Counselling and Psychotherapy Research

WILEY

ORIGINAL ARTICLE OPEN ACCESS

Clinicians' Perspectives on the Model of Trust Processes in Borderline Personality Disorder: A Qualitative Evaluation

Emanuele Preti[1,2] | Erika Fanti[1] | Rossella Di Pierro[1,2] | Alessandra Frigerio[3]

[1]University of Milano-Bicocca, Department of Psychology, Milan, Italy | [2]Personality Disorder Lab (PDLab), Parma, Italy | [3]Sigmund Freud University, Department of Psychology, Milan, Italy

Correspondence: Emanuele Preti (emanuele.preti@unimib.it)

Received: 18 July 2025 | **Revised:** 10 December 2025 | **Accepted:** 23 December 2025

Keywords: borderline personality disorders | clinicians' interview | interpersonal dysfunction | psychopathology | thematic analysis | therapeutic relationship | trust processes

ABSTRACT

Background: Trust is a fundamental aspect of human relationships, crucial for navigating social interactions and forming meaningful connections. While the ability to accurately assess and update trust appraisals is essential for adaptive functioning, individuals with borderline personality disorder (BPD) experience significant challenges in trust processing that contribute to interpersonal dysfunction and treatment complications. Despite growing recognition of trust difficulties in BPD, a comprehensive understanding of these processes remains limited.

Methods: The present study aimed to evaluate and refine this model through a qualitative analysis of interviews with 25 Italian therapists specialising in personality disorders. Using thematic analysis, we explored clinicians' observations of trust dynamics in BPD patients, focusing on developmental factors, situational perceptions, emotional influences, prior beliefs, behaviours and trust learning capacity. Our findings largely aligned with and extended the theoretical model of trust impairments in BPD proposed by Preti et al. in 2023.

Results: Results revealed a complex interplay between long-term and immediate factors in shaping trust processes: developmental experiences, rigid perceptions, emotional volatility, distorted beliefs, maladaptive behaviours, and impairment in trust learning were central to the trust dysfunctions in BPD. Two significant additions to the model emerged: emotional consequences of trust misattribution and behavioural activation triggered by emotional arousal.

Conclusions: These findings suggest that trust difficulties in BPD involve both planned behavioural consequences and immediate emotion-driven reactions. The study provides important implications for clinical practice, emphasising the need for interventions that address both primary trust issues and secondary emotional-behavioural responses.

1 | Background

Trust is a core interpersonal process that originates in early attachment relationships, where expectations about others' reliability, availability, and intentions are first shaped and internalised (Fonagy and Campbell 2017). Beyond its affective dimension, trust operates as a core social drive: it enables individuals to navigate interpersonal environments and sustain cooperative exchanges (Zhao et al. 2024). The capacity to evaluate others' intentions and revise these evaluations in light of new information is fundamental to effective social functioning (Fertuck and Preti 2023). However, individuals with borderline personality disorder (BPD) often show pervasive disruptions in trust-related processes (Preti et al. 2023). BPD is a complex disorder characterised by impulsivity, emotional instability, and profound difficulties in self-image and relationships (American

Summary

- Implications for Practice
  - Trust impairments in BPD should be understood as complex patterns involving both long-term and immediate factors, requiring clinicians to address both historical and present-day factors in treatment.
  - Treatment should address both primary trust issues and secondary responses, as emotional consequences of trust misattribution can trigger behavioural activation, leading to relationship deterioration.
  - Therapeutic interventions should specifically target both deliberate behavioural consequences (e.g., relationship avoidance) and immediate emotional-behavioural reactions (e.g., aggression, withdrawal) to trust issues.
- Implications for Policy
  - Mental health policies should promote clinical training and service guidelines that address trust impairments as a central feature of BPD, ensuring that clinicians are equipped to identify and treat both primary and secondary trust-related difficulties in BPD.

Psychiatric Association 2022). While earlier work emphasised affective and behavioural dysregulation (e.g., Kaufman et al. 2020; Kockler et al. 2022), emerging evidence highlights interpersonal dysfunction as a defining feature of the disorder stemming from attachment-related models of others as unreliable, rejecting, or threatening (e.g., Euler et al. 2021; Lazarus et al. 2014). Indeed, these trust disruptions reflect the internal working models shaped by early, inconsistent or frightening caregiving (Agrawal et al. 2004). Epistemic trust, the ability to treat others' communications as credible and personally relevant, also emerges from these early relational experiences and is easily compromised when caregivers are unpredictable or unresponsive (Sharp and Balzen 2025). In BPD, such developmental disruptions often result in models of others as unreliable or malevolent and of the self as unworthy or unsafe in relationships, amplifying mistrust and interpersonal instability (Sharp and Balzen 2025). This pattern is also evident in the therapeutic relationship, where these internal models shape how patients anticipate and interpret the clinician's intentions: disrupted trust patterns in BPD often evoke negative countertransference (Gabbard 1993), hinder therapeutic alliance formation (Richardson-Vejlgaard et al. 2013), and lead to polarised views of the therapist as either idealised or persecutory (Fertuck, Fischer, and Beeney 2018). These dynamics underscore the importance of clarifying trust processes in BPD to inform clinical practice, strengthen the therapeutic relationship and reduce dropout (Iliakis et al. 2021).

Impairments in trust can manifest in various forms, such as unstable trust appraisals, low trust ratings, or difficulty updating evaluations based on new experiences (Preti et al. 2023). Insights from developmental psychology (Morales-Muñoz et al. 2021), social psychology (Masland et al. 2020), and neuroscience (Fertuck, Grinband, et al. 2018) have helped clarify underlying mechanisms. However, an integrative conceptual framework that integrates these insights into a model of trust processes in BPD remains lacking. Other psychopathological models of BPD also emphasise the interaction between dispositional vulnerabilities and situational interpersonal triggers—for example, D'Agostino et al. (2018) link background dysphoria and negative interpersonal disposition to situational dysphoria and core BPD symptoms—but these frameworks do not address trust-specific processes.

Our model of trust processes in BPD (see Figure 1; Preti et al. 2023) is built on theories from psychology, sociology, anthropology, economics, and political science (e.g., Lewis and Weigert 1985; Shapiro 1987) and conceptualises trust in BPD as a dynamic, multi-stage process involving five core components. At its core, trust is a relational process that unfolds within interactions: individuals (trustors) rely on others' behaviours to achieve their goals, and those others (trustees) can either confirm or disconfirm their expectations. These dynamics are shaped by early attachment experiences and by the relational contexts in which interactions occur (Preti et al. 2023). Building on this perspective, the model by Preti et al. (2023) conceptualises trust not as a static disposition but as a multi-stage interpersonal process shaped by developmental, cognitive, emotional, and contextual influences.

First, developmental factors (*distal factors*), such as early caregiver bonds and attachment disruptions, trauma and temperamental vulnerabilities, shape the baseline dispositions individuals bring into relationships (Fertuck et al. 2023).

Second, Immediate Influences (*Proximal Factors*) include (Fertuck et al. 2023): (1) the perception of the situation/social context (e.g., situations with high or low social rejection potential; e.g., Jayawickreme et al. 2021), (2) emotional states (e.g., anger, calm, distress), and (3) core beliefs and dispositions brought into social encounters (e.g., deserving fairness, whether people are generally trustworthy). The model highlights that these situational cues can shift trust appraisals moment-to-moment, reflecting the sensitivity of borderline functioning to immediate interpersonal stressors (e.g., Miskewicz et al. 2015). These distal and proximal elements converge in the stage of trust appraisal, where individuals judge others' intentions and trustworthiness—for instance, a tendency to view neutral expressions as untrustworthy—and shape choices like whether to initiate or avoid relationships, how to revise trust judgements over time, and how consistently they trust across situations (Fertuck et al. 2023). Appraisal then gives rise to behavioural manifestations, including avoidance, testing, controlling behaviours, or abrupt shifts between idealisation and mistrust.

Finally, trust learning reflects the capacity to update trust appraisals over time and adjust trust based on others' actual behaviour, which is critical for forming healthy relationships (Fineberg et al. 2018). This includes recognising cooperation and adapting trust accordingly (Fertuck et al. 2023).

When trust is not updated in light of new interpersonal evidence—whether through persistent mistrust or unwarranted overtrust—the learning process breaks down. Over time, impaired trust learning prevents the revision of beliefs and

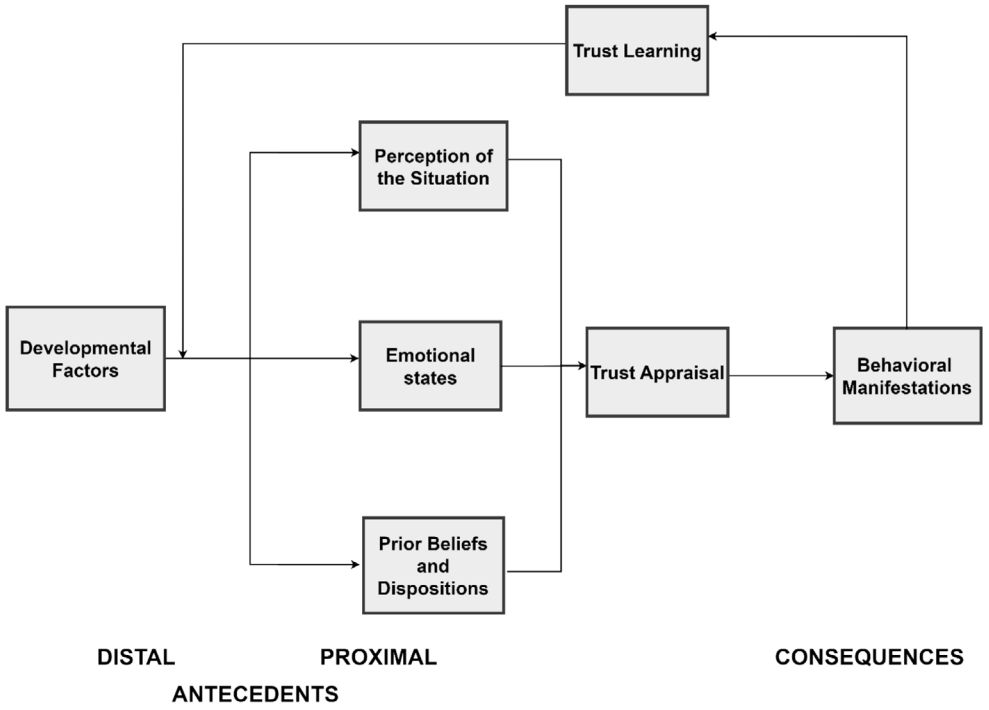

FIGURE 1 | A Model of Trust Processes in BPD (Adapted from Preti et al. 2023).

interpretations, reinforcing rigid relational expectations and hindering flexible engagement with others (Fertuck et al. 2023).

1.1 | Aims

The primary aim of this paper is to further develop a conceptual model of trust problems in individuals with BPD using clinician interviews. We interviewed clinicians with experience treating personality disorders, asking them to describe trust-specific challenges in BPD patients. The specific objectives include:

1. *Describe trust impairments*: Identify how individuals with BPD show difficulties in trust appraisal, decisions, and updating based on clinician reports.
2. *Characterise relational dysfunction*: Explore how trust disturbances lead to high-risk behaviour and fragile therapeutic relationships.
3. *Refine a conceptual framework*: Integrate developmental, social, cognitive and clinical insights into a multi-stage model of trust in BPD.
4. *Guide psychotherapy*: Offer foundations for targeted interventions that address trust issues and improve treatment retention.
5. *Inform future research*: Generate hypotheses for empirical work on how BPD trust processes differ from normative functioning.

By achieving these aims, this paper seeks to contribute to a deeper understanding of trust processes in BPD and to inform more nuanced and effective therapeutic interventions.

2 | Methods

2.1 | Participants

Participants were 25 Italian therapists ($M_{age} = 43 \pm 10.07$, range = 31–68; $N_{females} = 12$) with a specific interest in treating patients with personality disorders. Clinical experience ranged between 1 and 37 years ($M = 10.44$, SD = 10.88), with 2–100 patients treated ($M = 26.52$, SD = 26.43) in various settings (e.g., outpatients, inpatients, public mental health system). Five out of 25 were in training. Participants were trained in different therapeutic methods, with a predominance of psychodynamic psychotherapists. Such variability aligns with everyday clinical practice, where BPD patients are treated by clinicians with different levels of experience across diverse services. A full description of the participants can be found in Table S1.

2.2 | Procedure and Data Collection

Psychotherapists were contacted via email in May 2020 through a list of direct and indirect contacts. Those who agreed scheduled an interview. Twenty-three interviews were conducted via Zoom and two by phone, between May and July 2020. All clinicians consented to audio recording for *verbatim* transcription. Before each interview, participants were informed about the general topics to be discussed. At the end, those interested received further information about the broader research project. The study was approved by the local ethics committee.

The semi-structured ad hoc interview topic guide was developed based on the model of trust processes in BPD (Preti et al. 2023)

to gain insight into clinical experiences related to interpersonal trust in BPD patients. The interview guide (see Supporting Information S1) explored the areas of the trust processes model, that is, (1) a general introductory question about problems in trust appraisal; (2) developmental factors playing a role in others' trustworthiness; (3) specific situations reported by patients in which they had troubles or issues in trusting others; (4) specific emotional states that could have a role in attributing trust to others; (5) prior beliefs and dispositions playing a role in trust attribution; (6) issues in trust attribution that are connected to specific behaviours in the patient; (7) updating of trust attributions according to experience.

The interviews were conducted by two members of the research group, who received training in conducting and analysing qualitative data in an open and flexible manner, adapting the order of questions to the participants' responses. For each topic, interviewees were asked to recall situations reported by patients as well as examples and concrete experiences in the therapeutic relationship.

The interview guide was piloted through two test interviews. Interviews lasted an average of 45 min and were audio-recorded.

2.3 | Data Analysis

The data analysis was conducted according to Braun and Clarke's Thematic Analysis framework (Braun and Clarke 2006). In particular, we adopted a Theoretical Thematic Analysis based on a deductive or top-down approach where data are coded and interpreted using the lens of a specific theoretical model. No qualitative software was used, and coding was conducted manually following standard thematic analysis procedures. The identified themes closely aligned with the research questions and the study's aims guided the coding. Given our interest in trust attribution, we focused specifically on trust-related processes during coding.

After fully transcribing the interviews, we familiarised ourselves with the data through repeated readings. Two research team members generated codes deductively, based on the trust process model, while remaining open to integrating emergent aspects. Coding saturation was achieved after approximately 20 interviews; however, the analysis of the remaining interviews remained open to further specifications and nuances related to the already identified codes. Codes were refined into themes through group discussions to form concept clusters that both fit and enrich the theoretical model. A team member with expertise in qualitative methods supervised the process.

We acknowledge that the decision to adopt a specific theoretical framework to model the data analysis may be perceived as a methodological weakness. Nevertheless, the present study did not have a confirmatory objective, and—as will be discussed below—the emergence of new themes not included in the original model shows that the analysis was conducted in a way that remained adherent to the textual data, yet also in an open and flexible manner to permit the inclusion of new aspects that could potentially be integrated within the theory. The coding process was conducted with the objective of not restricting the discovery of novel insights outside the predefined theoretical lens (Karatsareas 2022). The theoretical framework provided the foundation for our analytical approach, which was informed by the structure of the trust processes model. However, this theoretical framework did not constrain the coding procedure or the conceptual categories that emerged from it. Illustrative quotes were selected for their representativeness and relevance. Data were analysed in the original language (Italian); quotations were translated into English and reviewed for accuracy. Participant anonymity was ensured using code numbers.

3 | Results

The analysis led to the identification of six themes (see Table 1), each reflecting different stages and components of the trust process. Notably, *Trust Appraisal*—the central construct in the Trust Processes Model—did not emerge as a standalone theme; rather, it is embedded throughout the data, as all interview material can be interpreted as specifying the difficulties BPD patients face in evaluating others' trustworthiness. Detailed examples of clinician quotes illustrating each theme are provided in Table S2. In this section, we report a subsample of quotes.

3.1 | Theme 1: Developmental Factors

Therapists described several developmental experiences shaping later trust ynamics in BPD. A first cluster involved *physical and sexual abuses*, often within the family:

"[...] the dad is involved with an extremely violent woman. The patient recounts that this woman used to physically abuse her and her sister by throwing objects and yelling at them. She still relives these experiences today, and when she feels criticized, she retrieves those memories" (23). Experiences of *abandonment* were also frequently mentioned as precursors of pervasive mistrust. Beyond overt trauma, clinicians emphasised the broader *quality of caregiving*—emotionally neglectful parenting, lack of recognition, inconsistent or ambiguous responses, family conflict, and role reversal—all seen as contributing to later difficulties in trusting others.

3.2 | Theme 2: Perception of the Situation

Clinicians highlighted how BPD patients often misinterpret social contexts involving trust. A first subtheme centers on situations involving three interrelated perceptions clearly associated with rejection sensitivity: *the perception of being ignored/excluded by others, the sense of being judged, and the fear of abandonment*.

> [therapist referring to the thoughts of a patient] When I see my university classmates talking among themselves, I might think, 'Oh, they're excluding me'. This feeling of discomfort then turns into anger, and I tend to interpret their behaviour more and more as, 'Look, they've left without involving me.'
>
> (25)

TABLE 1 | Themes emerged from the interview.

Developmental Factors
Physical and sexual abuses
Abandonment
Overall quality of caregiving
Perception of the Situation
Perception of being ignored/excluded, judged, abandoned
Rigid distinction between good people and bad people
Fluctuation between excessive trust and mistrust
Specificity of the relational situation
Emotions
Emotional antecedents
Emotional consequences
Prior beliefs and dispositions
Other as persecutor
Other person should satisfy rigidly all the patient's needs
Self-devaluation
Perfectionism
Deeply rooted mental schemas related to the disorder and its treatment
Distortion of communication signals
Behaviours
Consequences—avoidance
Consequences—destroying the relationship
Consequences—need to control
Consequences—demanding behaviours
Activation—aggression
Activation—social withdrawal
Activation—punishment
Trust learning
Overestimation of single negative episodes
Underestimation of positive interactions
Vicious circles

According to clinicians, individuals with BPD often interpret situations in which others gather without them as clear signs of exclusion, without considering alternative explanations. Therapists also noted that even minimal or momentary lapses in others' attention are experienced as evidence of carelessness or neglect. Social disapproval and criticism are frequently intolerable and felt as humiliating, reinforcing the expectation that social situations are inherently devaluing. Over time, this pattern fosters a pervasive loss of trust in people.

A second subtheme involves rigid, categorical interpretations of reality (*rigid distinction between good people and bad people*). People are often perceived as either absolutely bad and malevolent or with purely good intentions. Trust is granted only when the other is perceived as entirely benevolent—an expectation inconsistent with everyday social nuance. This dichotomy often produces persecutory fears, especially when self-disclosure raises concerns about privacy or betrayal. *Fluctuations between excessive trust and mistrust* were described as common and destabilising. Patients with BPD often swing from unjustified idealisation to abrupt devaluation, frequently triggered by minor misunderstandings, and these rapid shifts generate marked relational and emotional instability.

The *specificity of the relational situation* implicated was relevant in terms of trust attribution. Romantic relationships were described as particularly problematic due to fears of betrayal, jealousy, and defensive behaviours that create self-perpetuating mistrust. While close relationships confront patients with intimacy-related stress, superficial relationships are easier to maintain. When BPD patients encounter help or reciprocity, they often feel overwhelmed and withdraw. Family relationships are similarly difficult, as maintaining distance is hindered by rapid shifts between idealisation and devaluation. The same pattern appears in professional settings, especially in teamwork, which easily triggers embarrassment and fear of judgement. Finally, therapists identified the therapeutic relationship as a central trust-related context. Patients often struggle with its dual professional and human levels: session payment may be perceived as depreciation or exploitation, while pharmacological prescriptions can trigger paranoid fears of control or doubts about the clinician's confidence in their resources. At the same time, clinicians are frequently idealised as saviours expected to solve all problems, an unrealistic assumption that leads to disappointment and increases the risk of dropout, requiring ongoing efforts from therapists to maintain engagement.

3.3 | Theme 3: Emotions

Data analysis highlights the presence of both emotional antecedents and emotional consequences of the process of trust attribution. Intense emotional states are often indicators of mistrust: the stronger the emotion, the more it biases judgements about others' intentions. Shame and jealousy frequently act as antecedents precipitating abrupt drops in trust, while emotions such as anger and anxiety may function either as antecedents or consequences depending on the situation.

These *emotional consequences*, which are not included in the actual formulation of the model, do not influence trust processes but are an effect of the latter and are activated as a result of mistrust.

In the following example, the patient arrives angry due to an unrelated event, and this emotional activation leads her to perceive the clinician's interventions as inappropriate. Here, anger functions as an antecedent that shapes subsequent trust attribution:

> She came to the session extremely angry because, once again, her bike had been stolen. But even before I had said or thought of saying anything, my mere questioning was perceived by her as an attempt to formulate psychological hypotheses that she didn't want to hear. 'Now I'm angry about my bike, I don't want to hear any reflections [...] you're also being inappropriate at this moment when I'm angry.'
>
> (7)

In other cases, anger appears as a consequence of a relational episode involving trust, triggering aggressive acting-out. A similar dual role was observed for anxiety, which may precede distorted appraisals or result from feeling unprotected and unsupported by others.

3.4 | Theme 4: Prior Beliefs and Dispositions

Clinicians emphasised the distinction between a priori beliefs and how these beliefs distort social-signal processing, often producing biased interpretations of verbal and non-verbal cues. Prior beliefs are classifiable into beliefs about others and beliefs about oneself.

Patients frequently attribute *persecutory intentions* to others, perceiving behaviours as threatening, humiliating, or rejecting. This can range from mild suspiciousness to overt paranoia:

> She was convinced that there was a camera recording everything and that I had hidden it. I asked her where it was, she pointed to something I had on the bookshelf, and I told her to go see what it was. It was a rock, but she was absolutely convinced that it was a camera."
>
> (22)

Others are seen as harmful, intolerant of the patient's negative aspects, or unwilling to understand or help. This leads many patients to avoid sharing emotions, viewing disclosure as pointless or dangerous.

Another relevant theme regards the strong expectation that the *other person should meet and satisfy all the patient's needs and expectations*. Patients are often characterised by a sense of entitlement and low frustration tolerance, particularly when attention is not fully directed toward them. These demands collapse quickly when disconfirmed, prompting sharp declines in trust.

Concerning the beliefs about the self, the first theme corresponds to *self-devaluation*. In this respect, cognitive contents include the idea of having no values as people, being "unqualified" and inadequate for social relations, hopelessly destructive, unlucky, and incapable of controlling one's emotions. *Perfectionistic expectations* emerge as a defensive response to both self-perceived inadequacy and perceived persecution by others.

Clinicians also described *deeply rooted mental schemas related to the disorder and its treatment* (e.g., using the diagnosis to justify behaviours, believing change is impossible, or assuming the clinician has no genuine interest). This leads some patients to think that they do not need any treatment or that no therapy will be able to help them.

These beliefs often produce systematic *distortions in the interpretations of social communication signals, especially non-verbal ones*. Patients tend to overemphasise the significance of gestures, such as a laugh, a yawn or a gaze, interpreting them as clear signs of devaluation or disinterest. A paranoid thinking style reinforces self-referential interpretations, especially when decoding ambiguous cues, such as response timing or silence.

3.5 | Theme 5: Behaviours

Clinicians described two behavioural categories: behavioural consequences—those predicted by the model and following trust misattribution—and behavioural activations, which stem from the emotional arousal caused by mistrust. Behavioural consequences are of extreme importance for therapists, as they allow them to identify the enactment of dysfunctional dynamics and intervene early in specific problematic interactions.

A consequence of distrust is *avoidance*, expressed through withdrawal from relationships, emotional closure, fear of intimacy, or sudden elimination of ambivalent relationships:

> Blocking on all social media. The patient keeps blocking people, and I ask her, 'Why did you block them'? It's not clear, but 'they have to disappear'.
>
> (8)

A second consequence of distrust is trying to *destroy the relationship*, particularly within therapy, through inauthentic presence or attacks on the therapeutic setting.

Other trust-related behaviours include *controlling strategies*, such as compulsive seeking of information, constantly questioning the other person to test him/her, and manipulative attitudes. Linked to this is a strong need for reassurance, often expressed as demanding or intrusive requests for proof of helpfulness. Linked to the need for control is the *demanding behaviour* of BPD patients, as they tend to require, in invasive and compelling ways, continuous proof of the other's helpfulness and availability. Another signal of trust breakdown is the *denial of dependence*. Patients may devalue or blame the other to avoid feeling dependent—for instance, insisting on paying for each session as a way to signal emotional detachment.

Regarding behavioural activation, that is, behaviours that emerge as effects of the negative emotional activation due to trust misattribution, a first aspect is linked to *aggression*, both in terms of self-destructive behaviour and externally directed aggression, alongside *social withdrawal*, such as running away from conflicts, dropping out of therapy, or abruptly quitting jobs. The last category falling under the behavioural effects is related to the *punishment* of

others. BPD patients often react to confrontations or breakdown of trust, punishing the other person for what he/she has done.

3.6 | Theme 6: Trust Learning

Trust learning refers to the capability of updating one's own opinion about the possibility of trusting other people. The therapeutic relationship has the potential to make patients change their ideas about trustworthiness on the basis of therapeutic experience and information derived from it.

However, clinicians' responses showed that BPD patients have great difficulty in readjusting their impression of others' trustworthiness by including new data coming from relational experiences.

The first attributional bias that complicates this process is the tendency to *overestimate the significance of a single negative episode* when deciding whether to trust another person. This means that a single situation in which someone is perceived as lacking in the patient's regard is enough to undermine confidence.

> It was enough that one day it snowed, and there was 50 cm of snow, and his parents couldn't come to the community to visit him, and everything would start all over again, and he would say, 'See, they don't want me'. (12)

Together with the overestimation of negative episodes, clinicians frequently encounter the tendency to *underestimate positive interactions*, followed by a lack of updating patterns of trust attribution. To be taken into serious consideration, any expression of genuine interest by others has to be repeated consistently and for a long time; otherwise, it is downplayed as accidental and fortuitous.

Another barrier is the tendency to act out one's own fears. Patients may behave in ways that provoke other people to react exactly the way they fear, establishing a *vicious cycle* in which what is feared is provoked, confirming that suspiciousness and distrust were fully justified. By making impossible demands or setting unrealistic expectations, patients often elicit disappointment from others, reinforcing distrust and creating a self-confirming cycle.

4 | Discussion

4.1 | The Theoretical Model of Trust Impairments: Insights From a Clinical Perspective

The present findings should be viewed as clinically informed refinements of the trust-processing model introduced by Preti et al. (2023), adding clinical depth to its components through therapists' observations. The model suggests how trust impairments in BPD unfold across developmental, cognitive, emotional and behavioural domains, as consistently described by clinicians.

Developmental Factors highlight the influence of early trauma—such as abuse, abandonment, and poor caregiving—on trust development in BPD. This aligns with prior research (e.g., Fertuck et al. 2023; Preti et al. 2023) and theoretical accounts linking trust issues in adulthood to disrupted attachment (Ociskova et al. 2023) and distorted object relations (Kernberg 1967). This also parallels attachment-based conceptualisations in which early relational insecurity disrupts expectations of safety and availability, shaping chronic distrust—an interpersonal signature strongly associated with BPD (e.g., Euler et al. 2021; Lazarus et al. 2014), where trauma and disorganised attachment patterns are highly prevalent (from 30% up to 90%; Zanarini et al. 2006). Moreover, the results are consistent with Schema Therapy accounts describing how early experiences of abuse, abandonment, or emotional deprivation consolidate mistrust/abuse and abandonment schemas that later structure interpersonal expectations and frequently converge in the interpersonal instability typical of BPD (Kellogg and Young 2006).

Prior Beliefs and Dispositions reflect therapists' observations of a dichotomous thinking in BPD, with patients oscillating between idealisation and mistrust based on distorted self-other beliefs. This binary perception aligns with the model's depiction of how pre-existing beliefs and emotional states skew trust appraisal processes (Fertuck et al. 2023; Preti et al. 2023). Moreover, distortions of perceptions and communications seem particularly central in this process: BPD individuals tend to misunderstand and misattribute the meaning of others' behaviours, being impulsive, hostile, and having a biased negative evaluation of their partner's emotional expressions (Ociskova et al. 2023). This is also confirmed by neurobiological alterations in the amygdala and hippocampus, which generate distorted and frightening interpretations of social signals (Mainali et al. 2020).

Importantly, these proximal components of the model are not only rooted in distal developmental experiences but are also shaped by situational factors that operate in the present moment. Indeed, beyond a rigid pattern of splitting, where others are perceived as entirely trustworthy or untrustworthy (Kernberg 1967), and *'black and white'* thinking (Ociskova et al. 2023), the specific nature of different relationships emerged as a crucial factor in how trust issues manifest. This finding suggests that trust difficulties in BPD are not uniform across all relationships but rather are shaped by the unique demands and intimacy levels of each relational context, such that each context triggers specific trust-related concerns. In this framework, early trauma creates a vulnerability that is continually reactivated by current interpersonal stressors (Gunderson and Lyons-Ruth 2008), which modulate fluctuations in trust in everyday interactions and within the therapeutic relationship itself. Thus, distrust in BPD cannot be understood as a solely dispositional phenomenon: momentary interpersonal cues, perceived slights, and shifts in relational attunement also act as proximal triggers that amplify or dampen distrust.

Within this framework, the therapeutic relationship itself functions as an active proximal factor. Moment-to-moment therapist attunement may soothe or reignite the patient's expectancy of rejection (Fonagy et al. 2018), influencing trust appraisal processes. Notably, this dynamic emerged only marginally in our

interviews, possibly reflecting therapists' reluctance to consider themselves as active contributors to the activation of (pathological) relational processes characteristic of BPD.

Emotional Antecedents and Consequences were also critical. The model's emphasis on intense emotions (e.g., shame, jealousy) affecting trust is confirmed by clinicians. However, our study also revealed an additional layer not explicitly included in the original model—emotional consequences of trust misattribution. These emotional aftereffects, such as subsequent anger or anxiety, were noted by therapists as outcomes of the trust process rather than direct influencers of trust appraisal. This adds a layer of complexity: emotional fallout can perpetuate distrust and destabilise relationships. Given BPD individuals' heightened emotional reactivity (Mainali et al. 2020), this cycle of dysregulation and mistrust may become self-reinforcing, where emotional dysregulation influences trust appraisal, which, in turn, generates intense emotional responses that further compromise trust-related decisions, creating a cascade effect that perpetuates trust difficulties.

Behavioural Aspects extended the model by distinguishing behavioural consequences of trust misattribution—such as avoidance and relationship destruction (Ociskova et al. 2023) from a new category: behavioural activation, which encompasses behaviours stemming from the emotional arousal caused by trust misattribution. These immediate responses stem from emotional arousal rather than deliberate mistrust, echoing previous findings (Wrege et al. 2019). Intense emotional states (such as anger and anxiety) lead to specific behavioural responses (such as aggression, social withdrawal and punishment). Furthermore, this behavioural activation pathway appears to operate in parallel with the more deliberate behavioural consequences outlined in the original model. While behavioural consequences represent more planned responses to trust appraisal (such as conscious relationship-maintaining or destroying behaviours), behavioural activations emerge as more immediate, emotion-driven reactions. This distinction helps explain conflicting behaviours, where planned relationship-preserving efforts coexist with impulsive destructiveness.

The model's structure suggests a feedback loop through trust learning, where these behavioural patterns, influenced by emotional states, contribute to the formation of future trust-related experiences and beliefs. This cyclical nature, paired with inflexibility to update their trust appraisal, an overestimation of negative experiences, and an underestimation of positive interactions, helps explain the persistence of maladaptive trust patterns in BPD as emotional-behavioural sequences become reinforced through repeated experiences. For instance, BPD patients feel lower social connection even when accepted by others, reflecting deep-rooted disruptions in social belonging and trust (De Panfilis et al. 2015). BPD patients' altered social signal processing may create a persistent barrier to the natural updating mechanism of trust learning. This difficulty in updating trust also aligns with the notion of epistemic (dis)trust, whereby early relational unreliability weakens the capacity to treat social information as trustworthy and relevant, thereby contributing to the pervasive interpersonal suspicion typical of BPD (Nolte et al. 2023; Sharp and Balzen 2025).

4.2 | New Developments to the Model

Our findings offer two major clinically grounded refinements to the model (see Figure 2). First, emotional consequences—emotions experienced as a result of trust misattribution—provide a new dimension to understanding trust dynamics. Individuals with BPD tend to overreact both to neutral and negative stimuli, especially with anger and anxiety (Jacob et al. 2008). These reactions do not directly affect trust judgements but exacerbate relational instability and may trigger further maladaptive responses. Incorporating therapeutic techniques that address these emotional reactions (Transference-Focused Psychotherapy, TFP, Yeomans et al. 2015; Mentalization-Based Therapy, MBT, Bateman and Fonagy 2016; Dialectical Behaviour Therapy, DBT, Linehan 1993) may help patients better regulate them and prevent behavioural escalation.

Second, behavioural activation expands the model's behavioural dimension. Emotionally triggered behaviours—like aggression or withdrawal—reflect the interplay between arousal and impulsive action. Accordingly, emotional and behavioural dysregulation in individuals with BPD is fundamentally interconnected through "*emotional cascades*", suggesting that BPD individuals engage in dysregulated behaviours, such as non-suicidal self-injury and aggression, to distract from intolerable emotional states (Selby et al. 2008). Recognising this mechanism may help clinicians better anticipate and manage these reactions in therapy.

It is also important to address the absence of the original model's stage of trust appraisal in our thematic findings. While the specific stage of trust appraisal was not explicitly highlighted in our results, it is evident that the themes identified reflect issues related to the core of trust assessment. This suggests that the difficulties in trust appraisal may be more implicitly woven into the broader patterns of trust dynamics rather than manifesting as a distinct, standalone stage.

4.3 | Strengths and Implications

The study further extends Preti et al.'s (2023) theoretical model by providing a clinically grounded approach that emphasises the critical role of early trauma, emotional dysregulation, and distorted social cognition in BPD trust dynamics. First, the findings provide a comprehensive understanding of trust dynamics in BPD: clinicians should recognise that trust issues in BPD stem from early developmental trauma and attachment disruptions, manifest differently across various relationship types, are influenced by rigid, dichotomous thinking patterns, and include both emotional and behavioural components.

By corroborating the role of developmental factors, the results reinforce the importance of early intervention targets for BPD (Sharp and Fonagy 2015). In fact, this is especially crucial given the typical difficulty individuals with BPD face in adapting their perspectives in response to new social information. Moreover, results suggest that therapeutic interventions should address both the core components of trust impairments and the secondary emotional effects, behavioural responses, and impulsive

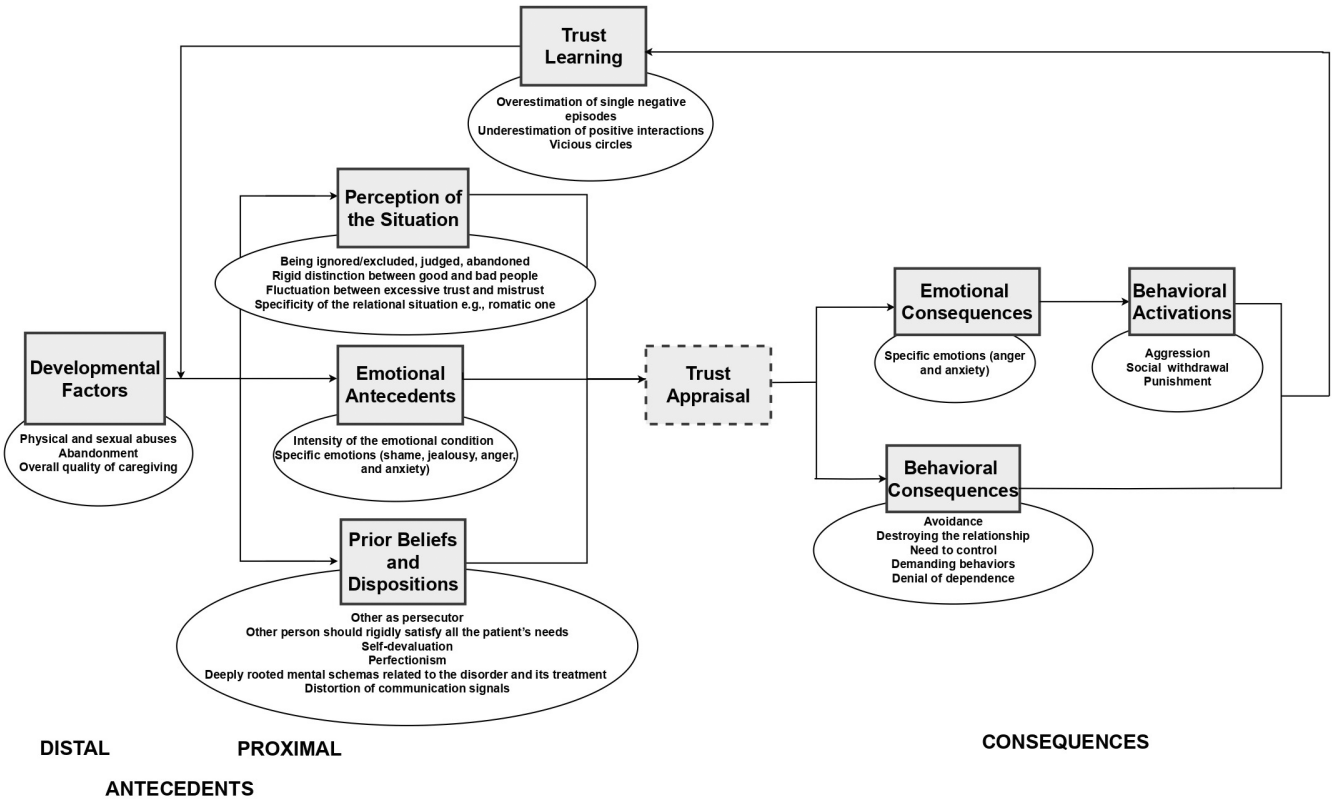

FIGURE 2 | Updated model of trust processes in BPD.

actions. Focusing on behavioural and emotional primary and secondary features may reduce emotional cascades that lead to maladaptive behavioural activation and impaired trust learning. Importantly, clinicians should help patients interpret social signal processing, challenging their distorted perceptions of relationships and integrating positive relational experiences.

Therapists working with individuals with BPD are advised to anticipate and prepare for trust-related challenges in therapy, maintain consistency in therapeutic boundaries, and address trust issues in the therapeutic relationship by deeply understanding and mirroring the patient's subjective experience. The practice of thoughtfully articulating and acknowledging the patient's experiences may, in itself, serve as a therapeutic tool to enhance trust, learning, and ease distrust, potentially enhancing the effectiveness of the therapeutic alliance. Hence, the therapeutic relationship may serve as a model for trust-building in BPD treatment, where patients can safely experience, test, and gradually modify their trust attributions through consistent and safe interactions with the therapist. On the other hand, clinicians are suggested to pay attention to aggressive or self-destructive behaviours and overall relationship sabotage patterns in response to trust issues with the therapist and to address them in therapy.

4.4 | Limitations and Future Directions

Several limitations must be acknowledged. First, relying solely on clinician reports may introduce bias, privileging therapists' interpretations over patients' lived experiences (Paterson et al. 2023). In addition, the qualitative nature of the study does not allow for the formulation of a new, data-driven model; rather, our contribution lies in refining and extending elements of the existing trust-processing framework. Moreover, therapists' focus on theoretical aspects risks emphasising theoretical viewpoints rather than the nuanced realities therapists face. Future research should include direct patient input and explore perspectives from therapists with diverse theoretical orientations. The variability in clinicians' experiences represents a limitation, as it may have influenced the depth of some observations. However, only a minority of participants had limited experience, and such heterogeneity nonetheless reflects the routine clinical practice in BPD care. Flexibility in interview structure may have led to uneven coverage across participants (Harvey-Jordan and Long 2001). Moreover, most clinicians were psychodynamically trained, emphasising early development and unconscious processes. This theoretical homogeneity may have biased interpretations and overlooked perspectives more central to other approaches. Broader inclusion of theoretical frameworks and international samples would enhance generalisability and depth.

5 | Conclusion

In summary, this study extends the theoretical model of trust impairments in BPD, highlighting the critical role of developmental factors, rigid perceptions, emotional influences, and distorted beliefs. The introduction of new concepts—emotional consequences and behavioural activation—provides a more nuanced understanding of trust dynamics and offers practical implications for therapeutic interventions. Addressing these expanded dimensions can enhance manualised treatments for BPD, such as TFP, MBT and DBT, by integrating a more comprehensive approach to managing trust-related issues in patients.

Acknowledgements

The authors have nothing to report.

Funding

The authors have nothing to report.

Ethics Statement

This study was reviewed and approved by the local commission for minimal-risk studies of the Department of Psychology (protocol n. RM-2021-439). Subjects read the study description and signed the informed consent sheet prepared in accordance with the current version of the Declaration of Helsinki.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The datasets used during the current study are available from the corresponding author on reasonable request.

References

Agrawal, H. R., J. Gunderson, B. M. Holmes, and K. Lyons-Ruth. 2004. "Attachment Studies With Borderline Patients: A Review." *Harvard Review of Psychiatry* 12, no. 2: 94–104. https://doi.org/10.1080/10673220490447218.

American Psychiatric Association. 2022. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5-TR)*. American Psychiatric Association Publishing. https://doi.org/10.1176/appi.books.9780890425787.

Bateman, A., and P. Fonagy. 2016. *Mentalization-Based Treatment for Personality Disorders: A Practical Guide*, 468. Oxford University Press. https://doi.org/10.1093/med:psych/9780199680375.001.0001.

Braun, V., and V. Clarke. 2006. "Using Thematic Analysis in Psychology." *Qualitative Research in Psychology* 3, no. 2: 77–101. https://doi.org/10.1191/1478088706qp063oa.

D'Agostino, A., A. Aportone, M. Petrini, E. Manganelli, M. Rossi Monti, and V. Starcevic. 2018. "Preliminary Validation of the Interpersonal Dysphoria Model of Borderline Personality Disorder." *Psychopathology* 51, no. 6: 390–399. https://doi.org/10.1159/000495843.

De Panfilis, C., P. Riva, E. Preti, C. Cabrino, and C. Marchesi. 2015. "When Social Inclusion Is Not Enough: Implicit Expectations of Extreme Inclusion in Borderline Personality Disorder." *Personality Disorders, Theory, Research, and Treatment* 6, no. 4: 301–309. https://doi.org/10.1037/per0000132.

Euler, S., T. Nolte, M. Constantinou, et al. 2021. "Interpersonal Problems in Borderline Personality Disorder: Associations With Mentalizing, Emotion Regulation, and Impulsiveness." *Journal of Personality Disorders* 35, no. 2: 177–193. https://doi.org/10.1521/pedi_2019_33_427.

Fertuck, E. A., S. Fischer, and J. Beeney. 2018. "Social Cognition and Borderline Personality Disorder: Splitting and Trust Impairment Findings." *Psychiatric Clinics of North America* 41, no. 4: 613–632. https://doi.org/10.1016/j.psc.2018.07.003.

Fertuck, E. A., J. Grinband, J. J. Mann, et al. 2018. "Trustworthiness Appraisal Deficits in Borderline Personality Disorder Are Associated With Prefrontal Cortex, Not Amygdala, Impairment." *NeuroImage: Clinical* 21: 101616. https://doi.org/10.1016/j.nicl.2018.101616.

Fertuck, E. A., and E. Preti. 2023. "Interpersonal Trust and Borderline Personality Disorder: Insights From Clinical Practice and Research: Introduction." *Journal of Personality Disorders* 37, no. 5: 469–474. https://doi.org/10.1521/pedi.2023.37.5.469.

Fertuck, E. A., E. Preti, and J. F. Clarkin. 2023. "Transference-Focused Psychotherapy and Trust Processing in BPD: Exploring Possible Mechanisms of Change." *Journal of Personality Disorders* 37, no. 5: 620–632. https://doi.org/10.1521/pedi.2023.37.5.620.

Fineberg, S. K., J. Leavitt, D. S. Stahl, et al. 2018. "Differential Valuation and Learning From Social and Nonsocial Cues in Borderline Personality Disorder." *Biological Psychiatry* 84, no. 11: 838–845. https://doi.org/10.1016/j.biopsych.2018.05.020.

Fonagy, P., and C. Campbell. 2017. "Mentalizing, Attachment and Epistemic Trust: How Psychotherapy Can Promote Resilience." *Psychiatria Hungarica* 32, no. 3: 283–287.

Fonagy, P., G. Gergely, and E. L. Jurist. 2018. *Affect Regulation, Mentalization and the Development of the Self*. Routledge. https://doi.org/10.4324/9780429471643.

Gabbard, G. O. 1993. "An Overview of Countertransference With Borderline Patients." *Journal of Psychotherapy Practice and Research* 2, no. 1: 7–18.

Gunderson, J. G., and K. Lyons-Ruth. 2008. "Bpd's Interpersonal Hypersensitivity Phenotype: A Gene-Environment-Developmental Model." *Journal of Personality Disorders* 22, no. 1: 22–41. https://doi.org/10.1521/pedi.2008.22.1.22.

Harvey-Jordan, S., and S. Long. 2001. "The Process and the Pitfalls of Semi-Structured Interviews." *Community Practitioner* 74, no. 6: 219.

Iliakis, E. A., G. S. Ilagan, and L. W. Choi-Kain. 2021. "Dropout Rates From Psychotherapy Trials for Borderline Personality Disorder: A Meta-Analysis." *Personality Disorders* 12, no. 3: 193–206. https://doi.org/10.1037/per0000453.

Jacob, G. A., C. Guenzler, S. Zimmermann, et al. 2008. "Time Course of Anger and Other Emotions in Women With Borderline Personality Disorder: A Preliminary Study." *Journal of Behavior Therapy and Experimental Psychiatry* 39, no. 3: 391–402. https://doi.org/10.1016/j.jbtep.2007.10.009.

Jayawickreme, E., W. Fleeson, E. D. Beck, A. Baumert, and J. M. Adler. 2021. "Personality Dynamics." *Personality Science* 2: e6179. https://doi.org/10.5964/ps.6179.

Karatsareas, P. 2022. "Semi-Structured Interviews." In *Research Methods in Language Attitudes*, edited by R. Kircher and L. Zipp, 1st ed., 99–113. Cambridge University Press. https://doi.org/10.1017/9781108867788.010.

Kaufman, E. A., J. Perez, S. Lazarus, S. D. Stepp, and S. L. Pedersen. 2020. "Understanding the Association Between Borderline Personality Disorder and Alcohol-Related Problems: An Examination of Drinking Motives, Impulsivity, and Affective Instability." *Personality Disorders, Theory, Research, and Treatment* 11, no. 3: 213–221. https://doi.org/10.1037/per0000375.

Kellogg, S. H., and J. E. Young. 2006. "Schema Therapy for Borderline Personality Disorder." *Journal of Clinical Psychology* 62, no. 4: 445–458. https://doi.org/10.1002/jclp.20240.

Kernberg, O. 1967. "Borderline Personality Organization." *Journal of the American Psychoanalytic Association* 15, no. 3: 641–685. https://doi.org/10.1177/000306516701500309.

Kockler, T. D., P. S. Santangelo, M. Eid, et al. 2022. "Self-Esteem Instability Might Be More Characteristic of Borderline Personality Disorder Than Affective Instability: Findings From an e-Diary Study With Clinical and Healthy Controls." *Journal of Psychopathology and Clinical Science* 131, no. 3: 301–313. https://doi.org/10.1037/abn0000731.

Lazarus, S. A., J. S. Cheavens, F. Festa, and M. Zachary Rosenthal. 2014. "Interpersonal Functioning in Borderline Personality Disorder: A Systematic Review of Behavioral and Laboratory-Based Assessments." *Clinical Psychology Review* 34, no. 3: 193–205. https://doi.org/10.1016/j.cpr.2014.01.007.

Lewis, J. D., and A. Weigert. 1985. "Trust as a Social Reality." *Social Forces* 63, no. 4: 967–985. https://doi.org/10.1093/sf/63.4.967.

Linehan, M. M. 1993. *Cognitive-Behavioral Treatment of Borderline Personality Disorder*, xvii–558. Guilford Press.

Mainali, P., T. Rai, and I. H. Rutkofsky. 2020. "From Child Abuse to Developing Borderline Personality Disorder Into Adulthood: Exploring the Neuromorphological and Epigenetic Pathway." *Cureus* 12, no. 7: e9474. https://doi.org/10.7759/cureus.9474.

Masland, S. R., S. E. Schnell, and T. V. Shah. 2020. "Trust Beliefs, Biases, and Behaviors in Borderline Personality Disorder: Empirical Findings and Relevance to Epistemic Trust." *Current Behavioral Neuroscience Reports* 7, no. 4: 239–249. https://doi.org/10.1007/s40473-020-00220-7.

Miskewicz, K., W. Fleeson, E. M. Arnold, M. K. Law, M. Mneimne, and R. M. Furr. 2015. "A Contingency-Oriented Approach to Understanding Borderline Personality Disorder: Situational Triggers and Symptoms." *Journal of Personality Disorders* 29, no. 4: 486–502. https://doi.org/10.1521/pedi.2015.29.4.486.

Morales-Muñoz, I., B. B. Durdurak, A. Bilgin, S. Marwaha, and C. Winsper. 2021. "Understanding the Relationship Between Sleep Problems in Early Childhood and Borderline Personality Disorder: A Narrative Review." *Nature and Science of Sleep* 13: 2175–2202. https://doi.org/10.2147/NSS.S311672.

Nolte, T., J. Hutsebaut, C. Sharp, C. Campbell, P. Fonagy, and A. Bateman. 2023. "The Role of Epistemic Trust in Mentalization-Based Treatment of Borderline Psychopathology." *Journal of Personality Disorders* 37, no. 5: 633–659. https://doi.org/10.1521/pedi.2023.37.5.633.

Ociskova, M., J. Prasko, F. Hodny, et al. 2023. "Black & White Relations: Intimate Relationships of Patients With Borderline Personality Disorder." *Neuroendocrinology Letters* 44, no. 5: 321–331.

Paterson, C., C. Roberts, and K. Bail. 2023. "'Paper Care Not Patient Care': Nurse and Patient Experiences of Comprehensive Risk Assessment and Care Plan Documentation in Hospital." *Journal of Clinical Nursing* 32, no. 3–4: 523–538. https://doi.org/10.1111/jocn.16291.

Preti, E., J. Richetin, A. Poggi, and E. Fertuck. 2023. "A Model of Trust Processes in Borderline Personality Disorder: A Systematic Review." *Current Psychiatry Reports* 25, no. 11: 555–567. https://doi.org/10.1007/s11920-023-01468-y.

Richardson-Vejlgaard, R., C. Broudy, B. Brodsky, E. Fertuck, and B. Stanley. 2013. "Predictors of Psychotherapy Alliance in Borderline Personality Disorder." *Psychotherapy Research: Journal of the Society for Psychotherapy Research* 23, no. 5: 539–546. https://doi.org/10.1080/10503307.2013.801001.

Selby, E. A., M. D. Anestis, and T. E. Joiner. 2008. "Understanding the Relationship Between Emotional and Behavioral Dysregulation: Emotional Cascades." *Behaviour Research and Therapy* 46, no. 5: 593–611. https://doi.org/10.1016/j.brat.2008.02.002.

Shapiro, S. P. 1987. "The Social Control of Impersonal Trust." *American Journal of Sociology* 93, no. 3: 623–658. https://doi.org/10.1086/228791.

Sharp, C., and K. Balzen. 2025. "Trust in Borderline Personality Disorder." In *Handbook of Trust and Social Psychology*, 270–287. Edward Elgar Publishing. https://www.elgaronline.com/edcollchap/book/9781803929415/chapter14.xml.

Sharp, C., and P. Fonagy. 2015. "Practitioner Review: Borderline Personality Disorder in Adolescence—Recent Conceptualization, Intervention, and Implications for Clinical Practice." *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 56, no. 12: 1266–1288. https://doi.org/10.1111/jcpp.12449.

Wrege, J. S., A. C. Ruocco, S. Euler, et al. 2019. "Negative Affect Moderates the Effect of Social Rejection on Frontal and Anterior Cingulate Cortex Activation in Borderline Personality Disorder." *Cognitive, Affective, & Behavioral Neuroscience* 19, no. 5: 1273–1285. https://doi.org/10.3758/s13415-019-00716-0.

Yeomans, F. E., J. F. Clarkin, and O. F. Kernberg. 2015. *Transference-Focused Psychotherapy for Borderline Personality Disorder: A Clinical Guide*. 1° edizione ed. Amer Psychiatric Pub Inc.

Zanarini, M. C., F. R. Frankenburg, J. Hennen, D. B. Reich, and K. R. Silk. 2006. "Prediction of the 10-Year Course of Borderline Personality Disorder." *American Journal of Psychiatry* 163, no. 5: 827–832. https://doi.org/10.1176/ajp.2006.163.5.827.

Zhao, M., Y. Li, J. Lin, et al. 2024. "The Relationship Between Trust and Well-Being: A Meta-Analysis." *Journal of Happiness Studies* 25, no. 5: 56. https://doi.org/10.1007/s10902-024-00737-8.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Table S1:** Clinicians' demographic information. **Table S2:** Themes and subthemes with Clinicians' Interviews Quotes.