How childhood maltreatment alters perception and cognition – the predictive processing account of borderline personality disorder

Philipp Herzog1,2,3, Tobias Kube3 and Eva Fassbinder2

1Department of Psychiatry and Psychotherapy, University of Lübeck, Ratzeburger Allee 160, D-23562 Lübeck, Germany; 2Department of Psychiatry and Psychotherapy, Christian-Albrechts-University of Kiel, Niemannsweg 147, D-24105 Kiel, Germany and 3Department of Psychology, University of Koblenz-Landau, Ostbahnstr. 10, 76829 Landau, Germany

Abstract

Borderline personality disorder (BPD) is a severe mental disorder, comprised of heterogeneous psychological and neurobiological pathologies. Here, we propose a predictive processing (PP) account of BPD to integrate these seemingly unrelated pathologies. In particular, we argue that the experience of childhood maltreatment, which is highly prevalent in BPD, leaves a developmental legacy with two facets: first, a coarse-grained, alexithymic model of self and others – leading to a rigidity and inflexibility concerning beliefs about self and others. Second, this developmental legacy leads to a loss of confidence or precision afforded beliefs about the consequences of social behavior. This results in an over reliance on sensory evidence and social feedback, with concomitant lability, impulsivity and hypersensitivity. In terms of PP, people with BPD show a distorted belief updating in response to new information with two opposing manifestations: rapid changes in beliefs and a lack of belief updating despite disconfirmatory evidence. This account of distorted information processing has the potential to explain both the instability (of affect, self-image, and interpersonal relationships) and the rigidity (of beliefs about self and others) which is typical of BPD. At the neurobiological level, we propose that enhanced levels of dopamine are associated with the increased integration of negative social feedback, and we also discuss the hypothesis of an impaired inhibitory control of the prefrontal cortex in the processing of negative social information. Our account may provide a new understanding not only of the clinical aspects of BPD, but also a unifying theory of the corresponding neurobiological pathways. We conclude by outlining some directions for future research on the behavioral, neurobiological, and computational underpinnings of this model, and point to some clinical implications of it.

Introduction

Borderline personality disorder (BPD) is a severe and heterogeneous mental disorder, characterized by a pervasive pattern of instability of interpersonal relationships, self-image, affects, and a marked impulsivity (Chapman, 2019). Despite recent advances in treating BPD (Storebø et al., 2020), an analysis of the effectiveness in routine clinical care for BPD pointed to the need of improving the reduction of BPD-specific symptoms (Herzog et al., 2020). To achieve this goal, a precise understanding of the psychobiological processes underlying BPD is required. Although a genetic vulnerability in BPD is unquestionable, genetics alone may have little influence on the development of BPD in a favorable environment as a large part of the explained variance is attributable to individually unique environmental factors (Skoglund et al., 2021). Hence, the genetic vulnerability may translate into disordered perceptions and cognitions in the presence of adverse environmental conditions (e.g. childhood maltreatment, CM), that, in turn, likely increase the risk for developing a BPD (Fontaine & Viding, 2008). To improve the understanding of these psychobiological processes in BPD, the present article seeks to apply current theories from neuroscience to provide a novel mechanistic model of BPD. In particular, we use a predictive processing (PP) framework to explain how the experience of CM, which is common in many people with BPD (Kleindienst, Vonderlin, Bohus, & Lis, 2020; Porter et al., 2020), alters perception and cognition, resulting in distinct psychobiological dysfunctions.

Predictive processing – an impactful theory of how the brain works

Considering the brain as a ‘phantastic', hypothesis-testing organ (Friston, Stephan, Montague, & Dolan, 2014), PP has become a prominent theory of fundamental working principles of the human brain (Clark, 2013, 2016a; Friston, 2005; Hohwy, 2014). Specifically, it has been theorized that the brain relies on principles of Bayesian inference to
minimize uncertainty by continuously testing hypotheses regarding the causes of sensory input. In doing so, a prior belief is combined with new information to compute a posterior belief. If new information critically deviates from the prior prediction, a prediction error (PE) is generated. The (healthy) brain seeks to refine its hypotheses and maximize evidence for its internal model of the world by minimizing PEs.

At the neurobiological level, PP has been conceived of in terms of cerebral hierarchies (Friston, 2008). Specifically, in neural representations of higher levels of cortical hierarchies, predictions are generated (encoded by synaptic activity), which then descend to lower levels. In superficial pyramidal cells, such descending top-down predictions are compared with neural representations at lower levels of the cortical hierarchy to compute a PE (Bastos et al., 2012). The PE (i.e. a mismatch signal) is sent back up the hierarchy where it updates prior beliefs or expectations (that generate top-down predictions), associated with the activity of deep pyramidal cells (Kanai, Komura, Shipp, & Friston, 2015). For a more detailed discussion of the neural architecture of PP, see Barrett and Simmons (2015), Parr and Friston (2018), and Shipp (2016).

PP models also account for the outcomes of action, with the outcomes of actions being predicted based on causal knowledge (Heil, Kwishtout, van Pelt, van Rooij, & Bekkering, 2018). The human brain chooses the actions that are expected to produce the most preferred outcome or provides the most salient information. Briefly, this process suggests that such action-perception cycles operate to minimize uncertainty and optimize an individual’s internal model of the world – a process referred to as active inference. Since perception in active inference is treated as a constructive process of hypothesis testing, where the brain aims to select the hypothesis that best explains sensory data, the brain must decide how much weight is given to new information relative to prior beliefs. This is referred to as precision-weighting, with precision being defined as ‘the certainty with which a model is believed to be true and the certainty of a particular afferent given an expectation’ (Paulus, Feinstein, & Khalsa, 2019). That is, precision can refer to both the reliability of new information and the confidence afforded to priors. Their balance critically determines the extent to which a prior is updated given new information. Put simply, if the prior is afforded low precision (referred to as ‘weak priors’), new information has more influence on the formation of the posterior, while the opposite is true for ‘strong priors’. In other words, if we are unsure about how much we can trust our beliefs, we prefer to rely on new information, provided that new information appears sufficiently valid. On the other hand, if priors are afforded overly much precision, they dominate perception such that information is consistent with prior predictions is prioritized and discrepant information is largely neglected (Powers, Mathys, & Corlett, 2017).

Neurobiologically, the precision of sensory information is thought to be signaled by neuromodulators such as dopamine and encoded based on synaptic gain control mechanisms (Fiorillo, Newsome, & Schultz, 2008; Galea, Bestmann, Beigi, Jahanshahi, & Rothwell, 2012; Iglesias et al., 2013). Aberrations in precision-weighting have recently been related to a number of psychopathological dysfunctions and mental disorders, such as depression (Barrett, Quigley, & Hamilton, 2016; Clark, Watson, & Friston, 2018; Kube, Schwarting, Rozenkrantz, Glombiewski, & Rief, 2020), stressors and psychological trauma (Krupnik, 2020; Linson, Parr, & Friston, 2020), PTSD (Kube, Berg, Klein, & Herzog, 2020a; Linson & Friston, 2019; Wilkinson, Dodgson, & Meares, 2017), hallucinations in psychosis (Corlett et al., 2019; Sterzer et al., 2018) and in PTSD (Lyndon & Corlett, 2020), autism (Lawson, Rees, & Friston, 2014; Pellicano & Burr, 2012), and somatization (Henningsen et al., 2018; Kube, Rozenkrantz, Rief, & Barsky, 2020b; van den Bergh, Witthöft, Petersen, & Brown, 2017). The belief updating process is illustrated in Fig. 1.

In the present article, we apply this account to BPD, proposing that BPD is related to an imbalance of hierarchical priors – specifically, imprecise prior beliefs, relative to sensory evidence – based upon a learning history of CM. Before we will lay out the central tenets of this account, we first briefly review the literature on normal infant learning from a PP perspective to understand how adverse environmental conditions (that is, CM) perturbs the perceptual system, which can ultimately result in psychopathological dysfunctions as manifested in BPD.

Normal infant learning

Research has shown that PP may provide a unifying perspective on infant learning, including statistical learning principles, motor and proprioceptive learning, and developing a basic understanding of the self and their physical and social environment (Köster, Kayhan, Langeloh, & Hoehl, 2020). Using Violation-of-Expectation paradigms (Sokolov, 1963, 1990), research has shown that infants use novel and unexpected experiences (i.e. PEs) to refine their predictive models of the world, as indicated e.g. in studies using event-related potential technique (Köster, Langeloh, & Hoehl, 2019; Langeloh, Buttelmann, Pauen, & Hoehl, 2020). Indeed, recent research has emphasized that infants build and update their predictive models of a changing environment early in their lives (Kayhan, Meyer, O’Reilly, Hunnius, & Bekkering, 2019), suggesting that reducing uncertainty in a changing world is an important developmental goal.

Childhood maltreatment in BPD

Learning from PEs in infancy can be hindered or disrupted by a harmful environment. For instance, a longitudinal study revealed that the experience of attachment disorganization and parental hostility in early childhood was associated with dysfunctions in several domains (i.e. attention, emotion, behavior, relationship, and self-representation) in middle childhood/early adolescence and, ultimately, symptoms of BPD in adulthood (Carlson, Egeland, & Sroufe, 2009). Harmful learning experiences in childhood, such as neglect and abuse, have been referred to as CM. CM constitutes as a transdiagnostic risk factor for mental disorders (Gilbert et al., 2009; Struck et al., 2020) and plays an important role in the development of personality disorders in general (Afifi et al., 2011; Battle et al., 2004) and BPD in particular (Brakemeier et al., 2018; Pietrek, Elbert, Weierstall, Müller, & Rockstroh, 2013; Quenneville et al., 2020; Wota et al., 2014). The majority of patients with BPD experienced some sort of maltreatment in their childhood (Battle et al., 2004; Carlson et al., 2009): A recent study showed that patients with BPD were over 13 times more likely to report childhood adversity than non-clinical controls and other clinical populations (Kleindienst et al., 2020; Porter et al., 2020). Among the most frequent subtypes of CM in BPD are emotional abuse and neglect compared to controls (Scheffers, van Vught, Lancôt, & Lemieux, 2019). While studies showed that BPD was associated with higher scores on sexual abuse besides of emotional abuse and neglect (Bradley, Jenei, & Westen, 2005; Hernandez, Arntz, Gaviria, Labad, & Gutiérrez-Zotes, 2012; Igarashi et al., 2010; Llobbestael, Arntz, &
Bernstein, 2010; Ogata et al., 1990; Weaver & Clum, 1993; Zhang et al., 2013), particular in women (de Aquino Ferreira, Queiroz Pereira, Neri Benevides, & Aguiar Melo, 2018), other studies indicated no independent relationship between sexual abuse and BPD (Cohen et al., 2014; Widom, Czaja, & Paris, 2009) or reported small effect sizes for sexual abuse (Hengartner, Ajdacic-Gross, Rodgers, Müller, & Rössler, 2013). In line, only emotional instability or vulnerability, impulsivity, and emotional abuse were found to be unique predictors of BPD (Bornovalova, Gratz, Delany-Brumsey, Paulson, & Lejuez, 2006). Moreover, social exclusion (ostracism) might be a psychosocial factor contributing to the development and persistence of BPD, in the sense of a vicious cycle where BPD increases the chance of being ostracized, and ostracism consolidates or even aggravates psychopathology (Reinhard et al., 2020). In line, only emotional instability or vulnerability, impulsivity, and emotional abuse were found to be unique predictors of BPD (Bornovalova, Gratz, Delany-Brumsey, Paulson, & Lejuez, 2006). Moreover, social exclusion (ostracism) might be a psychosocial factor contributing to the development and persistence of BPD, in the sense of a vicious cycle where BPD increases the chance of being ostracized, and ostracism consolidates or even aggravates psychopathology (Reinhard et al., 2020). For example, bullying and violence in schools and emotional abuse appear to be more salient markers of general personality pathology than other forms of childhood adversity (Hengartner et al., 2013). Insights from a sibling design showed that both probands and sisters reported similar prevalence of intrafamilial abuse, although later BPD patients reported more severe intrafamilial physical and emotional abuse, and higher prevalence of physical abuse by peers (Laporte, Paris, Guttman, Russell, & Correa, 2012). Furthermore, another study that prospectively followed a sample who had experienced childhood abuse showed an increase in risk for BPD primarily in children who experienced physical abuse and neglect (Widom et al., 2009). In particular, terrorizing predicted anxiety and somatic concerns, ignoring predicted depression scores and BPD features, and degradation predicted BPD features only (Allen, 2008).

**Psychological consequences of childhood maltreatment**

Co-regulation and social communication in infancy are thought to underpin emotional dysregulation and social cognition deficits across development and these mechanisms are further potentiated by maladaptive social experiences in a series of positive feedback loops (Winsper, 2018; Winsper et al., 2016). Interestingly, studies suggest an association between CM, especially emotional abuse and neglect, and emotion regulation difficulties in a way that emotion regulation difficulties influence the association between emotional abuse and acute symptomatology in BPD (e.g. Carvalho Fernando et al., 2014), supporting the Emotional Dysregulation theory (Linehan, 1993). While CM impacts
impulsivity and anger (Quenenneville et al., 2020), one study showed that difficulties in emotion regulation statistically mediated the effect of CM on impulsivity in BPD (Krause-Utz et al., 2019), further supporting this theory.

Consistent with these findings, a study showed a moderate relationship between low emotional awareness (EA) (especially difficulties in identifying and describing emotions) and BPD (Derks, Westerhof, & Bohlmeijer, 2017). Of note, EA has recently been conceptualized within an active inference model: the authors showed that it can successfully acquire a repertoire of emotion concepts in its ‘childhood’, as well as acquire new emotion concepts in synthetic ‘adulthood’, and that these learning processes depend on early experiences, environmental stability, and habitual patterns of selective attention (Smith, Parr, & Friston, 2019a). Distinct neurocomputational processes underlying EA have further been developed, e.g. mechanisms that (either alone or in combination) can produce phenomena – such as somatic misattribution, coarse-grained emotion conceptualization, and constrained reflective capacity – characteristic of low EA (Smith, Lane, Parr, & Friston, 2019a). Relatedly, while one study found no specific association between parenting style with BPD (Hernández et al., 2012), a meta-synthesis study found that maladaptive parenting is a well-established psychosocial risk factor for the development of BPD (Steele, Townsend, & Grenyer, 2019). A prominent study showed that family environment, parental psychopathology, and history of abuse all independently predicted BPD (Bradley et al., 2005).

Integrating the environmental factors from a lifespan perspective, literature suggests that vulnerability from mother to offspring may be partly transmitted via maladaptive parenting and maternal emotional dysfunction, i.e. mothers with BPD are more likely to engage in maladaptive interactions with their offspring characterized by insensitive, overprotective, and hostile parenting compared to mothers without BPD resulting in adverse offspring outcomes such as BPD symptoms, internalizing (e.g. depression) and externalizing problems, insecure attachment patterns, and emotional dysregulation (Eyden, Winsper, Wolke, Broome, & MacCallum, 2016). Importantly, CM and its link to BPD features exist already in children (Ibrahim, Cosgrave, & Woolgar, 2018). One important factor might be negative emotional reactivity that seems to be a marker of vulnerability that increases the risk for the development of BPD (Stepp, Scott, Jones, Whalen, & Hipwell, 2015). In the supplement, we also provide an extensive review of the psychological consequences of CM on social cognition and their implications for treatment (see online Supplemental Material 1). Here, we want to focus on how the effects of CM in individuals with BPD can be conceived of in PP terms.

**Childhood maltreatment and predictive processing**

In brief, we are proposing a developmental active inference and learning account of BPD that can be summarized as follows: Early traumatic experiences (i.e. CM) – particularly those involving disorganized attachment or emotional abuse/neglect – lead to an impoverished model of self and others, under which the consequences of social behavior are unpredictable. This has two consequences:

First, a loss of confidence or precision when selecting the course of action in social exchanges. This irreducible uncertainty – about ‘what would happen if I did that?’ – underwrites the emotional lability and impulsivity, characteristic of BDP. The implicit loss of precision, afforded prior beliefs about social narratives, renders belief updating overly sensitive to sensory evidence and social feedback.

To illustrate this key aspect of our argument, consider the example of a child being sad and crying because it feels alone, with the mother invalidating the child’s experience by saying there is no reason to be sad and leaving the room. In this case, the child experiences a lack of understanding and feel that their emotions and thoughts are called into question by their parents or significant others (Musser, Zalewski, Stepp, & Lewis, 2018), resulting in an impoverished model of the self in form of low emotional awareness (especially difficulties in identifying and describing emotions) (Smith et al., 2019b; Smith et al., 2022) that has been linked to BPD (Derks et al., 2017). Relatedly, early risk factors for BPD in adulthood include the maternal withdrawal in infancy and separation of 1 month or more from the mother in the first 5 years of life (Steele & Siever, 2010). Furthermore, an interesting candidate for a specific parent-child-relationship risk factor for BPD is parental inconsistency (Boucher et al., 2017), leading to the child’s perception that other people’s behavior is unpredictable. If such invalidating situations are frequently experienced, it is understandable that the child does not learn to place confidence in their beliefs and thus remains uncertain about the causes of their sensations. Put another way, such repeated experiences leave the child with the interpretation that their thoughts and emotions are ‘wrong’, leading in turn to increased uncertainty about their beliefs and perceptions that results in dysfunctional behavior (e.g. ineffective emotion regulation strategies). Indeed, a recent study indicated that BPD, compared to other mental disorders, is associated with a less frequent use of effective emotion regulation strategies (i.e. cognitive reappraisal, problem solving, and acceptance) and a more frequent use of dysfunctional emotion regulation strategies (i.e. suppression, rumination, and avoidance) (Daros & Williams, 2019).†

Second, the failure to learn a suitably expressive generative model, during neurodevelopment, leaves patients with BDP with impoverished, coarse-grained models of self and other. The ensuing interpersonal ‘alexithymia’ manifests as rigidity and inflexibility in social or affective interactions. In other words, everything is ‘black or white’, with no ‘shades of gray’ that would support a nuanced inference about the intention of others (and self). In the absence of an expressive social world model, the only explanations available – for unpredicted sensory or social outcomes – are ‘I am worthless’ (i.e. self-critical explanations) or ‘you are punitive’ (i.e. paranoid explanations).

This developmental account rests upon the intimate relationship between learning and inference. Here, a failure to learn a sufficiently expressive model of interpersonal narratives precludes prosocial inference and planning that precludes subsequent learning or updating of self-other models. This is not unlike some accounts of severe autism and the failure to attain central coherence (Happé & Frith, 2006; van de Cruys et al., 2014). In autism, this developmental failure is sometimes attributed to failure of sensory attenuation (i.e. a failure to attenuate sensory precision in relation to prior precision). This suggests interesting parallels between self-stimulation in autism and non-fatal self-harm in BPD.

†The notes appear after the main text.
In terms of the neurochemistry of uncertainty or precision encoding in the brain, our analysis speaks to the same kind of hyperdopaminergic state that may characterize certain forms of schizophrenia or delusional disorders. In our case, this can be read directly from the role of dopamine in active inference, as scoring the resolution of uncertainty about policies afforded by sensory evidence: in BPD, every piece of sensory (or social) information resolves uncertainty. This is because beliefs about the narrative currently being pursued are always uncertain or imprecise. The analogy with schizophrenia here may be useful in the sense that some people interpret delusions as an attempt to make sense of unattenuated sensory evidence (Sterzer et al., 2018).

The predictive processing account of BPD

While most models of BPD take a lifespan approach and consider the complex interplay of biological vulnerabilities (e.g. genetics), psychological factors and social influences (Stepp, Lazarus, & Byrd, 2016), BPD is currently viewed as a disorder of instability, i.e. instability of interpersonal relationships, self-image, affects, and, relatedly, a marked impulsivity (American Psychiatric Association, 2013). Here, we recast these core clinical aspects of BPD through the lens of a PP model (see Fig. 2). In doing so, we build upon a few previous accounts offering computational perspectives on BPD (Fineberg, 2019; Fineberg, Stahl, & Corlett, 2017; Fineberg, Steinfeld, Brewer, & Corlett, 2014b). For example, in an elaborated approach, Fineberg et al. (2017) put emphasis on the association between early disruption of mothers’ physical care and social dysfunction as a key feature in BPD, explained by social learning depending on reinforcement learning though embodied simulations.

Affective instability

BPD is characterized by an affective instability, in terms of rapid changes of mood, ranging from chronic feelings of emptiness to intense anger, anxiety, or dysphoria – and difficulties in handling these mood fluctuations. From a PP point of view, emotional states have been theorized to be a consequence of acting in the world with the aim of minimizing expected free energy, that is, the uncertainty about the future consequences of actions (Kiverstein, Miller, & Rietveld, 2020). In other words, emotional states reflect changes in the uncertainty about the somatic consequences of action (Joffily & Coricelli, 2013; Seth & Friston, 2016; Wager et al., 2015), with uncertainty relating to the precision with which motor and physiological states are predicted and inferred (Clark et al., 2018). By this view, painful emotions have been thought to accompany events that elicit beliefs of unpredictability, whereas pleasant emotions refer to events that resolve uncertainty and convey a sense of control and confidence (Barrett & Satpute, 2013; Clark et al., 2018; Gu, Hof, Friston, & Fan, 2013). Thus, the valence of emotional states relates to the resolution of uncertainty and the precision with which the consequences of action are predicted (Brown & Friston, 2012; Clark et al., 2018). Interestingly, BPD features predicted specific patterns of bias with regard to forecasting future (negative) emotional states (Hughes & Rizvi, 2019).

Linking this account with the corollaries of CM as discussed above, we suggest that BPD is related to the failure to resolve uncertainty and the dominance of predictions of unpredictability, resulting in intense painful emotions. Relatedly, we propose that people with BPD afford their beliefs about the consequences of action little precision; as a result, interoceptive experience is largely influenced by sensory evidence. Hence, the exact emotional experience may vary greatly with respect to the nature and the source of sensory information. In other words, because people with BPD have learned to be uncertain about their consequences of action, their emotional experience fluctuates significantly depending on the (perceived) situational circumstances and other people’s behavior. For example, as people with BPD have often grown up in an invalidating environment, they may have learnt that experiencing a certain emotion (e.g. sadness) and expressing a corresponding emotional need (e.g. being consoled by their parents) has no influence on their parents’ actual response (e.g. neglect perceived as a negative consequence). Thus, they cannot resolve uncertainty and experience fluctuating intense emotions, by likewise generating a distorted internal model of emotions (e.g. no access to the concept of being consoled) and a reduced self-efficacy. In addition to the uncertainty inherent to beliefs of people with BPD, the experience of CM and the deficit in adaptive emotion regulation strategies may lead to the pervasive belief that emotions are uncontrollable. Due to the high precision afforded to it, this belief is difficult to be revised through new experiences.

Self-image instability

Besides affective instability, BPD is characterized by identity instability which leads to a markedly and persistently unstable, yet often negative, self-concept as well as a lack of a self-coherence. This cognitive instability is related to uncertainty in at least two of the following life domains: self-image, sexual orientation, long-term goals or career choice, type of friends desired, and values. From a PP point of view, this self-image instability manifests through low precision that is afforded to prior beliefs. As a result, beliefs are hastily updated in line with novel information. In other words, because the beliefs of people with BPD – about both the self and others – are fraught with uncertainty, they are often updated based on fairly thin evidence. This account is well in line with evidence from experimental studies, particularly research that has examined how people with BPD update their beliefs in response to social feedback. For instance, Korn, la Rosée, Heekeren, & Roepke (2016a) found that people with BPD, in contrast to healthy people (Korn, Prehn, Park, Walter, & Heekeren, 2012), adjusted their beliefs about themselves significantly in line with undesirable social feedback, that is, a single negative interpersonal experience (Korn et al., 2012, 2016a). In line with that, Liebke et al. (2018) showed in a virtual group-interaction paradigm (where participants interacted with a group of computer-controlled avatars, although they believed them to be real human co-players) that people with BPD updated their beliefs about being socially accepted in response to negative but not positive social feedback (Liebke et al., 2018). Interestingly, the authors demonstrated that people with BPD even behaved less cooperatively in a subsequent trust game if they had previously received positive social feedback. Similarly, another study found that whereas healthy people focus on the positive aspects of social feedback, people with BPD focus more on negative feedback, thereby maintaining negative self-views (Van Schie, Chiu, Rombouts, Heiser, & Elzinga, 2020). This volatility of belief updating – with more adjustments in response to negative feedback – can again be linked to the effects of early learning experiences and CM. In particular, children who experienced CM often receive negative social feedback, e.g.
parents’ scolding, which children seek to integrate in their attempt to satisfy their parents. When unexpectedly receiving positive feedback, children might be confused about how to interpret that, thus increasing the uncertainty in the children’s perceptions and predictions. In computational terms, the world (i.e., the social feedback by significant others) seems unpredictable as the child is permanently exposed to PEs. Given the prevailing negative social feedback, though, people with CM learn to afford higher precision to negative than to positive feedback, while at the same time affording high precision to negative prior beliefs about the self. Both contributes to the above-described asymmetry in belief updating in BPD, with more rapid changes in line with novel negative information and the persistence of strong negative beliefs about the self. This affects also social decision making and learning in a way that BPD patients expect higher volatility than control which underpins social and non-social belief updating in BPD (Fineberg et al., 2018b).

**Instability of interpersonal relationships**

As a consequence of self-image instability and impaired social decision making, BPD is also associated with a pattern of unstable and intense interpersonal relationships, while alternating between extremes in the perception of other people, i.e., idealization and devaluation. For example, patients’ perceptions of their therapist can change rapidly, from beliefs such as ‘My therapist is the only person who has ever understood me’ to ‘My therapist wants to get rid of me’. From a PP perspective, idealization and devaluation is related to hasty changes of beliefs about other people, resulting from weak priors and high precision afforded to new information. This instability likely leads to difficulties in interpersonal relationships, with rapid alternations of desperate efforts to avoid abandonment and a premature relationship termination. In line with this notion, research has shown that BPD patients exhibit less trust during interpersonal interactions (Unoka, Seres, Áspán, Bódi, & Kéri, 2009), reflecting the high degree of uncertainty people with BPD experience in social interactions.

In a transdiagnostic clinical sample, it has been shown that a reduced confidence in how to act, rather than increased emotional conflicts, explains maladaptive approach-avoidance behaviors, i.e., a greater decision uncertainty during approach-avoidance conflicts (Smith et al., 2020). Consistent with that, a recent study found a 2-fold larger preferred interpersonal distance in BPD patients than in healthy people (Fineberg et al., 2018a).

In addition to instability in their relationships, people with BPD also form strong negative beliefs about relations to other people in their attempt to reduce uncertainty, as touched upon above. This rigidity in beliefs about other people may underlie another symptom cluster of BPD: transient, stress-related paranoid ideation (i.e., the belief that harm is intended by others). In line with this account, a recent study found that paranoia may make it harder to update beliefs and it is linked with an increased risk of violence towards oneself (Reed et al., 2020). More precisely, this study showed that uncertainty may be sufficient to elicit learning differences in paranoid individuals, even without social threat; and paranoia is associated with a stronger prior on volatility, accompanied by elevated sensitivity to perceived changes in the task environment (Reed et al., 2020). In other words, people with paranoia expect the world to change frequently, change their minds repeatedly, and have a harder time learning in response to changing circumstances. This finding is in line with the general high uncertainty of BPD patients and may explain paranoid ideation in BPD, further highlighting the interdependence between inference and learning in BPD. That
is, BPD patients with higher scores on paranoid ideations might have more difficulty updating their beliefs in interpersonal situations (i.e. learning difficulties in response to changing circumstances), leading to more interpersonal problems. In fact, difficulties with trusting others and volatile impressions of others’ moral character are often problems that result in the premature termination of a relationship. Of note, the moral inference differed: In patients with BPD, beliefs about harmful agents were more certain and less amenable to updating relative to healthy controls (Siegel, Curwell-Parry, Pearce, Saunders, & Crockett, 2020). For instance, when interacting with therapists within a mental health care setting, one specific focus is to build a strong therapeutic alliance. Conceivably, there are two possible outcomes: a weak therapeutic alliance will probably lead to premature treatment discontinuation – a common problem in the treatment of BPD (Barnicot, Katsakou, Marougka, & Priebe, 2011). On the other hand, in the case of a strong therapeutic alliance, patients strive to avoid changes in the therapeutic setting (e.g. changing therapists due to leaves) as they often distrust other therapists – despite previous positive experiences and the fact that other therapists will probably also be kind to them. As such, this example highlights once more the lack of a suitably expressive social world model and the use of an impoverished, coarse-grained model of self and other that leads to rigidity and inflexibility in social interactions preventing a more nuanced inference about the intention of others. In line with that notion, interpersonal functioning was found to predict non-delusional paranoia in BPD (Oliva, Dalmotto, Pirfo, Furlan, & Picci, 2014) – some type of non-delusional paranoia was reported by 87% in a BPD sample (Zanarini, Frankenbug, Wedig, & Fitzmaurice, 2013).

**Impulsivity**

BPD is also characterized by a variety of impulsive behaviors, such as promiscuity and substance abuse. The most prominent example of such behaviors, which aim at reducing emotional tension, is non-suicidal self-injury (NSSI). According to PP, such impulsive behaviors may seek to resolve uncertainty: if intense aversive emotions reflect the prediction of unpredictability (Clark et al., 2018), any behavior aimed at reducing such intense aversive emotional states may convey a sense of control. In other words, people with BPD have learnt that certain behaviors, such as NSSI, lower their emotional tension, thereby increasing the certainty of the consequences of action.

In our account, impulsive behavior (such as NSSI) is considered more as a consequence of affective instability. In fact, NSSI may be reinforced by its affect stabilization function (Vansteelandt et al., 2017). In line, a recent study showed that affective instability was significantly greater in adolescents engaging in NSSI, and the number of BPD criteria met was positively correlated with affective instability in the NSSI group (Santangelo et al., 2017). Particularly, higher levels of momentary negative affect predicted greater subsequent urges to self-injure, but only when self-concept clarity was low, supporting interactive effects (Scala et al., 2018). Remarkably, impulsivity as a personality trait per se is genetically influenced and heritable (Balestri, Calati, Serretti, & de Ronchi, 2014; Bevilaqua & Goldman, 2013; Bezdjian, Baker, & Tuvblad, 2011; Fineberg et al., 2014a; Khandka et al., 2014). The basic assumptions of the PP account of BPD are displayed in Fig. 2.

**Neural specification of this account**

Contemporary etiological theories of BPD assume that biological predispositions (i.e. genetic factors) are potentiated by environmental risk factors (i.e. CM). Studies estimating the heritability of the basic dimensions of personality disorders report approximately between 35% and 56% (Jang, Livesley, Vernon, & Jackson, 1996; Livesley, Jang, Jackson, & Vernon, 1993). Twin and family studies estimated a moderate to high heritability in BPD indicating a general genetic risk factor, but highlight also individually unique environmental influences (Distel et al., 2008, 2010; Skoglund et al., 2021). Indeed, results from a longitudinal discordant twin design show that there might be a genetic influence underlying the association of traumatic events with BPD, rather than BPD being directly caused by a trauma (Bornovolova et al., 2013). Likely contributing biological factors include genes linked to dopamine, serotonin, the hypothalamic-pituitary-adrenal axis, and neuropeptides (Steele & Siever, 2010). Research points to abnormalities in the dopamine system in people with BPD (Friedel, 2004; Oquendo & Mann, 2000). The efficacy of dopamine D2 receptor (DRD2) blocking antipsychotic drugs in BPD treatment also suggests the involvement of the dopamine system in the neurobiology of BPD (Nemoda et al., 2010). Moreover, dopamine transporter (DAT1) gene variants increase the risk of BPD (Amad, Ramoz, Thomas, Jardri, & Gorwood, 2014; Joyce, Stephenson, Kennedy, Mulder, & McHugh, 2013). This crucial role of dopamine is well in line with the PP account of BPD, because in PP dopamine is thought to encode the precision of beliefs that underwrite choices and behavior (Schwartenbeck, FitzGerald, Mathys, Dolan, & Friston, 2015), and dopamine has been shown to modulate belief updating (Sharot et al., 2012). Drawing on this previous work, we suggest that the increased integration of negative social feedback in BPD corresponds to enhanced levels of dopamine in the respective synaptic gains, accounting for the increased use of that information to update the prior prediction. Computationally, this is associated with high precision with which new information is encoded.

Another psychobiological mechanism that might be involved in the psychopathology of BPD refers to the connectivity of different brain areas. Specifically, the increased integration of negative social feedback in BPD might be related to a reduced inhibitory control of the prefrontal cortex (PFC) in PE processing. In particular, previous research has shown that PE processing in the reward system (i.e. ventral striatum) can be suppressed by the PFC, resulting in a lack of belief updating, as demonstrated for pain perception (Schenk, Sprenger, Onat, Colloca, & Büchel, 2017) and reward processing in depression (Greenberg et al., 2015). More specifically, research in the neuronal coding of PE (Schultz & Dickinson, 2000) has indicated that neurons in the dorsolateral prefrontal, orbitofrontal, and anterior cingulate cortex are activated in relation to errors in the prediction of reward (Niki & Watanabe, 1979; Watanabe, 1989), and these PEs are processed in combination with neurons in the striatum that code rewards relative to their unpredictability (Apicella, Legallet, & Trouche, 1997) and neurons in the amygdala signaling reward-predicting stimuli (Nishijo, Ono, & Nishino, 1988). Drawing on this work, we suggest that in BPD, the PFC might fail to execute inhibitory control over the processing of PEs from social feedback, thereby contributing to its increased integration.

Furthermore, research has focused on the neural base of distorted affective processing in BPD as supported by five
meta-analytic reviews (de-Almeida et al., 2012; Mitchell, Dickens, & Picchioni, 2014; Rodrigues et al., 2011; Ruocco, Amirthavasagam, & Zakzanis, 2012; Ruocco, Amirthavasagam, Choi-Kain, & McMain, 2013). In this line of (mostly MRI) research, investigators have examined how people with BPD and healthy participants process negative emotional stimuli relative to neutral stimuli. Taken together, such neuroimaging studies suggest that dysfunctional fronto-limbic brain regions underlie the emotional dysregulation in BPD (Krause-Utz, Winter, Niedtfeld, & Schmahl, 2014). Specifically, individuals with BPD showed structural and functional abnormalities in a fronto-limbic network including regions involved in emotion processing (e.g. amygdala, insula) and frontal brain regions implicated in regulatory control processes (e.g. anterior cingulate cortex, medial frontal cortex, orbitofrontal cortex, and dorsolateral PFC) (Krause-Utz et al., 2014).

In line with the assumption of BPD as an emotion dysregulation disorder, a more recent multimodal meta-analysis of neuroimaging studies synthesized that BPD is related to abnormal activation of dorsolateral prefrontal and limbic brain regions (Hazlett, 2016; Schulze, Schmahl, & Niedtfeld, 2016). In particular, BPD patients showed an increased activation of the left amygdala and posterior cingulate cortex, along with debilitating responses of the bilateral dorsolateral PFC while processing negative emotional stimuli. Interestingly, the functional corticolimbic connectivity, in particular between the right amygdala and right dorsolateral PFC, mediates the relationship between childhood adversities and symptom severity in BPD (Vai et al., 2018). Another fMRI study (Scherpier et al., 2014) found abnormalities not only in the perception but also in the anticipation of negative emotional stimuli. Collectively, this line of research is well in line with the PP account of BPD in that it provides neural evidence for the hypothesis BPD is related to the anticipation of negative events and experiences, increasing the likelihood of actually experiencing intense negative emotions, which is reflected by abnormal activity in fronto-limbic networks.

Although depending on the type, frequency and timing of exposure, CM has an influence on the child brain development with functional and structural changes observed even decades later in adulthood (Jedd et al., 2015). Associations between CM and brain structures have been widely documented with the most frequent alterations related to CM in the function and structure of lateral and ventromedial fronto-limbic brain areas and neural networks (i.e. deficits in structural interregional connectivity) that mediate behavioral, cognitive and affect control (Hart & Rubia, 2012; Lim, Radua, & Rubia, 2014). In particular, CM is related to a reduced volume of the (adult) hippocampus (Paquola, Bennett, & Lagopoulos, 2016; Teicher, Anderson, & Polcari, 2012), anterior cingulate and ventromedial and dorsolateral prefrontal and orbitofrontal cortices, as well as to the development of the corpus callosum (Teicher & Samson, 2016; Teicher, Samson, Anderson, & Ohashi, 2016). Moreover, a review of fMRI studies found that CM is associated with altered functioning in a range of neurocognitive systems (i.e. threat processing, reward processing, emotion regulation and executive control) (McCrorry, Gerin, & Viding, 2017). In fact, an association was found with an enhanced amygdala response to threatening stimuli (Dannlowski et al., 2012), reduced ventral striatal response to the anticipation or receipt of reward, decreased connectivity between prefrontal regions and the amygdala, and increased volume and network centrality of the precuneus (Teicher et al., 2016). In line with research on CM and brain alterations, a PET study showed a correlation between a dysfunction of the dorsolateral and medial PFC (including anterior cingulate) with the recall of CM specifically in BPD (Schmahl, Vermetten, Elzinga, & Bremner, 2004). Indeed, CM is associated with structural impairment in the (right) ventrolateral PFC and aggressiveness in patients with BPD (Morandotti et al., 2013). Further, some researchers found a hypoconnectivity between structures associated with emotion regulation and structures associated with social cognitive responses in BPD: Higher levels of CM were associated with reduced levels of brain connectivity, with different types of CM having differential effects on connectivity in BPD patients (Duque-Alarcón, Alcalá-Lozano, González-Olvera, Garza-Villarreal, & Pellicer, 2019). In support of our account of CM, there is some evidence for complex gene-environment interactions involving CM that determine the risk or protection against BPD pathology (Goodman, New, & Siever, 2004). Indeed, CM impacts biological processes epigenetically (Prados et al., 2015). Providing evidence of epigenome × environment interactions, epigenetic modifications of the glucocorticoid receptor gene (i.e. hGR methylation) were strongly associated with an increased vulnerability to psychopathology in CM (Radtke et al., 2015). More and more evidence points to a severe relationship dysfunction being the core epigenetic expression of BPD (Steele & Siever, 2010).

On a hormone level, early CM was associated with reduced plasma oxytocin (Klucznik et al., 2019), highlighting the role of oxytocin in BPD (Herpertz & Bertsch, 2015). In the pathogenesis of BPD, a beneficial effect of oxytocin on threat processing and stress responsiveness was found (Bertsch & Herpertz, 2018), despite considerable heterogeneity in the literature (Amad, Thomas, & Perez-Rodriguez, 2015; Bertsch & Herpertz, 2018).

The neural specification of our PP account of BPD is displayed in Fig. 3.

Novelty of this account
The present article is the first to apply current thinking in neuroscience and computational psychiatry to provide a novel mechanistic model of BPD. Specifically, by highlighting how the experience of CM is thought to alter perception, our account provides a coherent explanation as to why people with BPD show rapid fluctuations of intense emotions, cognitions, and inconsistent behaviors. Moreover, proposing that CM leads to a dysbalance between the precision of prior beliefs and data, our account can also explain why people with BPD are highly sensitive to negative social feedback, whereas information processing in other domains is intact. Thus, in contrast to previous cognitive theories of BPD (Arntz, 1994; Arntz, Dietzel, & Dreesen, 1999; Arntz, Dreesen, Schouten, & Weertman, 2004; Baer, Peters, Eisenlohr-Moul, Geiger, & Sauer, 2012; Beck, Davis, & Freeman, 2015; Butler, Brown, Beck, & Grisham, 2002), we suggest that BPD may not primarily be related to the presence of dysfunctional beliefs per se (i.e. their contents), but to aberrant precision afforded to them. This is consistent with recent theories from computational psychiatry (Kube & Rozenkrantz, 2021; Paulus et al., 2019). Moreover, our account goes beyond previous theories of BPD as a disorder of emotion regulation (Linehan, 1993) by assuming that the instability of affect, cognition, and behavior has one underlying pathology, that is, overly weak prior predictions. At the same time, our model can also account for the rigidity (or: lack of flexibility) of BPD patients by conceiving of it as an
expression of the patients’ impoverished generative model resulting from CM.

**Implications for the treatment and prevention of BPD**

On the basis of our review, two major clinical implications can be drawn. First, the crucial role of CM on the perception and cognition of a child and its emotional development. People generate a model of its environment (including the external world and the body) which predicts future sensory inputs and is updated by PEs, depending on how precise these error signals are. Expectation and attention increase the integration of top-down and bottom-up signals in perception (Gordon, Tsuchiya, Koenig-Robert, & Hohwy, 2019). Interventions focusing on both enhancing sensitivity to the validity of sensory input/evidence and therefore PE; and attention on expectations vs. sensory input to disentangle imbalance might be promising treatment approaches. In general, individuals tend to have an optimism bias, processing desirable information more frequently than undesirable information. Findings suggest that BPD patients appear initially more pessimistic about their personal future than healthy people but they might be able to overcome their pessimism when provided with relevant information (Korn et al., 2016b). Indeed, optimism training has potential to change individuals with mild dysphoria perceptions about the future (Yoshimura & Hashimoto, 2020). By sampling precise empirical evidence, BPD patients might enhance their predictability of the world and reduce uncertainty in their beliefs. Before, traditional cognitive techniques such as the, Downward Arrow™-technique – sometimes referred to as Vertical Descent (Leahy, 2017) – might be useful to identify relevant expectations (‘predictions’). Individual treatment should then focus on techniques to increase the attention of BPD patients to PEs by an elevated propensity to counterbalance and weight perceptual beliefs (priors) over sensory evidence and interventions to empirically examine the credibility of one’s beliefs in order to strengthen the precision of one’s priors and enhance belief updating. Attention alters PP indicating that there are top-down effects of attention on perception (Clark, 2016b). Some researchers have proposed that salience is something that is afforded to actions that realize epistemic affordance, while attention per se is afforded to precise sensory evidence – or beliefs about the causes of sensations (Parr & Friston, 2017). Indeed, attention and PE suggest that information sought by top-down-attention is prioritized (Ramamoorthy, Parker, Plaisted-Grant, Muhl-Richardson, & Davis, 2020). Attention optimizes the expected precision of predictions by modulating the synaptic gain of PE units, that is, attention increases the selectivity for mismatch information in the neural response to a surprising stimulus indicating that attention optimizes precision expectations during hierarchical inference by increasing the gain of PEs (Smout, Tang, Garrido, & Mattingley, 2019). Furthermore, it is widely accepted that predictions across different stimulus
attributes (e.g. time and content) facilitate sensory processing. Content (‘what’) and temporal (‘when’) predictions engage complementary neural mechanisms in different brain regions, suggesting domain-specific prediction signaling along the cortical hierarchy (Auksztulewicz et al., 2018). Moreover, as BPD patients perceive the world as ambiguous, building new expectations should be based on repetitions and hints which have been shown to facilitate perceptual experience of ambiguous images (Hertz, Blakemore, & Frith, 2020). Yet, people have two main pathways to determine the veracity of the sensory input: the perceived credibility of the source and direct-evaluation via first-hand evidence, i.e. testing the advice against observation. Beliefs are interpreted in light of the perceived credibility of the source in form of credibility-led biased interpretations of evidence (whether belief or suspicion confirming) that lead to further polarization of the perceived credibility highlighting the crucial role of credibility in belief updating (Pilditch, Madsen, & Custers, 2020), while cues including valence and relevance influence these credibility judgments suggesting a utility and credibility trade off during decision making (Gugerty & Link, 2020). Therefore, creating a more nuanced credibility picture in BPD patients might also be promising treatment target. By providing the brain with an intense inflow of salient and unambiguous bottom-up sensory input to shift the brains mapping of the observed body state, interoceptive exposure techniques (including exteroceptive as well as interoceptive feedback and imagined scenarios in naturalistic approach (Miano, Fertuck, Roepke, & Dziobeck, 2017) and computer-based social rejection paradigms [e.g. virtual-reality-enhanced cyberball paradigm (McCull, 2016) or social feedback and imagined scenarios in NoOneLikesYou! paradigm (D’Astolfo, Kirchner, & Rief, 2020)] or social interactive paradigms to quantify social behavior [e.g. multi-round economic exchange games such as the established trust game (Rilling & Sanfey, 2011; Tzieropoulos, 2013)]

| Hypothesis | Empirical examination |
|------------|-----------------------|
| Influence of CM on adult BPD is mediated by altered perception and cognition in adolescence characterized through high uncertainty of beliefs (i.e. weak priors). | Longitudinal study on the influence of CM on perception and cognition in adulthood |
| Specific hypotheses:  
- The higher the degree of perceived social rejection, the less trust BPD patients have towards others compared to healthy controls, regardless of actual social exclusion.  
- The more volatile social interactions are, the more BPD patients update their beliefs as a function of sensory input compared to healthy controls.  
- The higher the degree of ambiguity in social interactions, the more BPD patients update their beliefs as a function of sensory input compared to healthy controls.  
- The higher the degree of perceived interpersonal rupture, the less cooperative and the less able to repair broken cooperations BPD patients are. | Experimental studies investigating (negative) social feedback using mixed methods for symptom provocation in the lab, e.g. personalized narratives in the form of script-driven imagery (Kraus et al., 2010) or naturalistic approach (Miano, Fertuck, Roepke, & Dziobeck, 2017) and computer-based social rejection paradigms [e.g. virtual-reality-enhanced cyberball paradigm (McCull, 2016) or social feedback and imagined scenarios in NoOneLikesYou! paradigm (D’Astolfo, Kirchner, & Rief, 2020)] or social interactive paradigms to quantify social behavior [e.g. multi-round economic exchange games such as the established trust game (Rilling & Sanfey, 2011; Tzieropoulos, 2013)] |
| BPD is related to reduced precision of prior predictions and increased precision afforded to new information, entailing enhanced updating in line with new information. | Using computational modeling to assess the influence of prior precision and data precision on the posterior |
| Elevated levels of dopamine are related to the increased integration of sensory information. | Combining neurochemical analyses with computational modeling |
| Dopamine encodes the precision of PEs rather than PEs per se. | Administering dopamine-decreasing drugs in social rejection paradigms |
| Providing drugs that decrease the dopamine system (DRD2-blocking antipsychotics) weakens the precision afforded to and the increased integration of (negative) social feedback into beliefs. | |
| The lack of inhibitory control of PE processing from social feedback executed by the PFC leads to an enhanced integration of social information into beliefs. | Connectivity analysis in fMRI |

Note: BPD, Borderline personality disorder; CM, Childhood maltreatment; PE, prediction error; PFC, prefrontal cortex.
predictions of the body state and somatic error might naturally van-
ish as the mind attempts to function with low-precision priors trig-
nering fewer regulatory responses and that allows the entire pro-
predictive model to be driven by incoming sensory input from
the present moment in time. Traditionally not a standard tool in
psychotherapy, new interventions such as the floatation-REST
(Reduced Environmental Stimulation Therapy)\(^4\) might be able to
enrich current state-of-the-art treatments such as DBT (Feinstein
et al., 2018). Furthermore, whole body hyperthermia (WBH)
(Janssen et al., 2016), modulation of muscle tension via Swedish
massage (Rapaport et al., 2016), yoga (Jeter, Slutsky, Singh, &
Khalsa, 2015), and exercise (Smits, Berry, Tart, & Powers, 2008);
repeated brief exposures to high doses of CO\(_2\) (Wolpe, 1987); and
cyclic activation of the sympathetic nervous system and suppression
of the immune response through cold immersion and
CO\(_2\)-modulation using alternating cycles of hyperventilation fol-
lowed by breath holding (Kox et al., 2014) might also yield a future
potential in this regard (Paulus et al., 2019).

Second, this account highlights the role of prevention of CM in
risk populations with low educational skills, and in parents suffer-
ing also from mental disorders, in particular emotion regulation
deficits (such as in the case of mothers with BPD). Both, parent-
ing programs as well as mother-child-interventions might serve as
important preventive strategies to reduce the occurrence of CM.
For example, in some clinics there exist group psychotherapy
for mothers suffering from BPD, and evidence-based parenting
programs such as the Triple p – Positive Parenting Program
(Bodenmann, Cina, Ledermann, & Sanders, 2008; Sanders,
1999, 2008, 2012). In a critical and sensitive phase in child devel-

dopment, children must be encouraged to build appropriate confi-
dence (i.e. precision) in their expectations and beliefs (i.e. priors)
by a validating environment that see their emotions and fulfill
their needs appropriately leading to fewer PEs (i.e. minimizing
uncertainty and surprise in the environment to make the world
and behavior of others more predictable to them).

Limitations and future directions

Although the majority of patients with BPD report maltreatment
in their childhood (Battle et al., 2004; Carlson et al., 2009) and
patients with BPD were over 13 times more likely to report child-
hood adversity than non-clinical controls and other clinical popu-
lations (Kleindienst et al., 2020; Porter et al., 2020), a major
limitation of our model is that it only explains aberrant belief
updating in patients with BPD that have experienced some sort
of CM. For BPD patients without CM in the past, genetics
might be a more relevant risk factor in the pathogenesis of BPD as
studies found a moderate to high heritability in BPD (Distel
et al., 2008, 2010; Skoglund et al., 2021). However, despite the
potentially different etiology, we believe that a similar psychobi-
ological pathology in the sense of PP (i.e. the role of priors, preci-
sion and likelihood) underlies also in those patients with BPD.
Furthermore, although particularly relevant in BPD (Brakemeier
et al., 2018; Pietrek et al., 2013; Quenneville et al., 2020; Wota
et al., 2014), CM is considered as a transdiagnostic factor
(Gilbert et al., 2009; Struck et al., 2020) and future studies should
investigate the specific causal contribution of CM to the patho-
genesis of BPD. Though our account can build upon the wealth
of literature on childhood adversities, no consensus has been
reached about how to measure poor parenting or the invalidating
environment to quantify the extent to which these types of spe-
cific factors contribute to BPD (Musser et al., 2018). Also,
different forms of CM may have different consequences. Similarly,
there is a lack of assessment tools to measure a key element of our account: the precision of prior beliefs and new
information. Therefore, much of what we proposed here as
conceptually-based evidence still needs to be empirically tested
in future work using rigorous experimental designs and computa-
tional modeling to generate data-based evidence.

Despite strong arguments in favor of PP (Seth, Millidge,
Buckley, & Tschantz, 2020; van de Cruys, Friston, & Clark,
2020), it should be noted there have also been some critical voices
regarding PP as a ‘theory-of-everything’ (Hutto, 2020; Litwin &
Milkowski, 2020; Sun & Firestone, 2020). Indeed, there seem over-
lapping features and similarities to traditional theories (F.A.
França, 2020) and some researchers urge for refinements of the
PP theory in order to increase its explanatory power (Gilead,
Trope, & Liberman, 2020; Vilas, Melloni, & Melloni, 2020).
This critique is in part also relevant to the present article as it
applies one set of theoretical assumptions to explain a fairly het-
erogeneous disorder. Yet, we believe that PP has the potential to
inspire future research and allows to derive some novel hypotheses
about the psychopathology of BPD that may contribute to a more
nuanced understanding of this complex disorder, particularly in
view of the developmental psychopathology perspective PP offers.

For an overview of specific hypotheses that can be derived
from our account, see Table 1.

Conclusions

In this article, we proposed a Bayesian account of BPD that relies
on the observation that many people with BPD experienced some
sort of maltreatment in childhood. We argued that such adverse
experiences lead to pervasive alterations in perception. In essence,
we suggested that CM impairs the continuous refinement of pre-
dictive models of the world and precipitates the formation of weak
and strong prior predictions, relative to sensory information. This
entails a distorted belief updating in response to novel informa-
tion, resulting in a marked instability of affect, cognitions, and
behavior – that is, the core features of BPD.

Supplementary material. The supplementary material for this article can
be found at https://doi.org/10.1017/S0033291722002458.

Acknowledgements. Figure 3 was created using a template from Servier
Medical Art by Servier (http://smart.servier.com/), licensed under a Creative
Common Attribution 3.0 Unported License (https://creativecommons.org/
licenses/by/3.0/).

Author contributions. P. H. conceived the main original ideas. All authors
contributed to the development of the theoretical concept and design of the
model. P. H. drafted the paper, and T. K. and E. F. provided critical revisions.
All authors approved the final version of the paper for submission.

Financial support. E. F. obtained funding from the Else Kröner-Fresenius-
Stiftung (2018_A152). This research did not receive any other specific grant
from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest. E. F. has provided trainings, presentations and pub-
lished books/chapters on the treatment of BPD. P. H. and T. K. have no con-
licts of interest to declare.

Notes

1 In addition to these emotion regulation deficits, CM may also contribute to
aberrant social cognition in BPD as discussed in further detail in the supple-
ment (see online Supplemental Material 1).
In the online Supplemental Material 2, we provide an explanation for the other end of the emotional spectrum in BPD, i.e. chronic emptiness. Of note, in domains that are not core to BPD, such as the integration of allostasis and interoception in depression. Not all predictions are equal: A sensory-reduced intervention that attenuates exteroceptive sensory input to brain by reducing the impact of predictive models through the act of floating brain by reducing the impact of predictive models through the act of floating

References

Afifi, T. O., Mather, A., Boman, J., Fleisher, W., Enns, M. W., MacMillan, H., & Sareen, J. (2011). Childhood adversity and personality disorders: Results from a nationally representative population-based study. Journal of Psychiatric Research, 45(6), 814–822. https://doi.org/10.1016/j.jpsychires.2010.11.008.

Allen, B. (2008). An analysis of the impact of diverse forms of childhood psychological maltreatment on emotional adjustment in early adulthood. Child Maltreatment, 13(3), 307–312. https://doi.org/10.1177/1077559508318394.

Amad, A., Ramoz, N., Thomas, P., Jardri, R., & Gornwood, P. (2014). Genetics of borderline personality disorder: Systematic review and proposal of an integrative model. Neuroscience and Biobehavioral Reviews, 40, 6–19. https://doi.org/10.1016/j.neubiorev.2014.01.003.

Amad, A., Thomas, P., & Perez-Rodriguez, M. (2015). Borderline personality disorder and oxytocin: Review of clinical trials and future directions. Current Pharmaceutical Design, 21(23), 3311–3316. https://doi.org/10.1080/1070–

Arrtzt, A. (1994). Treatment of borderline personality disorder: A challenge for cognitive-behavioural therapy. Behaviour Research and Therapy, 32(4), 419–430. https://doi.org/10.1016/0005-7967(94)90005-1.

Arrtzt, A., Dietel, R., & Dreessen, L. (1999). Assumptions in borderline personality disorder: Specificity, stability and relationship with etiological factors. Behaviour Research and Therapy, 37(6), 545–557. https://doi.org/10.1016/S0005–7967(98)00152-1.

Arrtzt, A., Dreessen, L., Schouten, E., & Weertman, A. (2004). Beliefs in personality disorders: A test with the Personality Disorder Belief Questionnaire. Behaviour Research and Therapy, 42(10), 1215–1225. https://doi.org/10.1016/j.brat.2003.08.004.

Auksztulewicz, R., Schwidetzik, C. M., Thesen, T., Doyle, W., Devinsky, O., Nobre, A. C., ... Melloni, L. (2018). Not all predictions are equal: ‘what’ and ‘when’ predictions modulate activity in auditory cortex through different mechanisms. Journal of Neuroscience, 38(40), 8680–8693. https://doi.org/10.1523/JNEUROSCI.3669-18.2018.

Bair, R. A., Peters, J. R., Eisenlohr-Moul, T. A., Geiger, P. J., & Sauer, S. E. (2012). Emotion-related cognitive processes in borderline personality disorder: A review of the empirical literature. Clinical Psychology Review, 32(5), 359–369. https://doi.org/10.1016/j.cpr.2012.03.002.

Balestri, M., Calati, R., Serretti, A., & de Ronchi, D. (2014). Genetic modulation of personality traits: A systematic review of the literature. International Journal of Psychopharmacology, 29(1), 1–15. https://doi.org/10.1097/YIC.0b013e328364590b.

Barnicot, K., Katsakou, C., Marouga, S., & Priebe, S. (2011). Treatment completion in psychotherapy for borderline personality disorder – a systematic review and meta-analysis. Acta Psychiatica Scandinavica, 123(5), 327–338. https://doi.org/10.1111/j.1600-0447.2010.01652.x.

Barrett, L. F., Quigley, K. S., & Hamilton, P. (2016). An active inference theory of allostasis and interoception in depression. Philosophical Transactions of the Royal Society B: Biological Sciences, 371(1708). https://doi.org/10.1098/rstb.2016.0011.

Barrett, L. F., & Satpute, A. B. (2013). Large-scale brain networks in affective and social neuroscience: Towards an integrative functional architecture of the brain. Current Opinion in Neurobiology, 23(3), 361–372. https://doi.org/10.1016/j.conb.2012.12.012.

Barrett, L. F., & Simmons, W. K. (2015). Interoceptive predictions in the brain. Nature Reviews Neuroscience, 10(7), 419–429. https://doi.org/10.1038/nrn3950.

Bastos, A. M., Usrey, W. M., Adams, R. A., Manguin, G. R., Fries, P., & Friston, K. J. (2012). Canonical microcircuits for predictive coding. Neuron, 76(4), 695–711. https://doi.org/10.1016/j.neuron.2012.10.038.

Batte, C. L., Shea, M. T., Johnson, D. M., Yen, S., Zlotnick, C., Zanarini, M. C., ... Morey, L. C. (2004). Childhood maltreatment associated with adult personality disorders: Findings from the collaborative longitudinal personality disorders study. Journal of Personality Disorders, 18(2), 193–211. https://doi.org/10.1521/pedi.18.2.193.32777.

Beck, A. T., Davis, D. D., & Freeman, A. (2015). Cognitive therapy of personality disorders. New York, NY, USA: Guilford Press.

Bertsch, K., & Herpertz, S. C. (2018). Oxytocin and borderline personality disorder. In R. Hurlermein & V. Grinevich (Eds.), Behavioral pharmacology of neuropeptides: Oxytocin. Current topics in behavioral neurosciences (Vol. 35, pp. 499–514). Cham: Springer Verlag. https://doi.org/10.1007/978–584_2017_26.

Bevilaqua, L., & Goldman, D. (2013). Genetics of impulsive behaviour. Philosophical Transactions of the Royal Society B: Biological Sciences, 368(1615). https://doi.org/10.1098/rstb.2012.0380.

Benzdijian, S., Baker, L. A., & Turblad, C. (2011). Genetic and environmental influences on impulsivity: A meta-analysis of twin, family and adoption studies. Clinical Psychology Review, 31(7), 1209–1223. https://doi.org/10.1016/j.cpr.2011.07.005.

Bodenmann, G., Cina, A., Ledermann, T., & Sanders, M. R. (2008). The efficacy of the Triple p–Positive Parenting Program in improving parenting and child behavior: A comparison with two other treatment conditions. Behaviour Research and Therapy, 46(4), 411–427. https://doi.org/10.1016/j.brat.2008.01.001.

Bonnova, A. M., Gratz, K. L., Delany-Brumsey, A., Paulson, A., & Lejuez, C. W. (2006). Temperamental and environmental risk factors for borderline personality disorder among inner-city substance users in residential treatment. Journal of Personality Disorders, 20(3), 218–231. https://doi.org/10.1521/pedi.2006.20.3.218.

Bonnova, A. M., Hulbregtse, B. M., Hicks, B. M., Keyes, M., McGue, M., & Iacono, W. (2013). Tests of a direct effect of childhood abuse on adult borderline personality disorder traits: A longitudinal discordant twin design. Journal of Abnormal Psychology, 122(1), 180–194. https://doi.org/10.1037/a0028328.

Boucher, M.-E., Puglise, J., Allard-Chapais, C., Locours, S., Ahoundova, L., Chounnard, R., & Gaham, S. (2017). Parent-child relationship associated with the development of borderline personality disorder: A systematic review. Personality and Mental Health, 11(4), 229–255. https://doi.org/10.1002/pmh.1385.

Bradley, R., Jenei, J., & Westen, D. (2005). Etiology of borderline personality disorder. The Journal of Nervous and Mental Disease, 193(1), 24–31. https://doi.org/10.1097/00005990–1223. https://doi.org/10.1159/0000848481.

Brown, H. R., & Friston, K. J. (2012). Dynamic causal modelling of precision and synaptic gain in visual perception – an EEG study. Neuroimage, 63(1), 223–231. https://doi.org/10.1016/j.neuroimage.2012.06.044.

Butler, A. C., Brown, G. K., Beck, A. T., & Grisham, J. R. (2002). Assessment of dysfunctional beliefs in borderline personality disorder. Behaviour Research and Therapy, 40(10), 1231–1240. https://doi.org/10.1016/S0006–7967(02)00301-1.

Carlson, E. A., Eggeland, B., & Sroufe, L. A. (2009). A prospective investigation of the development of borderline personality symptoms. Development and Psychopathology, 21(4), 1311–1334. https://doi.org/10.1017/S095457940990174.
Siegel, J. Z., Curwell-Parry, O., Pearce, S., Saunders, K. E. A., & Crockett, M. J. (2020). A computational phenotype of disrupted moral inference in borderline personality disorder. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 5(12), 1134–1141. https://doi.org/10.1016/j.bpsc.2020.07.013.

Skoglund, C., Tiger, A., Rück, C., Petrovic, P., Asherson, P., Hellner, C., … Kuja-Halkola, R. (2021). Familial risk and heritability of diagnosed borderline personality disorder: A register study of the Swedish population. *Molecular Psychiatry*, 26(3), 999–1008. https://doi.org/10.1038/s41386-019-0442-0.

Smith, R., Kirlic, N., Stewart, J., Touthang, J., Kuglicki, R., Khalsa, S., … Aupperle, R. (2020). Greater decision uncertainty characterizes a transdiagnostic patient sample during approach-avoidance conflict: A computational modeling approach. *Journal of Psychiatry & Neuroscience: JPN*, 46(1), 200032. https://doi.org/10.31234/osf.io/t2dnh.

Smith, R., Lane, R. D., Parr, T., & Friston, K. J. (2019a). Neurocomputational mechanisms underlying emotional awareness: Insights afforded by deep active inference and their potential clinical relevance. *Neuroscience and Biobehavioral Reviews*, 107, 473–491. https://doi.org/10.1016/j.nebiorev.2019.09.002.

Smith, R., Parr, T., & Friston, K. J. (2019b). Simulating emotions: An active inference model of emotional state and emotion concept learning. *Frontiers in Psychology*, 10. https://doi.org/10.3389/fpsyg.2019.02844.

Smith, R., Steklis, H. D., Steklis, N., Weihs, K. L., Allen, J. J., & Lane, R. D. (2022). Lower emotional awareness is associated with greater early adversity and faster life history strategy. *Evolutionary Behavioral Sciences*. Advance online publication. https://doi.org/10.1007/s41324-021-00251-6.

Smits, J. A. J., Berry, A. C., Tart, C. D., & Powers, M. B. (2008). The efficacy of cognitive-behavioral interventions for reducing anxiety sensitivity: A meta-analytic review. *Behaviour Research and Therapy*, 46(9), 1047–1054. https://doi.org/10.1016/j.brat.2008.06.010.

Smout, C. A., Tang, M. F., Garrido, M. I., & Mattingley, J. B. (2019). Attention promotes the neural encoding of prediction errors. *PLoS Biology*, 17(2). https://doi.org/10.1371/journal.pbio.2006812.

Sokolov, E. N. (1963). *Perception and the conditioned reflex*. Oxford: Pergamon Press.

Sokolov, E. N. (1990). The orienting response, and future directions of its study. In P. N. Glaser (Ed.), *The Pavlovian Journal of Biological Science* (Vol. 25, pp. 5–12). Oxford: Pergamon Press.

Steele, H., & Siever, L. (2010). An attachment perspective on borderline personality disorder: A register study of the Swedish population. *Molecular Psychiatry*, 26(3), 999–1008. https://doi.org/10.1038/s41386-019-0442-0.

Sun, Z., & Firestone, C. (2020). The dark room problem. *Trends in Cognitive Sciences*, 24, 346–348. https://doi.org/10.1016/j.tics.2020.02.006.

Teicher, M. H., Anderson, C. M., & Polcari, A. (2012). Childhood maltreatment is associated with reduced volume in the hippocampal subfields CA3, dentate gyrus, and subiculum. *Proceedings of the National Academy of Sciences of the United States of America*, 109(9), E563–E572. https://doi.org/10.1073/pnas.1115396109.

Teicher, M. H., & Samson, J. A. (2016). Annual Research Review: Enduring neurobiological effects of childhood abuse and neglect. *Journal of Child Psychology and Psychiatry*, 57(3), 241–266. https://doi.org/10.1111/jcpp.12507.

Teicher, M. H., Samson, J. A., Anderson, C. M., & Olahsi, K. (2016). The effects of childhood maltreatment on brain structure, function and connectivity. *Nature Reviews Neuroscience*, 17(10), 652–666. https://doi.org/10.1038/nrn.2016.111.

Tziropoulos, H. (2013). The trust game in neuroscience: A short review. *Social Neuroscience*, 8(5), 407–416. https://doi.org/10.1080/17470919.2013.832375.

Unoka, Z., Seres, L., Áspán, N., Bodí, N., & Kéri, S. (2009). Trust game reveals biological effects of childhood abuse and neglect. *Journal of Personality Disorders*, 23(4), 399–409. https://doi.org/10.1521/pedi.2009.23.4.399.

Vai, B., Sforzini, L., Visintini, R., Riberto, M., Bulgarelli, C., Ghiglino, D., … Benedetti, F. (2018). Corticolimbic connectivity mediates the relationship between adverse childhood experiences and symptom severity in borderline personality disorder. *Neuropsychobiology*, 76(2), 105–115. https://doi.org/10.1159/000487961.

van de Cruys, S., Evers, K., van der Hallen, R., van Eylen, L., Boets, B., de-Wit, L., & Wagemans, J. (2014). Precise minds in uncertain worlds: Predictive coding in autism. *Psychological Review*, 121(4), 649–675. https://doi.org/10.1037/a0037665.

van de Cruys, S., Friston, K. J., & Clark, A. (2020). Controlled optimism: Reply to Sun and firestone on the dark room problem. *Trends in Cognitive Sciences*, 24(9), 680–681. https://doi.org/10.1016/j.tics.2020.05.012.

van den Bergh, O., Witthöft, M., Petersen, S., & Brown, R. J. (2017). Symptoms and the body: Taking the inferential leap. *Neuroscience & Biobehavioral Reviews*, 74(Pt A), 185–203. https://doi.org/10.1016/j.neubiorev.2017.01.015.

Van Schie, C. C., Chiu, D. C., Rombouts, S. A. R. B., Heiser, W. J., & Elzinga, B. M. (2020). Stuck in a negative me: FMRI study on the role of disturbed self-views in social feedback processing in borderline personality disorder. *Psychological Medicine*, 50(4), 625–635. https://doi.org/10.1017/S0033291719000448.

Vansteelandt, K., Houben, M., Claes, L., Berens, A., Sleuwaegen, E., Sienaert, P., & Kuppens, P. (2017). The affect stabilization function of nonsuicidal self injury in borderline personality disorder: An ecological momentary assessment study. *Behaviour Research and Therapy*, 92, 41–50. https://doi.org/10.1016/j.brat.2017.02.003.

Vilas, M. G., Melloni, L., & Melloni, L. (2020). A challenge for predictive coding: Representational or experiential diversity? *Behavioral and Brain Sciences*, 43. https://doi.org/10.1017/S0140525X19003157.

Wager, T. D., Kang, J., Johnson, T. D., Nichols, T. E., Satpute, A. B., & Barrett, L. F. (2015). A Bayesian model of category-specific emotional brain responses. *PLOS Computational Biology*, 11(4), e1004066. https://doi.org/10.1371/journal.pcbi.1004066.

Watanabe, M. (1989). The appropriateness of behavioral responses coded in post-traual activity of primate prefrontal units. *Neuroscience Letters*, 101(1), 113–117. https://doi.org/10.1016/0304-3940(89)90450-3.

Weaver, T. L., & Clum, G. A. (1993). Early family environments and traumatic experiences associated with borderline personality disorder. *Journal of Consulting and Clinical Psychology*, 61(6), 1068–1075. https://doi.org/10.1037/0022-006x.61.6.1068.

Widom, C. S., Czaia, S. J., & Paris, J. (2009). A prospective investigation of borderline personality disorder symptoms in childhood and adolescence. *Biological Psychiatry*, 76(4), 392–400. https://doi.org/10.1016/j.biopsych.2008.02.018.

Winsper, C., Marwaha, S., Lereya, S. T., Thompson, A., Eyden, J., & Singh, S. P. (2016). A systematic review of the neurobiological underpinnings of borderline personality disorder (BPD) in childhood and adolescence. *Reviews in the Neurosciences*, 27(8), 827–847. https://doi.org/10.1515/rvre-2016-0026.
Wolpe, J. (1987). Carbon dioxide inhalation treatments of neurotic anxiety: An overview. *Journal of Nervous and Mental Disease, 175*(3), 129–133. https://doi.org/10.1097/00005053-198703000-00001.

Wota, A. P., Byrne, C., Murray, I., Ofuafor, T., Nisar, Z., Neuner, F., & Hallahan, B. P. (2014). An examination of childhood trauma in individuals attending an adult mental health service. *Irish Journal of Psychological Medicine, 31*(4), 259–270. https://doi.org/10.1017/ipm.2014.49.

Yoshimura, S., & Hashimoto, Y. (2020). The effect of induced optimism on the optimistic update bias. *BMC Psychology, 8*(1). https://doi.org/10.1186/s40359-020-0389-6.

Zanarini, M. C., Frankenburg, F. R., Wedig, M. M., & Fitzmaurice, G. M. (2013). Cognitive experiences reported by patients with borderline personality disorder and axis ii comparison subjects: A 16-year prospective follow-up study. *American Journal of Psychiatry, 170*(6), 671–679. https://doi.org/10.1176/appi.ajp.2013.13010055.

Zhang, T. H., Chow, A., Wang, L. L., Yu, J. H., Dai, Y. F., Lu, X., … Xiao, Z. P. (2013). Childhood maltreatment profile in a clinical population in China: A further analysis with existing data of an epidemiologic survey. *Comprehensive Psychiatry, 54*(7), 856–864. https://doi.org/10.1016/j.comppsych.2013.03.014.