2004).

Based on these findings, there are good reasons to suppose that different attachment histories are associated with attachment styles that differ in terms of the associated background level of activation of the attachment system, and the point at which the switch from more prefrontal controlled to more automatic mentalizing occurs (Luyten, Fonagy, et al., 2009). In particular, neuroimaging studies support the notion that attachment history affects the setting of the "switch," which turns the mentalizing system from planned, controlled and organized cognition to automatic processing with narrowed, poorly sustained attention, and increased vigilance for attachment disruptions such as rejection and abandonment.

For example, Gillath, Bunge, Shaver, Wendelken, and Mikulincer (2005) reported an fMRI study of women with high and low scores on attachment anxiety and avoidance who were asked to think about or stop thinking about various relationship scenarios. When thinking about negative scenarios, women with high levels of attachment anxiety showed more activation in the hippocampus (memory retrieval) and in emotion-related brain regions (the anterior temporal pole and the dorsal anterior cingulate), and less activation in the orbitofrontal cortex. It seems that in contemplating negative scenarios those with an anxious attachment style may underrecruit the brain areas associated with emotion regulation and show enhanced retrieval of negative memories. In a prospective study,

Strathearn and colleagues (2008, in press) examined whether differences in attachment security of 30 first-time mothers assessed using the Adult Attachment Interview (AAI; George, Kaplan, & Main, 1985) before the birth of their child, were related to brain reward and peripheral oxytocin response to infant cues. Mothers about 10 months after birth of their child viewed their own or others' infants' smiling and crying faces during fMRI scanning. Mothers with secure attachment showed greater activation of brain reward regions, including the ventral striatum, and the oxytocin-associated hypothalamus/pituitary region. Peripheral oxytocin response during infant contact was also significantly higher in secure mothers, and the size of change from baseline oxytocin levels was positively correlated with brain activation to own infants in both regions. The important differences based on attachment history emerged when the mothers viewed their infants' sad faces. Securely attached mothers continued to show greater activation in reward processing regions, whereas insecure/dismissing mothers, congruent with findings described earlier, showed reward system deactivation and insular activation in response to seeing their own infant's sad faces. The insula may be a region associated with feelings of unfairness, pain, and disgust (see review by Montague & Lohrenz, 2007). Mothers with insecure/dismissing attachment histories appeared less able to downregulate the sad feelings evoked in them by their infant's sad faces, possibly because they felt overwhelmed by sad memories of their own past. These findings suggest that for securely attached mothers, infant cues (whether positive or negative in affect) may act as an important affective signal of "incentive salience" (Berridge, 2007), reinforcing and motivating responsive maternal care. However, insecure mothers may experience a negative subjective reaction that would make them reflect (mirror) their infant's sadness without being able to create a symbolic/mentalizing distance between their infant's and their own states of mind.[1]

The finding of reduced "reward" activation in mothers with insecure/dismissing attachment is also consistent with a recent study of responses to smiling adult faces and positive task feedback (Vrticka, Andersson, Grandjean, Sander, & Vuilleumier, 2008). This study reported a negative correlation between dismissing attachment scores and activation of the ventral striatum. This is certainly congruent with the typical distant and detached relational style of avoidant individuals, and their tendency to deny the importance of attachment relationships. In both studies, challenges to the attachment system yielded greater ventral striatal activation in securely attached individuals, whereas in avoidant/dismissing individuals it was associated with a relative downregulation of reward-related activity. Consistent with the claims above about the complementary activation of the attachment and mentalizing systems, in the Vrticka et al. study, avoidant attachment was positively related to activation in the mPFC and the ventral (vACC), areas that have been implicated in controlled social cognition and mentalization, social rejection, and emotion suppression. Moreover, Vrticka and colleagues also reported that secure attachment was positively related to the activation of the ventral striatum in response to positive reinforcement, but negatively with activation of the amygdala to negative reinforcement. This is in line with previous suggestions we have made that secure attachment requires the simultaneous (paradoxical) activation of components that are normally reciprocally activated, mentalizing and reward-salience associated regions of the brain (Fonagy & Bateman, 2006b). These observations are also congruent with the assumption that secure attachment consists of a combination of low anxiety and low avoidance.

Based on these and other findings, we propose that a combination of characteristics strongly associated with attachment history is likely to determine whether an individual "switches" in a particular context from more controlled reflective to automatic mentalization (see Figure 1, and for a detailed discussion, see Luyten, Fonagy, et al., 2009).

1. The phenomenon of insecure mothers experiencing and therefore perhaps directly mirroring their infant rather than in a "marked" (nearly, but clearly not the same) manner may be one of the mechanisms behind the intergenerational transmission of patterns of attachment (van IJzendoorn & Bakermans-Kranenburg, 1997).

The studies reviewed suggest that the *anxious–preoccupied attachment strategies* that are characteristic of many BPD patients are associated with a lowered threshold for attachment system activation and, simultaneously, a lower threshold for controlled mentalization deactivation (see Figure 2). Thus, more automatic, subcortical systems, including the amygdala, have a low threshold for responding to stress in BPD patients. This hypothesis in and of itself could offer a comprehensive explanation for one of the central dynamic features of BPD patients, that is, their tendency to form attachments easily and quickly, often resulting in many disappointments. This pattern would be because of their low threshold for activation of the attachment system, and their low threshold for deactivation of neural systems associated with controlled social cognition, including neural systems involved in judging the trustworthiness of others. The vicious interpersonal cycles that are so characteristic of many BPD patients thus can be understood in terms of excitatory feedback loops leading to increased vigilance for stress-related cues in anxious attachment, particularly attachment characterized by high anxiety and high avoidance (Mikulincer & Shaver, 2007). These vicious cycles are also related to their hypervigilance concerning emotional states in others and their failure to deactivate a hyperactive TESS system, which further feeds into their lack of self-other differentiation, setting up a likely sequence of further failures in understanding their own internal world, that of others, and the relationship between the two (see Figure 2).

We assume that the arousal of the attachment system, beyond more general interpersonal stress induced arousal, leads to a general loss of mentalization. Any trauma arouses the attachment system (seeking for protection) and attachment trauma may do so chronically. In seeking proximity to the traumatizing attachment figure as a consequence of trauma, the child may naturally be further traumatized. The prolonged activation of the attachment system may be an additional problem, because the arousal of attachment may have specific inhibitory consequences for mentalization in addition to that which might be expected as a consequence of increased emotional arousal. In BPD it is therefore possible not only that trauma-related rapid triggering of fright–flight may account for the inhibition of mentalization, but more specifically that the hyperactivation of the attachment system in BPD may be a consequence of maltreatment specifically in an attachment context. Attachment theorists, in particular Mary Main and Erik Hesse, have suggested that maltreatment leads to the disorganization of the child's attachment to the caregiver because of the irresolvable internal conflict created by the need for reassurance from the very person who also (by association perhaps) generates an experience of lack of safety. The activation of the attachment system by the threat of maltreatment is followed by proximity seeking, which drives the child closer to an experience of threat leading to further (hyper)activation of the attachment system (Hesse, 2008). This irresolvable conflict leaves the child with an overwhelming sense of helplessness and hopelessness, and leads to what has been considered a hyperactivation of the attachment system. There can be no resolution of the anxiety, because looking for reassurance and protection generates more fear through the (mental) proximity of the maltreating figure. The ready triggering of the attachment system in BPD may be a residue of trauma history and manifests both as the rapidly accelerated tempo of intimacy in interpersonal relations and vulnerability to the temporary loss of mentalization.

In contrast, in individuals with *secure attachment*, the activation of the attachment system predictably involves a relaxation of normal strategies of interpersonal caution. Congruent with this assumption, expressions in most languages associate love with various severe forms of sensory handicap, particularly blindness. There is good evidence that intense activation of the neurobehavioral system underpinning attachment is associated with deactivation of arousal and affect regulation systems (Luyten, Mayes, et al., 2009), as well as deactivation of neurocognitive systems likely to generate interpersonal suspicion, that is, those involved in social cognition or mentalization, including the LPFC, mPFC, LPAC, mPAC, mTL, and rACC (Bartels & Zeki, 2000, 2004; Lieberman, 2007; Satpute & Lieberman, 2006). Moreover,

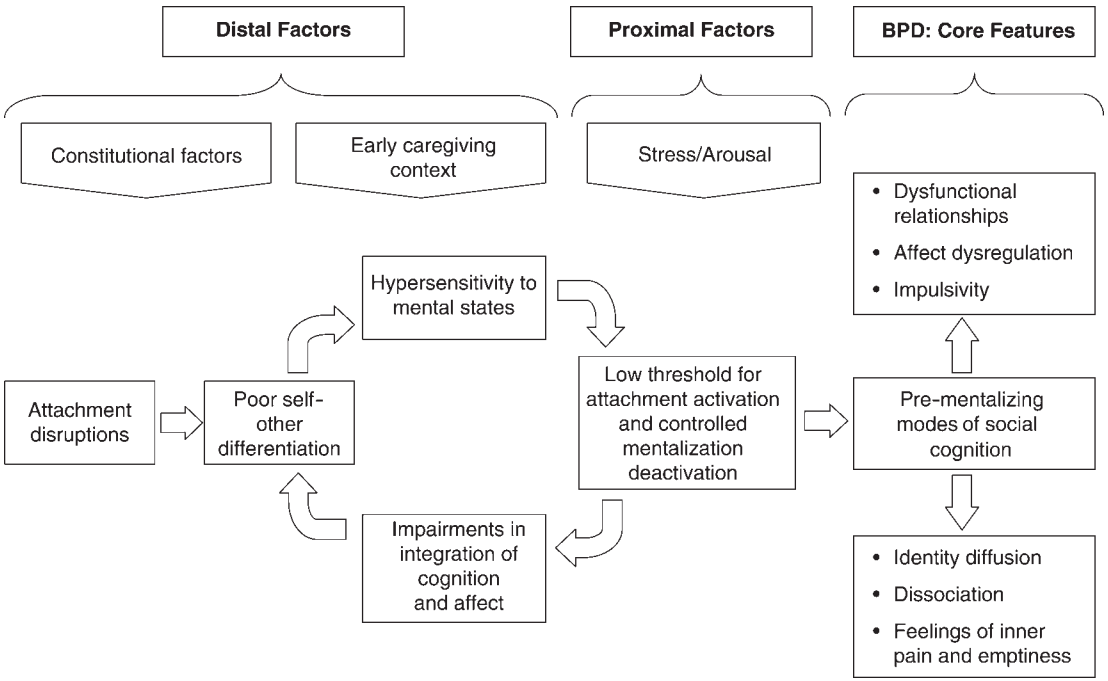

Figure 2. A mentalization-based model of BPD.

the neuropeptides oxytocin and vasopressin play key roles in two aspects of creating attachment relationships: (a) by activating the reward/attachment system (the "push" mechanism involved in attachment), and (b) by deactivating neurobehavioral systems that are involved in mediating social avoidance (the "pull" mechanism involved in attachment). For instance, oxytocin and vasopressin have been shown to inhibit aversion of both female and male rodents to infant pups, as well as leading to a number of affiliative behaviors, including caregiving behavior (Insel & Young, 2001). Oxytocin also reduces behavioral and neuroendocrine responses to social stress and seems both to enable animals to overcome their natural avoidance of proximity and to inhibit defensive behavior, thereby facilitating approach behavior. Vasopressin has primarily been implicated in male-typical social behaviors, including aggression and pair-bond formation, and mediates anxiogenic effects (Heinrichs & Domes, 2008). Thus, in the context of secure attachment, the activation of the attachment system will generate increased experience of reward, increased sensitivity, decreased social avoidance, but also the potential for the reward to override lack of trust.

This complex set of associations with social behavior may help us to account for the puzzling combination of facilitative and inhibitory associations between attachment history and social cognition. In two separate imaging studies, Bartels and Zeki (2000, 2004), for instance, reported that the activation of areas mediating maternal and/or romantic attachments appeared simultaneously to suppress brain activity in several regions mediating different aspects of cognitive control and including those associated with making social judgments and mentalizing. The medial prefrontal, inferior parietal, and medial temporal cortices mainly in the right hemisphere, as well as the posterior cingulate cortex, may be part of the circuitry specialized for attention and long-term memory (Cabeza & Nyberg, 2000), perhaps specifically responsible for integrating emotion and cognition (e.g., emotional encoding of episodic memories; Maddock, 1999) and subserving mood-mediated inhibition or enhancement of cognitive processing (Mayberg et al., 1999). The second set of brain areas observed to be deactivated by the activation of attachment concerns, included the temporal poles, parietotemporal junction, amygdala, and mesial prefrontal cortex, areas consistently linked to explicit and internally focused mentalization including judgments of social trustworthiness, moral judgments, ToM tasks, and attention to one's own emotions.

Individuals who use *attachment deactivation strategies*, typical of dismissive attached individuals, are able to keep the neural systems involved in controlled mentalization "on-line" for longer, including neural systems involved in judging the trustworthiness of other individuals (i.e., the "pull mechanism" associated with attachment; Vrticka et al., 2008). The distinction from securely attached individuals is clear. Secure individuals are able to keep the controlled mentalizing system "on line" even in the context of increased stress, which is less likely to trigger the attachment system, whereas dismissive individuals, for whom mild stress is not likely to trigger the attachment system, may be able to keep mentalization going until the stress becomes severe, at which point the deactivating strategy is likely to fail. If securely attached individuals are those who are able to retain a relatively high activation of prefrontal areas in the presence of the activation of the dopaminergic mesolimbic pathways (attachment and reward system), then differences in mentalization between securely attached and avoidantly/dismissively attached individuals may only show themselves under increasing stress, and this seems concordant with experimental studies (Mikulincer & Shaver, 2007). Although the threshold of avoidant individuals for switching from controlled to automatic mentalization might be elevated, studies have shown that under increasing levels of stress, these deactivating strategies tend to fail, leading to a strong reactivation of feelings of insecurity, heightened reactivation of negative self-representations, and increased stress (Mikulincer, Gillath, & Shaver, 2002).

By contrast, a low threshold for the stress-induced activation of the attachment system may translate as easy deactivation of the "pull-mechanism" of attachment, and a low threshold for activation of the "push-mechanism." In addition, we hypothesize that, if all other factors are constant, the greater an individual's use of

hyperactivating strategies, the lower will be their threshold for the activation of automatic mentalization and thus the stronger the relationship between stress and a switch to automatic mentalization will be (see Figure 2; Luyten, Fonagy, et al., 2009). Moreover, we predict that greater use of hyperactivating strategies will also be associated with increased time to recovery of mentalization and that deactivating strategies might be associated with relatively rapid recovery of the capacity for mentalization (Mikulincer & Shaver, 2007), but these predictions remain to be investigated.

However, this model would explain why mentalization deficits in BPD are more likely to be observed when the attachment system is triggered, such as in studies collecting AAI narratives (e.g., Fonagy et al., 1996), and why BPD patients who mix deactivating and hyperactivating strategies, as is characteristic of *disorganized attachment*, show a tendency for both hypermentalization and a failure of mentalization. Because attachment deactivating strategies are typically associated with minimizing and avoiding affective contents, BPD patients often have a tendency for hypermentalization, that is, continuing attempts to mentalize, but without integrating cognition and affect. At the same time, as the use of hyperactivating strategies is associated with a decoupling of controlled mentalization, this leads to failures of mentalization as a result of an overreliance on models of social cognition that antedate full mentalizing (Bateman & Fonagy, 2006). Similar conclusions have been drawn from an fMRI study in BPD patients where TAT cards elicited hyperactivation of the anterior cingulate and medial prefrontal cortices, suggesting an overly sensitive switch between emotionally salient and neutral information processing (Schnell, Dietrich, Schnitker, Daumann, & Herpertz, 2007).

Finally, it is important to note that BPD is in some ways at the opposite end of the spectrum from interpersonal resilience (Gunderson & Lyons-Ruth, 2008; Higgitt & Fonagy, 1992) and understanding the association between resilience and mentalization (Fonagy & Target, 1997) also provides helpful insight into BPD. Studies suggest that the ability to continue to mentalize even under considerable stress leads to so-called "broaden and build" (Fredrickson, 2001) cycles of attachment security, which reinforce feelings of secure attachment, personal agency, and affect regulation ("build"), and lead one to be pulled into different and more adaptive environments ("broaden"; Mikulincer & Shaver, 2007). Congruent with these assumptions, studies on resilience have shown that positive attachment experiences are related to resilience in part through relationship recruiting, that is, the capacity of resilient individuals to become attached to caring others (Hauser, Allen, & Golden, 2006). Hence, high levels of mentalization and the associated use of security-based attachment strategies when faced with stress are good candidates to explain the effect of relationship recruiting and resilience in the face of stress. Attachment hyperactivation and deactivation strategies, typically employed by anxious–preoccupied and dismissively attached individuals respectively, in contrast, have been shown to limit the ability to "broaden and build" in the face of stress. Moreover, they have also been shown to inhibit other behavioral systems that are involved in resilience, such as exploration, affiliation, and caregiving (Insel & Young, 2001; Mikulincer & Shaver, 2007; Neumann, 2008), which might also partially explain these individuals' difficulties in entering lasting relationships (including relationships with mental health care professionals) and the intergenerational transmission of psychopathology.

One major implication of these findings is that treatment should aim to find the optimal balance between attachment activation and mentalization (Fonagy & Bateman, 2006b). Moreover, this implies that treatments that focus on gaining insight into one's past, and particularly into one's traumatic past, which typically involves high levels of stress, might be contraindicated in patients with serious impairments in mentalization. As trauma is typically associated with attachment insecurity, and anxious and disorganized attachment in particular, insight-oriented treatments might be especially harmful for this subgroup of BPD patients (Fonagy & Bateman, 2006a). For instance, studies have shown that reflecting on oneself and one's autobiographical past and self-concept, capacities that are typically called upon in many forms of insight-oriented psychotherapy, are associated with activation of the

mPFC, which is deactivated under increasing arousal levels (Lieberman, 2007).

Relationship-specific mentalization: Mentalization in context

The association between mentalization, stress, and attachment reviewed in the previous section suggests that we should expect differences in the quality of mentalization depending on the quality of the relationship within which the aspect of interpersonal cognition is observed (Allen et al., 2008; Luyten, Fonagy, et al., 2009). This means that some revision of previously held assumptions is called for. Attachment theory and research is largely based on the assumption that cognitive–affective models of relationships (internal working models Bowlby, 1980) are relatively stable over time and are activated in a wide array of interpersonal relationships, including relationships with parents, partners, and friends (Fraley, 2007; Overall, Fletcher, & Friesen, 2003). Similarly, mentalization has often been assumed to be invariant across different relationships. For instance, current assessment methods of mentalization, such as the Reflective Functioning Scale as scored on the AAI (Fonagy, Target, Steele, & Steele, 1998), Child Attachment Interview (Ensink, 2003), and Parent Development Interview (Slade, 2005), involve the aggregation of mentalization across a number of attachment relationships. Moreover, because the Reflective Functioning Scale involves scoring mentalization based on states of mind with regard to past attachment experiences, research with this scale makes the further assumption that mentalization elicited in relation to past relationships will be related to current and even future relationships (Fonagy, Steele, Moran, Steele, & Higgitt, 1991).

However, research has called into question the assumption that working models are traitlike. The substantial within-person variation in internal working models of others (e.g., father vs. mother) (e.g., Fraley, 2007; Pierce & Lydon, 2001) supports a view of internal working models as hierarchically organized networks that contain both global and relationship-specific representations (Fonagy, 2001; Overall et al., 2003) or as distributed processes within a connectionistic network (Fraley, 2007). As mentalization is likely to be a function loosely coupled with the attachment system (Fonagy & Target, 1997), it is doubtful that individual differences in mentalization are best thought of as essentially traitlike. Mentalization is likely to show considerable fluctuations over time and across relationship contexts, not just as a function of stress but also as a function of the quality of particular relationships. Several sets of empirical observations are relevant to this assumption. First, general and relationship-specific mentalization, although overlapping, appear to be distinct (Happé & Frith, 1996; Humfress, O'Connor, Slaughter, Target, & Fonagy, 2002; O'Connor & Hirsch, 1999), and there is increasing agreement that key mentalization-linked experiences associated with social relationships, such as trust and fairness, are relationship-specific (Bugental & Johnston, 2000; Fiddick & Cummins, 2007). Second, several independent research traditions have focused on mentalization with regard to one's own infant as a relationship-specific form of mentalization (for a review, see Sharp & Fonagy, 2008a). Third, and most pertinent in the present context, is the line of evidence suggesting that quality of mentalization of therapists may be specific to particular patients and, even within that, may be variable across sessions of psychotherapy (Diamond, Stovall-McClough, Clarkin, & Levy, 2003; Diamond & Yeomans, 2008; Vermote et al., 2009).

These findings are congruent with neuroimaging studies suggesting an inverse relationship between amygdala activation and the activation of mentalizing areas (particularly mPFC and TPJ) in response to pictures or names of more or less familiar and personally connected individuals (Bartels & Zeki, 2000, 2004; Gobbini & Haxby, 2007; Gobbini, Leibenluft, Santiago, & Haxby, 2004; Ortigue, Bianchi-Demicheli, Hamilton, & Grafton, 2007). In line with these and the above observations concerning the situation and relationship specificity of mentalization, we suggest that any anomalies in relation to mentalization are unlikely to be manifest in patients with BPD unless the relationship in which mentalization is being observed "pulls" for the activation of these areas. The stronger the attachment in a particular relationship at a particular moment, the more likely

that anomalies in mentalization will emerge in these patients. Certainly, clinical evidence strongly implies that as the attachment bond between therapist and client intensifies, the quality of BPD patients' mentalization will tend to deteriorate. Thus, initial assessment of clients can leave therapists with the impression that they are working with an individual with relatively high psychological mindedness and someone highly suitable for insight oriented psychotherapy. Typically, as treatment progresses and transference intensifies, 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 et al., 2008).

The perspective we are taking also implies that in both research and clinical practice, assessment of mentalization of patients with BPD will be potentially unhelpful without regard to context, or based on specific categories of relationships. Interviews such as the AAI and the recently developed Child Attachment Interview (Shmueli-Goetz, Target, Fonagy, & Datta, 2008), as well as the Object Relations Inventory (Diamond, Blatt, Stayner, & Kaslow, 1991) allow scoring for mentalization with regard to self and others, as well as mentalization in specific relationships, as they elicit narratives about different significant others, and therefore provide a promising avenue for further research in this area.

Conclusions

Although there is increasing consensus regarding the core features of BPD, there is much disagreement over the features that should take predominance in explaining the etiopathogenesis of BPD. In this paper we suggest that the core features of BPD reflect impairments in different facets of mentalization, each related to impairments in relatively distinct neural circuits underlying these facets. Hence, we aim to provide a comprehensive account of the core features of BPD at multiple levels of analysis by showing how these core features are related with each other in theoretically meaningful ways. More specifically, we argue that BPD is associated with low thresholds for activation of the attachment system under stress, which, in combination with low thresholds for deactivation of the capacity for controlled mentalization, particularly with regard to differences in mental states of self versus others, renders the interpersonal world incomprehensible and leads to a cascade of impairments in other types of mentalization. This theory offers a framework for understanding BPD patients' propensity to become involved in vicious interpersonal cycles, their marked affective dysregulation, and high levels of impulsivity, as well as identity diffusion, dissociative experiences, and profound feelings of inner pain. Our formulations translate into a coherent treatment approach, which may also inform treatment of BPD across various theoretical orientations.

References

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.

Aichhorn, M., Perner, J., Kronbichler, M., Staffen, W., & Ladurner, G. (2006). Do visual perspective tasks need theory of mind? *NeuroImage*, *30*, 1059–1068.

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

Arnsten, A. F. T., Mathew, R., Ubriani, R., Taylor, J. R., & Li, B.-M. (1999). Alpha-1 noradrenergic receptor stimulation impairs prefrontal corical cognitive function. *Biological Psychiatry*, *45*, 26–31.

Arntz, A., Bernstein, D., Oorschot, M., Robson, K., & Schobre, P. (2006). *Theory of mind in borderline and Cluster-C personality disorder: No evidence for deficits*. Unpublished manuscript, Maastricht University.

Aviram, R. B., Brodsky, B. S., & Stanley, B. (2006). Borderline personality disorder, stigma, and treatment implications. *Harvard Review of Psychiatry*, *14*, 249–256.

Barnow, S., Ruge, J., Spitzer, C., & Freyberger, H. J. (2005). Temperament and character in persons with borderline personality disorder. *Nervenarzt*, *76*, 838–846.

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: Oxford University Press.

Barry, R. A., Kochanska, G., & Philibert, R. A. (2008). G × E interaction in the organization of attachment: Mothers' responsiveness as a moderator of children's genotypes. *Journal of Child Psychology and Psychiatry*, *49*, 1313–1320.

Bartels, A., & Zeki, S. (2000). The neural basis of romantic love. *NeuroReport*, *11*, 3829–3834.

Bartels, A., & Zeki, S. (2004). The neural correlates of maternal and romantic love. *NeuroImage*, *21*, 1155–1166.

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

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

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

Berridge, K. C. (2007). The debate over dopamine's role in reward: The case for incentive salience. *Psychopharmacology (Berlin)*, *191*, 391–431.

Binder, R. L., McNiel, D. E., & Goldstone, R. L. (1996). Is adaptive coping possible for adult survivors of childhood sexual abuse? *Psychiatric Services*, *47*, 186–188

Birch, S. A. J., & Bloom, P. (2003). Children are cursed: An asymmetric bias in mental-state attribution. *Psychological Science*, *14*, 283–286.

Birch, S. A. J., & Bloom, P. (2004). Understanding children's and adults' limitations in mental state reasoning. *Trends in Cognitive Science*, *8*, 255–260.

Birch, S. A. J., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science*, *18*, 382–386.

Black, D. W., Blum, N., Pfohl, B., & Hale, N. (2004). Suicidal behavior in borderline personality disorder: Prevalence, risk factors, prediction, and prevention. *Journal of Personality Disorders*, *18*, 226–239.

Black, D. W., Gunter, T., Allen, J., Blum, N., Arndt, S., Wenman, G., et al. (2007). Borderline personality disorder in male and female offenders newly committed to prison. *Comprehensive Psychiatry*, *48*, 400–405.

Blair, J. (2008). Empathic dysfunction in psychopathy. In C. Sharp, P. Fonagy, & I. Goodyer (Eds.), *Social cognition and developmental psychopathology*. Oxford: Oxford University Press.

Blank, H., Nestler, S., von Collani, G., & Fischer, V. (2008). How many hindsight biases are there? *Cognition*, *106*, 1408–1440.

Blatt, S. J. (2008). *Polarities of experience: Relatedness and self definiton in personality depvelopment, psychopathology, and the therapeutic process*. Washington, DC: American Psychological Association.

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

Blatt, S. J., & Luyten, P. (in press). A structural–developmental psychodynamic approach to psychopathology: Two polarities of experience across the life span. *Development and Psychopathology*.

Bloom, P. (2004). *Descartes' baby*. New York: Basic Books.

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

Bradley, R., Jenei, J., & Westen, D. (2005). Etiology of borderline personality disorder: Disentangling the contributions of intercorrelated antecedents. *Journal of Nervous and Mental Disease*, *193*, 24–31.

Bradley, R., & Westen, D. (2005). The psychodynamics of borderline personality disorder: A view from developmental psychopathology. *Development and Psychopathology*, *17*, 927–957.

Brass, M., Derrfuss, J., Forstmann, B., & von Cramon, D. Y. (2005). The role of the inferior frontal junction area in cognitive control. *Trends in Cognitive Science*, *9*, 314–316.

Brass, M., & Haggard, P. (2008). The what, when, whether model of intentional action. *Neuroscientist*, *14*, 319–325.

Brass, M., Schmitt, R. M., Spengler, S., & Gergely, G. (2007). Investigating action understanding: Inferential processes versus action simulation. *Current Biology*, *17*, 2117–2121.

Bretherton, I., Bates, E., Benigni, L., Camaioni, L., & Volterra, V. (1979). Relationships between cognition, communication, and quality of attachment. In E. Bates, L. Benigni, I. Bretherton, L. Camaioni, & V. Volterra (Eds.), *The emergence of symbols: Cognition and communication in infancy* (pp. 223–269). New York: Academic Press.

Buchheim, A., Erk, S., George, C., Kachele, H., Ruchsow, M., Spitzer, M., et al. (2006). Measuring attachment representation in an fMRI environment: A pilot study. *Psychopathology*, *39*, 144–152.

Bugental, D. B., & Johnston, C. (2000). Parental and child cognitions in the context of the family. *Annual Review of Psychology*, *51*, 315–344.

Cabeza, R., & Nyberg, L. (2000). Neural bases of learning and memory: Functional neuroimaging evidence. *Current Opinion in Neurology*, *13*, 415–421.

Caldwell, H. K., Lee, H. J., Macbeth, A. H., & Young., W. S. 3rd. (2008). Vasopressin: Behavioral roles of an "original" neuropeptide. *Progress in Neurobiology*, *84*, 1–24.

Carpendale, J. I. M., & Lewis, C. (2006). *How children develop social understanding (Understanding children's worlds)*. London: Blackwell.

Chabrol, H., & Leichsenring, F. (2006). Borderline personality organization and psychopathic traits in nonclinical adolescents: Relationships of identity diffusion, primitive defense mechanisms and reality testing with callousness and impulsivity traits. *Bulletin of the Menninger Clinic*, *70*, 160–170.

Champagne, F., Diorio, J., Sharma, S., & Meaney, M. J. (2001). Naturally occurring variations in maternal behavior in the rat are associated with differences in estrogen-inducible central oxytocin receptors. *Proceedings of the National Academy of Sciences of the United States of America, A 98*, 12736–12741.

Champagne, F. A., Chretien, P., Stevenson, C. W., Zhang, T. Y., Gratton, A., & Meaney, M. J. (2004). Variations in nucleus accumbens dopamine associated with individual differences in maternal behavior in the rat. *Journal of Neuroscience*, *24*, 4113–4123.

Chanen, A. M., Jovev, M., & Jackson, H. J. (2007). Adaptive functioning and psychiatric symptoms in adolescents with borderline personality disorder. *Journal of Clinical Psychiatry*, *68*, 297–306.

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.

Cicchetti, D., Rogosch, F. A., Maughan, A., Toth, S. L., & Bruce, J. (2003). False belief understanding in maltreated children. *Development and Psychopathology*, *15*, 1067–1091.

Cohen, P. (2008). Child development and personality disorder. *Psychiatric Clinics of North America*, *31*, 477–493.

Cohen, P., Chen, H., Gordon, K., Johnson, J., Brook, J., & Kasen, S. (2008). Socioeconomic background and the developmental course of schizotypal and borderline personality disorder symptoms. *Development and Psychopathology*, *20*, 633–650.

Coid, J. W. (1993). An affective syndrome in psychopaths with borderline personality disorder? *British Journal of Psychiatry*, *162*, 641–650.

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

Crick, N. R., Murray-Close, D., & Woods, K. (2005). Borderline personality features in childhood: A short-term longitudinal study. *Development and Psychopathology*, *17*, 1051–1070.

Critchfield, K. L., Levy, K. N., Clarkin, J. F., & Kernberg, O. F. (2008). The relational context of aggression in borderline personality disorder: Using adult attachment style to predict forms of hostility. *Journal of Clinical Psychology*, *64*, 67–82.

Csibra, G., & Gergely, G. (2006). Social learning and social cognition: The case for pedagogy. In M. H. Johnson & Y. M. Munakata (Eds.), *Processes of change in brain and cognitive developement. Attention and performance* (Vol. 21, pp. 249–274). Oxford: Oxford University Press.

De Clercq, B., De Fruyt, F., Van Leeuwen, K., & Mervielde, I. (2006). The structure of maladaptive personality traits in childhood: A step toward an integrative developmental perspective for *DSM-V*. *Journal of Abnormal Psychology*, *115*, 639–657.

Decety, J., & Grezes, J. (2006). The power of simulation: imagining one's own and other's behavior. *Brain Research*, *1079*, 4–14.

Deschamps, P. K., & Vreugdenhil, C. (2008). Stability of borderline personality disorder from childhood to adulthood: A literature review. *Tijdschrift voor Psychiatrie*, *50*, 33–41.

Diamond, A., Blatt, S. J., Stayner, D., & Kaslow, N. (1991). *Self–other differentiation of object representations*. Unpublished manuscript, Yale University.

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.

Diamond, D., & Yeomans, F. (2008). The patient–therapist relationship: Implications of attachment theory, reflective functioning, and research. *Santé Mentale au Québec*, *33*, 61–87.

Dimaggio, G., Lysaker, P. H., Carcione, A., Nicolo, G., & Semerari, A. (2008). Know yourself and you shall know the other . . . to a certain extent: Multiple paths of influence of self-reflection on mindreading. *Conscious Cognition*, *17*, 778–789.

Distel, M. A., Trull, T. J., Derom, C. A., Thiery, E. W., Grimmer, M. A., Martin, N. G., et al. (2008). Heritability of borderline personality disorder features is similar across three countries. *Psychological Medicine*, *38*, 1219–1229.

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.

Ensink, K. (2003). *Assessing theory of mind, affective understanding and reflective functioning in primary school aged children*. Unpublished PhD dissertation, University of London.

Fearon, P., Target, M., Fonagy, P., Williams, L., McGregor, J., Sargent, J., et al. (2006). Short-term mentalization and relational therapy (SMART): An integrative family therapy for children and adolescents. In J. Allen & P. Fonagy (Eds.), *Handbook of mentalisaiton based treatments*. London: Wiley.

Fertuck, E. A., Marsano-Jozefowicz, S., Stanley, B., Tryon, W. W., Oquendo, M., Mann, J. J., et al. (2006). The impact of borderline personality disorder and anxiety on neuropsychological performance in major depression. *Journal of Personality Disorders*, *20*, 55–70.

Fiddick, L., & Cummins, D. (2007). Are perceptions of fairness relationship-specific? The case of noblesse oblige. *Quarterly Journal of Experimental Psychology (Colchester)*, *60*, 16–31.

Fonagy, P. (1991). Thinking about thinking: Some clinical and theoretical considerations in the treatment of a borderline patient. *International Journal of Psycho-Analysis*, *72*, 1–18.

Fonagy, P. (2001). *Attachment theory and psychoanalysis*. New York: Other Press.

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

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

Fonagy, P., & Bateman, A. W. (2006b). 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: Other Press.

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

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

Fonagy, P., Steele, H., Moran, G., Steele, M., & Higgitt, A. (1991). The capacity for understanding mental states: The reflective self in parent and child and its significance for security of attachment. *Infant Mental Health Journal*, *13*, 200–217.

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

Fonagy, P., & Target, M. (2006). The mentalization-focused approach to self pathology. *Journal of Personality Disorders*, *20*, 544–576.

Fonagy, P., Target, M., Steele, H., & Steele, M. (1998). *Reflective-functioning manual, version 5.0, for application to Adult Attachment Interviews*. London: University College London.

Fraley, R. C. (2007). A connectionist approach to the organization and continuity of working models of attachment. *Journal of Personality*, *75*, 1157–1180.

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist*, *56*, 218–226.

Frith, C. D. (2007). The social brain? *Philosophical Transactions of the Royal Society of London B Biological Science*, *362*, 671–678.

Frith, C. D., & Frith, U. (2006). The neural basis of mentalizing. *Neuron*, *50*, 531–534.

Frodi, A., & Smetana, J. (1984). Abused, neglected, and nonmaltreated preschoolers' ability to discriminate emotions in others: The effects of IQ. *Child Abuse & Neglect*, *8*, 459–465.

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

Gabbard, G. O., & Wilkinson, S. M. (1994). *Management of contertransference with borderline patients*. Washington, DC: American Psychiatric Press.

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Science*, *8*, 396–403.

Garner, B., Chanen, A. M., Phillips, L., Velakoulis, D., Wood, S. J., Jackson, H. J., et al. (2007). Pituitary volume in teenagers with first-presentation borderline personality disorder. *Psychiatry Research*, *156*, 257–261.

George, C., Kaplan, N., & Main, M. (1985). *The Adult Attachment Interview*. Unpublished manuscript, University of California, Berkeley.

Gergely, G. (2004). The role of contingency detection in early affect–regulative interactions and in the development of different types of infant attachment. *Social Behavior*, *13*, 468–478.

Gergely, G., Bekkering, H., & Kiraly, I. (2002). Rational imitation in preverbal infants. *Nature*, *415*, 755.

Gergely, G., Koós, O., & Watson, J. S. (2002). Perception causale et role des comportements imitatifs des parents dans le développement socio-émotionnel précoce [Contingency perception and the role of contingent parental reactivity in early socio-emotional development]. In J. Nadel & J. Decety (Eds.), *Imiter pour découvrir l'human: Psychologie, neurobiology, robotique et philosophie de l'esprit* (pp. 59–82). Paris: Presses Universitaires de France.

Gergely, G., & Watson, J. (1996). The social biofeedback model of parental affect-mirroring. *International Journal of Psycho-Analysis*, *77*, 1181–1212.

Gilbert, S. J., Williamson, I. D., Dumontheil, I., Simons, J. S., Frith, C. D., & Burgess, P. W. (2007). Distinct regions of medial rostral prefrontal cortex supporting social and nonsocial functions. *Social Cognitive and Affective Neuroscience*, *2*, 217–226.

Gillath, O., Bunge, S. A., Shaver, P. R., Wendelken, C., & Mikulincer, M. (2005). Attachment-style differences in the ability to suppress negative thoughts: Exploring the neural correlates. *NeuroImage*, *28*, 835–847.

Gobbini, M. I., & Haxby, J. V. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia*, *45*, 32–41.

Gobbini, M. I., Leibenluft, E., Santiago, N., & Haxby, J. V. (2004). Social and emotional attachment in the neural representation of faces. *NeuroImage*, *22*, 1628–1635.

Gopnik, A., & Meltzoff, A. (1997). *Words, thoughts, and theories*. Cambridge, MA: MIT Press.

Grant, B. F., Chou, S. P., Goldstein, R. B., Huang, B., Stinson, F. S., Saha, T. D., et al. (2008). Prevalence, correlates, disability, and comorbidity of *DSM-IV* borderline personality disorder: results from the Wave 2 National Epidemiologic Survey on Alcohol and Related Conditions. *Journal of Clinical Psychiatry*, *69*, 533–545.

Gratz, K. L., Tull, M. T., & Gunderson, J. G. (2008). Preliminary data on the relationship between anxiety sensitivity and borderline personality disorder: The role of experiential avoidance. *Journal of Psychiatric Research*, *42*, 550–559.

Grootens, K. P., van Luijtelaar, G., Buitelaar, J. K., van der Laan, A., Hummelen, J. W., & Verkes, R. J. (2008). Inhibition errors in borderline personality disorder with psychotic-like symptoms. *Progress in Neuropsychopharmacology and Biological Psychiatry*, *32*, 267–273.

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

Happé, F., & Frith, C. D. (1996). Theory of mind and social impairment in children with conduct disorder. *British Journal of Developmental Psychology*, *14*, 385–398.

Hauser, S. T., Allen, J. P., & Golden, E. (2006). *Out of the woods: Tales of resilient teens*. Cambridge, MA: Harvard University Press.

Heinrichs, M., & Domes, G. (2008). Neuropeptides and social behaviour: Effects of oxytocin and vasopressin in humans. *Progress in Brain Research*, *170*, 337–350.

Hesse, E. (2008). The Adult Attachment Interview: Protocol, method of analysis, and empirical studies. In J. Cassidy & P. R. Shaver (Eds.), *Handbook of attachment theory and research* (2nd ed., pp. 552–558). New York: Guilford Press.

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

Hill, J., Pilkonis, P., Morse, J., Feske, U., Reynolds, S., Hope, H., et al. (2008). Social domain dysfunction and disorganization in borderline personality disorder. *Psychological Medicine*, *38*, 135–146.

Hobson, P. (2002). *The cradle of thought: Explorations of the origins of thinking*. Oxford: Macmillan.

Holm, A. L., & Severinsson, E. (2008). The emotional pain and distress of borderline personality disorder: A review of the literature. *International Journal of Mental Health Nursing*, *17*, 27–35.

Holmes, J. (2005). Notes on mentalization—Old hat or new wine? *British Journal of Psychotherapy*, *19*, 690–710.

Horwitz, A. V., Widom, C. S., McLaughlin, J., & White, H. R. (2001). The impact of childhood abuse and neglect on adult mental health: Prospective study. *Journal of Health and Social Behaviour*, *42*, 184–201.

Humfress, H., O'Connor, T. G., Slaughter, J., Target, M., & Fonagy, P. (2002). General and relationship-specific models of social cognition: Explaining the overlap and discrepancies. *Journal of Child Psychology and Psychiatry*, *43*, 873–883.

Insel, T., & Young, L. (2001). The neurobiology of attachment. *Nature Review Neuroscience*, *2*, 129–136.

Jogems-Kosterman, B. J., de Knijff, D. W., Kusters, R., & van Hoof, J. J. (2007). Basal cortisol and DHEA levels in women with borderline personality disorder. *Journal of Psychiatric Research*, *41*, 1019–1026.

Johnson, J. G., Cohen, P., Chen, H., Kasen, S., & Brook, J. S. (2006). Parenting behaviors associated with risk for offspring personality disorder during adulthood. *Archives of General Psychiatry*, *63*, 579–587.

Johnson, J. G., Cohen, P., Gould, M. S., Kasen, S., Brown, J., & Brook, J. S. (2002). Childhood adversities, interpersonal difficulties, and risk for suicide attempts during late adolescence and early adulthood. *Archives of General Psychiatry*, *59*, 741–749.

Johnson, J. G., Cohen, P., Smailes, E., Skodol, A., Brown, J., & Oldham, J. (2001). Childhood verbal abuse and risk for personality disorders during adolescence and early adulthood. *Comprehensive Psychiatry*, *42*, 16–23.

Kahl, K. G., Bens, S., Ziegler, K., Rudolf, S., Dibbelt, L., Kordon, A., et al. (2006). Cortisol, the cortisol–dehydroepiandrosterone ratio, and pro-inflammatory cytokines in patients with current major depressive disorder comorbid with borderline personality disorder. *Biological Psychiatry*, *59*, 667–671.

Kantojarvi, L., Joukamaa, M., Miettunen, J., Laksy, K., Herva, A., Karvonen, J. T., et al. (2008). Childhood family structure and personality disorders in adulthood. *European Psychiatry*, *23*, 205–211.

Kilner, J. M., Paulignan, Y., & Blakemore, S. J. (2003). An interference effect of observed biological movement on action. *Current Biology*, *13*, 522–525.

King-Casas, B., Sharp, C., Lomax-Bream, L., Lohrenz, T., Fonagy, P., & Montague, P. R. (2008). The rupture and repair of cooperation in borderline personality disorder. *Science*, *321*, 806–810.

Klonsky, E. D. (2008). What is emptiness? Clarifying the 7th criterion for borderline personality disorder. *Journal of Personality Disorders*, *22*, 418–426.

Kochanska, G. (1997). Mutually responsive orientation between mothers and their young children: Implications for early socialization. *Child Development*, *68*, 94–112.

Koós, O., & Gergely, G. (2001). The "flickering switch" hypothesis: A contingency-based approach to the etiology of disorganized attachment in infancy. *Bulletin of the Menninger Clinic*, *65*, 397–410.

Korzekwa, M. I., Dell, P. F., Links, P. S., Thabane, L., & Webb, S. P. (2008). Estimating the prevalence of borderline personality disorder in psychiatric outpatients using a two-phase procedure. *Comprehensive Psychiatry*, *49*, 380–386.

Kullgren, G. (1988). Factors associated with completed suicide in borderline personality disorder. *Journal of Nervous and Mental Disease*, *176*, 40–44.

Lee, R., Geracioti, T. D. Jr., Kasckow, J. W., & Coccaro, E. F. (2005). Childhood trauma and personality disorder: Positive correlation with adult CSF corticotropin-releasing factor concentrations. *American Journal of Psychiatry*, *162*, 995–997.

Levine, A., Zagoory-Sharon, O., Feldman, R., & Weller, A. (2007). Oxytocin during pregnancy and early postpartum: Individual patterns and maternal-fetal attachment. *Peptides*, *28*, 1162–1169.

Levy, K. N., Edell, W. S., & McGlashan, T. H. (2007). Depressive experiences in inpatients with borderline personality disorder. *Psychiatric Quarterly*, *78*, 129–143.

Levy, K. N., Meehan, K. B., Weber, M., Reynoso, J., & Clarkin, J. F. (2005). Attachment and borderline personality disorder: Implications for psychotherapy. *Psychopathology*, *38*, 64–74.

Lieberman, M. D. (2007). Social cognitive neuroscience: A review of core processes. *Annual Review of Psychology*, *58*, 259–289.

Lim, M. M., Wang, Z., Olazabal, D. E., Ren, X., Terwilliger, E. F., & Young, L. J. (2004). Enhanced partner preference in a promiscuous species by manipulating the expression of a single gene. *Nature*, *429*, 754–757.

Linehan, M. M. (1993). *Cognitive–behavioural treatment of borderline personality disorder*. New York: Guilford Press.

Ludolph, P. S., Westen, D., Misle, B., Jackson, A., Wixom, J., & Wiss, F. C. (1990). The borderline diagnosis in adolescents: Symptoms and developmental history. *American Journal of Psychiatry*, *147*, 470–476.

Luyten, P., Fonagy, P., Mayes, L., & Van Houdenhove, B. (2009). *Mentalization as a multidimensional concept*. Manuscript submitted for publication.

Luyten, P., Mayes, L., Fonagy, P., & Van Houdenhove, B. (2009). *The interpersonal regulation of stress*. Unpublished manuscript.

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., Holmes, B. M., Sasvari-Szekely, M., Ronai, Z., Nemoda, Z., & Pauls, D. (2007). Serotonin transporter polymorphism and borderline or antisocial traits among low-income young adults. *Psychiatric Genetics*, *17*, 339–343.

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.

MacDonald, H. Z., Beeghly, M., Grant-Knight, W., Augustyn, M., Woods, R. W., Cabral, H., et al. (2008). Longitudinal association between infant disorganized attachment and childhood posttraumatic stress symptoms. *Development and Psychopathology*, *20*, 493–508.

Maddock, R. J. (1999). The retrosplenial cortex and emotion: new insights from functional neuroimaging of the human brain. *Trends in Neuroscience*, *22*, 310–316.

Mayberg, H. S., Liotti, M., Brannan, S. K., McGinnis, S., Mahurin, R. K., Jerabek, P. A., et al. (1999). Reciprocal limbic–cortical function and negative mood: Converging PET findings in depression and normal sadness. *American Journal of Psychiatry*, *156*, 675–682.

Mayes, L. C. (2000). A developmental perspective on the regulation of arousal states. *Seminars in Perinatology*, *24*, 267–279.

Mayes, L. C. (2006). Arousal regulation, emotional flexibility, medial amygdala function, and the impact of early experience: Comments on the paper of Lewis et al. *Annals of the New York Academy of Sciences*, *1094*, 178–192.

Meins, E., Fernyhough, C., Wainwright, R., Das Gupta, M., Fradley, E., & Tuckey, M. (2002). Maternal mind-mindedness and attachment security as predictors of theory of mind understanding. *Child Development*, *73*, 1715–1726.

Mikulincer, M., Gillath, O., & Shaver, P. R. (2002). Activation of the attachment system in adulthood: Threat-related primes increase the accessibility of mental representations of attachment figures. *Journal of Personality and Social Psychology*, *83*, 881–895.

Mikulincer, M., & Shaver, P. R. (2007). *Attachment in adulthood: Structure, dynamics and change*. New York: Guilford Press.

Miller, A. L., Muehlenkamp, J. J., & Jacobson, C. M. (2008). Fact or fiction: Diagnosing borderline personality disorder in adolescents. *Clinical Psychology Review*, *28*, 969–981.

Minzenberg, M. J., Fan, J., New, A. S., Tang, C. Y., & Siever, L. J. (2007). Fronto-limbic dysfunction in response to facial emotion in borderline personality disorder: An event-related fMRI study. *Psychiatry Research*, *155*, 231–243.

Minzenberg, M. J., Grossman, R., New, A. S., Mitropoulou, V., Yehuda, R., Goodman, M., et al. (2006). Blunted hormone responses to Ipsapirone are associated with trait impulsivity in personality disorder patients. *Neuropsychopharmacology*, *31*, 197–203.

Montague, P. R., & Lohrenz, T. (2007). To detect and correct: Norm violations and their enforcement. *Neuron*, *56*, 14–18.

Morrison, A. P., Frame, L., & Larkin, W. (2003). Relationships between trauma and psychosis: A review and integration. *British Journal of Clinical Psychology*, *42*, 331–353.

Nakash-Eisikovits, O., Dutra, L., & Westen, D. (2002). Relationship between attachment patterns and personality pathology in adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry*, *41*, 1111–1123.

Neumann, I. D. (2008). Brain oxytocin: A key regulator of emotional and social behaviours in both females and males. *Journal of Neuroendocrinology*, *20*, 858–865.

Ni, X., Sicard, T., Bulgin, N., Bismil, R., Chan, K., McMain, S., et al. (2007). Monoamine oxidase a gene

is associated with borderline personality disorder. *Psychiatric Genetics*, *17*, 153–157.
Northoff, G., Heinzel, A., de Greck, M., Bermpohl, F., Dobrowolny, H., & Panksepp, J. (2006). Self-referential processing in our brain—A meta-analysis of imaging studies on the self. *NeuroImage*, *31*, 440–457.
O'Connor, T. (2006). The persisting effects of early experiences on social development. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology: Vol. 3. Risk, disorder and adaptation* (2nd ed., pp. 202–234). New York: Wiley.
O'Connor, T. G., & Hirsch, N. (1999). Intra-individual differences and relationship-specificity of social understanding and mentalising in early adolescence. *Social Development*, *8*, 256–274.
Onishi, K. H., Baillargeon, R., & Leslie, A. M. (2007). 15-Month-old infants detect violations of pretend scenarios. *Acta Pychologica*, *124*, 106–128.
Ortigue, S., Bianchi-Demicheli, F., Hamilton, A. F., & Grafton, S. T. (2007). The neural basis of love as a subliminal prime: An event-related functional magnetic resonance imaging study. *Journal of Cognitive Neuroscience*, *19*, 1218–1230.
Overall, N. C., Fletcher, G. J., & Friesen, M. D. (2003). Mapping the intimate relationship mind: Comparisons between three models of attachment representations. *Personality and Social Psychology Bulletin*, *29*, 1479–1493.
Paris, J. (2004). Sociocultural factors in the treatment of personality disorders. In W. J. Livesley (Ed.), *Handbook of personality disorders: Theory and practice* (pp. 135–147). New York: Wiley.
Pears, K. C., & Fisher, P. A. (2005). Emotion understanding and theory of mind among maltreated children in foster care. *Development and Psychopathology*, *17*, 47–65.
Philipsen, A., Richter, H., Schmahl, C., Peters, J., Rusch, N., Bohus, M., et al. (2004). Clonidine in acute aversive inner tension and self-injurious behavior in female patients with borderline personality disorder. *Journal of Clinical Psychiatry*, *65*, 1414–1419.
Piaget, J., & Inhelder, B. (1956). *The child's conception of space*. London: Routledge. (original work published 1948)
Pierce, T., & Lydon, J. E. (2001). Global and specific relational models in the experience of social interactions. *Journal of Personality and Social Psychology*, *80*, 613–631.
Pliszka, S. R. (2003). Non-stimulant treatment of attention-deficit/hyperactivity disorder. *CNS Spectrums*, *8*, 253–258.
Pronin, E., Gilovich, T., & Ross, L. (2004). Objectivity in the eye of the beholder: Divergent perceptions of bias in self versus others. *Psychological Review*, *111*, 781–799.
Raikes, H. A., & Thompson, R. A. (2006). Family emotional climate, attachment security, and young children's emotion knowledge in a high-risk sample. *British Journal of Developmental Psychology*, *24*, 89–104.
Reisch, T., Ebner-Priemer, U. W., Tschacher, W., Bohus, M., & Linehan, M. M. (2008). Sequences of emotions in patients with borderline personality disorder. *Acta Psychiatrica Scandinavica*, *118*, 42–48.
Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169–192.
Robbins, T. W. (1997). Arousal systems and attentional processes. *Biological Psychology*, *45*, 57–71.
Rochat, P. (2009). *Others in Mind—Fear of rejection and the social origin of self-consciousness*. Cambridge: Cambridge University Press.
Rogosch, F. A., & Cicchetti, D. (2005). Child maltreatment, attention networks, and potential precursors to borderline personality disorder. *Development and Psychopathology*, *17*, 1071–1089.
Rogosch, F. A., Cicchetti, D., & Aber, J. L. (1995). The role of child maltreatment in early deviations in cognitive and affective processing abilities and later peer relationship problems. *Development and Psychopathology*, *7*, 591–609.
Ross, C. A. (2007). Borderline personality disorder and dissociation. *Journal of Trauma and Dissociation*, *8*, 71–80.
Rusch, N., Lieb, K., Gottler, I., Hermann, C., Schramm, E., Richter, H., et al. (2007). Shame and implicit self-concept in women with borderline personality disorder. *American Journal of Psychiatry*, *164*, 500–508.
Sabbagh, M. A. (2004). Understanding orbitofrontal contributions to theory-of-mind reasoning: Implications for autism. *Brain Cognition*, *55*, 209–219.
Sanislow, C. A., Grilo, C. M., Morey, L. C., Bender, D. S., Skodol, A. E., Gunderson, J. G., et al. (2002). Confirmatory factor analysis of *DSM-IV* criteria for borderline personality disorder: Findings from the collaborative longitudinal personality disorders study. *American Journal of Psychiatry*, *159*, 284–290.
Sar, V., Akyuz, G., Kugu, N., Ozturk, E., & Ertem-Vehid, H. (2006). Axis I dissociative disorder comorbidity in borderline personality disorder and reports of childhood trauma. *Journal of Clinical Psychiatry*, *67*, 1583–1590.
Satpute, A. B., & Lieberman, M. D. (2006). Integrating automatic and controlled processes into neurocognitive models of social cognition. *Brain Research*, *1079*, 86–97.
Schmahl, C. G., Elzinga, B. M., Ebner, U. W., Simms, T., Sanislow, C., Vermetten, E., et al. (2004). Psychophysiological reactivity to traumatic and abandonment scripts in borderline personality and posttraumatic stress disorders: A preliminary report. *Psychiatry Research*, *126*, 33–42.
Schnell, K., Dietrich, T., Schnitker, R., Daumann, J., & Herpertz, S. C. (2007). Processing of autobiographical memory retrieval cues in borderline personality disorder. *Journal of Affective Disorders*, *97*, 253–259.
Shamay-Tsoory, S. G., & Aharon-Peretz, J. (2007). Dissociable prefrontal networks for cognitive and affective theory of mind: A lesion study. *Neuropsychologia*, *45*, 3054–3067.
Shamay-Tsoory, S. G., Aharon-Peretz, J., & Levkovitz, Y. (2007). The neuroanatomical basis of affective mentalizing in schizophrenia: Comparison of patients with schizophrenia and patients with localized prefrontal lesions. *Schizophrenia Research*, *90*, 274–283.
Shamay-Tsoory, S. G., Aharon-Peretz, J., & Perry, D. (2009). Two systems for empathy: A double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain*, *132*, 617–627.
Sharp, C., & Fonagy, P. (2008a). The parent's capacity to treat the child as a psychological agent: Constructs, measures and implications for developmental psychopathology. *Social Development*, *17*, 737–754.
Sharp, C., & Fonagy, P. (2008b). Social cognition and attachment-related disorders. In C. Sharp, P. Fonagy & I. Goodyer (Eds.), *Social cognition and developmental psychopathology*. Oxford: Oxford University Press.
Sharp, C., Fonagy, P., & Goodyer, I. (Eds.). (2008). *Social cognition and developmental psychopathology*. Oxford: Oxford University Press.

Sharp, C., & Romero, C. (2007). Borderline personality disorder: A comparison between children and adults. *Bulletin of the Menninger Clinic*, *71*, 85–114.

Shipman, K., Edwards, A., Brown, A., Swisher, L., & Jennings, E. (2005). Managing emotion in a maltreating context: A pilot study examining child neglect. *Child Abuse and Neglect*, *29*, 1015–1029.

Shmueli-Goetz, Y., Target, M., Fonagy, P., & Datta, A. (2008). The Child Attachment Interview: A psychometric study of reliability and validity. *Developmental Psychology*, *44*, 939–956.

Silbersweig, D., Clarkin, J. F., Goldstein, M., Kernberg, O. F., Tuescher, O., Levy, K. N., et al. (2007). Failure of frontolimbic inhibitory function in the context of negative emotion in borderline personality disorder. *American Journal of Psychiatry*, *164*, 1832–1841.

Slade, A. (2005). Parental reflective functioning: An introduction. *Attachment and Human Development*, *7*, 269–281.

Smith, M., & Walden, T. (1999). Understanding feelings and coping with emotional situations: A comparison of maltreated and nonmaltreated preschoolers. *Social Development*, *8*, 93–116.

Soloff, P. H., Feske, U., & Fabio, A. (2008). Mediators of the relationship between childhood sexual abuse and suicidal behavior in borderline personality disorder. *Journal of Personality Disorders*, *22*, 221–232.

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

Stiglmayr, C. E., Ebner-Priemer, U. W., Bretz, J., Behm, R., Mohse, M., Lammers, C. H., et al. (2008). Dissociative symptoms are positively related to stress in borderline personality disorder. *Acta Psychiatrica Scandinavica*, *117*, 139–147.

Strathearn, L., Fonagy, P., Amico, J., & Montague, R. (in press). Adult attachment predicts maternal brain and oxytocin response to infant cues. *Neuropsychopharmacology*.

Strathearn, L., Li, J., Fonagy, P., & Montague, R. (2008). What's in a smile? Maternal brain responses to infant facial cues. *Pediatrics*, *122*, 40–51.

Stuart, J., Westen, D., Lohr, N. E., Benjamin, J., Becker, S., & Vorus, N. (1990). Object relations in borderlines, depressives, and normals: An examination of human responses on the Rorschach. *Journal of Personality Assessment*, *55*, 296–318.

Target, M., & Fonagy, P. (1996). Playing with reality II: The development of psychic reality from a theoretical perspective. *International Journal of Psycho-Analysis*, *77*, 459–479.

Uddin, L. Q., Iacoboni, M., Lange, C., & Keenan, J. P. (2007). The self and social cognition: the role of cortical midline structures and mirror neurons. *Trends in Cognitive Science*, *11*, 153–157.

van IJzendoorn, M. H. & Bakermans-Kranenburg, M. J. (1997). Intergenerational transmission of attachment: A move to the contextual level. In L. Atkinson & K. J. Zucker (Eds.), *Attachment and psychopathology* (pp. 135–170). New York: Guilford Press.

Vanderwal, T., Hunyadi, E., Grupe, D. W., Connors, C. M., & Schultz, R. T. (2008). Self, mother and abstract other: An fMRI study of reflective social processing. *NeuroImage*, *41*, 1437–1446.

Vermote, R., Fonagy, P., Vertommen, H., Verhaest, Y., Stroobants, R., Vandeneede, B., et al. (2009). Outcome and outcome trajectories of personality disordered patients during and after a psychoanalytic hospitalization-based treatment. *Journal of Personality Disorder*, *23*, 293–306.

Vorria, P., Papaligoura, Z., Sarafidou, J., Kopakaki, M., Dunn, J., Van Ijzendoorn, M. H., et al. (2006). The development of adopted children after institutional care: A follow-up study. *Journal of Child Psychology and Psychiatry*, *47*, 1246–1253.

Vrticka, P., Andersson, F., Grandjean, D., Sander, D., & Vuilleumier, P. (2008). Individual attachment style modulates human amygdala and striatum activation during social appraisal. *PLoS ONE*, *3*, e2868.

Walter, M., Bureau, J. F., Holmes, B. M., Bertha, E. A., Hollander, M., Wheelis, J., et al. (2008). Cortisol response to interpersonal stress in young adults with borderline personality disorder: A pilot study. *European Psychiatry*, *23*, 201–204.

Watson, S., Chilton, R., Fairchild, H., & Whewell, P. (2006). Association between childhood trauma and dissociation among patients with borderline personality disorder. *Australian and New Zealand Journal of Psychiatry*, *40*, 478–481.

Wertz, A. E., & German, T. C. (2007). Belief-desire reasoning in the explanation of behavior: Do actions speak louder than words? *Cognition*, *105*, 184–194.

Westen, D., Dutra, L., & Shedler, J. (2005). Assessing adolescent personality pathology. *British Journal of Psychiatry*, *186*, 227–238.

Westen, D., Ludolph, P., Lerner, H., Ruffins, S., & Wiss, F. C. (1990). Object relations in borderline adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry*, *29*, 338–348.

Williams, L., Fonagy, P., Target, M., Fearon, P., Sargent, J., Bleiberg, E., et al. (2006). Training psychiatry residents in mentalization-based therapy. In J. Allen & P. Fonagy (Eds.) *Handbook of mentalisation based treatments*. London: Wiley.

Wingenfeld, K., Lange, W., Wulff, H., Berea, C., Beblo, T., Saavedra, A. S., et al. (2007). Stability of the dexamethasone suppression test in borderline personality disorder with and without comorbid PTSD: A one-year follow-up study. *Journal of Clinical Psychology*, *63*, 843–850.

Zanarini, M., Gunderson, J. G., & Frankenburg, F. R. (1990). Discriminating borderline personality disorder from other Axis II disorders. *American Journal Psychiatry*, *147*, 161–167.

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

Zanarini, M. C., Frankenburg, F. R., DeLuca, C. J., Hennen, J., Khera, G. S., & Gunderson, J. G. (1998). The pain of being borderline: Dysphoric states specific to borderline personality disorder. *Harvard Review of Psychiatry*, *6*, 201–207.

Zanarini, M. C., Frankenburg, F. R., Jager-Hyman, S., Reich, D. B., & Fitzmaurice, G. (2008). The course of dissociation for patients with borderline personality disorder and Axis II comparison subjects: A 10-year follow-up study. *Acta Psychiatrica Scandinavica*, *118*, 291–296.

Zanarini, M. C., Frankenburg, F. R., Ridolfi, M. E., Jager-Hyman, S., Hennen, J., & Gunderson, J. G. (2006). Reported childhood onset of self-mutilation among borderline patients. *Journal of Personality Disorders*, *20*, 9–15.

Zittel Conklin, C., & Westen, D. (2005). Borderline personality disorder in clinical practice. *American Journal of Psychiatry*, *162*, 867–875.

Zweig-Frank, H., & Paris, J. (1991). Parents' emotional neglect and overprotection according to the recollections of patients with borderline personality disorder. *American Journal of Psychiatry*, *148*, 648–651.