HHS Public Access

Author manuscript

Curr Opin Psychol. Author manuscript; available in PMC 2019 June 01.

Published in final edited form as:

Curr Opin Psychol. 2018 June ; 21: 94–104. doi:10.1016/j.copsyc.2017.09.011.

Interpersonal dysfunction in borderline personality: a decision neuroscience perspective

Michael N. Hallquist[1], **Nathan T. Hall**[1], **Alison M. Schreiber**[1], and **Alexandre Y. Dombrovski**[2]

[1]Department of Psychology, Penn State University

[2]Department of Psychiatry, University of Pittsburgh

Abstract

Borderline personality disorder (BPD) is characterized by disadvantageous decisions that are often expressed in close relationships and associated with intense negative emotions. Although functional neuroimaging studies of BPD have described regions associated with altered social cognition and emotion processing, these correlates do not inform an understanding of *how* brain activity leads to maladaptive choices. Drawing on recent research, we argue that formal models of decision-making are crucial to elaborating theories of BPD that bridge psychological constructs, behavior, and neural systems. We propose that maladaptive interactions between Pavlovian and instrumental influences play a crucial role in the expression of interpersonal problems. Finally, we articulate specific hypotheses about how clinical features of BPD may map onto neural systems that implement separable decision processes.

Introduction

Borderline personality disorder (BPD) is defined by emotion dysregulation, interpersonal dysfunction, and impulsivity [1]. Many problematic behaviors in BPD (e.g., suicide attempts) occur in response to negative interpersonal events such as perceived abandonment [2,3]. Clinical theory and experimental data indicate that individuals with BPD are interpersonally hypersensitive, often distorting social cues, forming extreme opinions of others, and making negative attributions about others' actions and even facial expressions [4–6]. These biases are linked with alterations in brain function, including heightened amygdala responses to negative emotional stimuli [7,8] and greater dorsal anterior cingulate cortex (dACC; BA 24/32) and medial prefrontal cortex (mPFC; BA 10) activity to social rejection [9]. More generally, regions involved in emotion regulation and reward processing are affected, including subgenual anterior cingulate cortex (sgACC; BA 25), medial

Corresponding author: Michael Hallquist, Department of Psychology, Penn State University, 140 Moore Building, University Park, PA 16802. michael.hallquist@psu.edu.

Conflict of interest statement

None

orbitofrontal cortex/ventromedial prefrontal cortex (mOFC/vmPFC; BA 10–14 [10]), and the amygdala [11].

This literature has informed the perspective that interpersonal hypersensitivity and emotion dysregulation in BPD reflect the predominance of ascending limbic influences over networks supporting social cognition and self-regulation [12]. However, the notion that emotions interfere with self-regulation and deliberation in BPD [13,14] provides few specific hypotheses about the neural underpinnings of salient features such as aggression [15] or rejection sensitivity [16]. In part, this reflects the absence of a testable theory linking BPD symptoms, underlying neurocognitive processes, and neural representation (see Box 1). Moreover, several other mental disorders (e.g., depression) prominently feature abnormalities in frontolimbic circuits [17], further undermining the neurobiological specificity of this perspective.

Box 1

Levels of analysis in borderline personality research

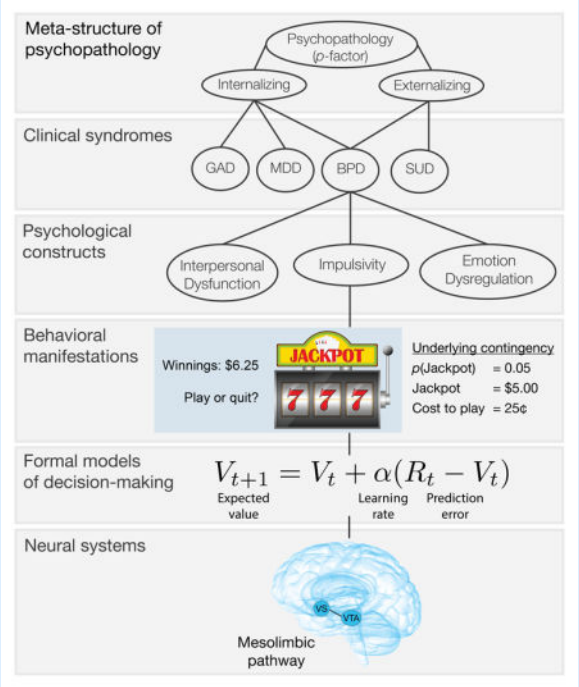
Empirical research on the modern construct of borderline personality disorder (BPD) emerged from the neo-Kraepelinian descriptive psychiatry movement of the 1970s [86]. Drawing on psychodynamic theories of BPD [27], scientists articulated a view of BPD as a *clinical syndrome* with specific symptoms (e.g., intense, unstable relationships) that could be diagnosed reliably by clinicians [87]. Although there have been a few revisions to the diagnostic criteria for BPD over the past four decades, the DSM-5 [88] retained the definition of the DSM-IV, requiring that five of nine criteria be present for a diagnosis.

Problems with boundary overlap and co-occurrence among personality disorders (PDs) [89] have motivated an alternative view that reframes borderline personality in terms of dimensional variation on a number of *psychological constructs* such as impulsivity [90]. Indeed, proponents of a dimensional view question the diagnostic utility of BPD [91]. Furthermore, in order to develop an empirically based taxonomy of psychopathology, it is essential to consider borderline personality vis-à-vis other PDs and mental disorders. For example, in analyses of the *meta-structure of psychopathology*, Eaton and colleagues [92] found that borderline personality had both internalizing and externalizing aspects. These two superordinate dimensions explain covariation among many mental disorders [93] and may be nested under a general psychopathology (aka '*p*') liability factor [94]. Indeed, Sharp and colleagues [95] found that BPD symptoms were largely explained by a general personality pathology factor.

Debates about the latent structure and ontological status of borderline personality arise in part from the observational nature of interview and self-report methods. Experimentation is essential to elaborate a mechanistic account of cognitive processes and neural abnormalities that underlie the *behavioral manifestations* of borderline personality. We focus specifically on disadvantageous decision-making such as desperately seeking intimacy with a partner even when it causes the partner to withdraw emotionally. Early experimental research on decision-making used paradigms such as the Iowa Gambling Task (IGT) to compare summary statistics (e.g., number of risky bets) between patient and control groups. This work identified abnormalities across various mental disorders,

but no disorder specificity [38]. Crucially, summary statistics provide limited insight into the cognitive processes underlying each behavioral output during the experiment (e.g., a single decision to gamble). As detailed in the main text, *formal models of decision-making*, such as reinforcement learning (Box 3) offer a generative account of how underlying cognitive processes lead to specific choices.

Formal models provide a crucial bridge linking latent decision processes with the dynamics of underlying *neural systems* [20] measured at the level of neuronal populations (e.g., fMRI or EEG). This is in contrast to the oversimplified view of a one-to-one correspondence between psychological constructs and neural systems (e.g., the neural correlates of neuroticism), which leaps across multiple levels of analysis and has not substantially advanced theories of personality disorders [96,97]. Formal models instantiate testable hypotheses about the latent processes underlying decision-making that are proximate to specific behavioral outputs. The structure of such models mirrors the organization of neural populations into input-output units that are responsible for processing specific information. While we see value in studying borderline personality at many levels of analysis, we believe that formal models decision-making paired with functional neuroimaging can provide new insights into BPD [for a cogent discussion of the importance of behavior in neuroscience, see 98].



Box 2

Using formal models to elucidate cognitive representation and neural systems

Over the past three decades, cognitive neuroscientists have used functional neuroimaging to describe neural activity elicited by stimuli presented sequentially during an experiment. In traditional fMRI paradigms, regional activation to one type of stimulus (e.g., faces) can be compared to another stimulus (e.g., houses) or to no task (also called the resting baseline). For example, a region involved in reward processing may exhibit greater activity to the receipt of a reward than at rest. Such analyses identify regions that are sensitive to differences in experimental conditions (e.g., faces compared to houses), but they do not necessarily elucidate the underlying cognitive architecture or its neural implementation.

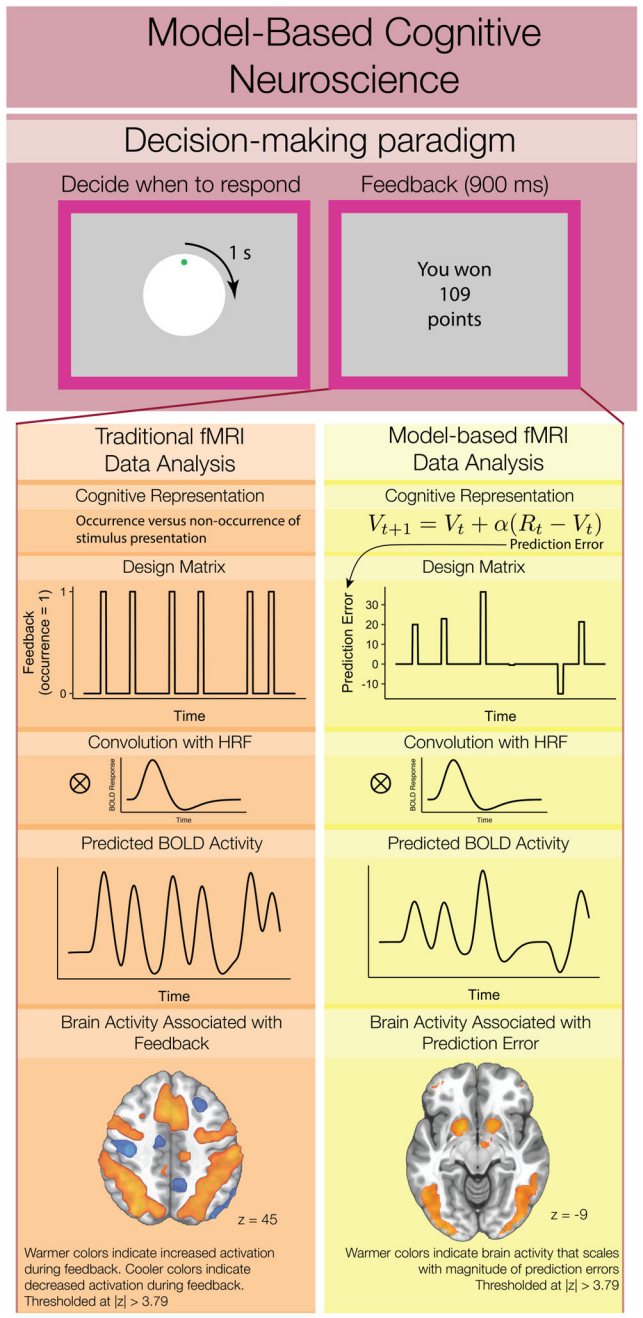
By contrast, formal, algorithmic models of behavior seek to predict what an individual will do. In such models, equations and quantitative parameters specify an explicit mapping between latent cognitive processes and task-relevant behavior. Individual differences in parameter values may reflect intermediate phenotypes that are associated with broader psychological constructs.

The use of formal models for bridging concurrent behavioral and neural data has been termed *model-based cognitive neuroscience* [99]. In model-based fMRI analyses, time-varying estimates of a cognitive process are convolved with a hemodynamic response function to derive the predicted BOLD activity for that process. One can then extend standard voxel-wise general linear model analyses to identify regions whose activity scales *parametrically* with the model-based regressors. Such regions are likely to be involved in the *representation* of a cognitive process instantiated in a theoretically motivated model. Importantly, a model may be falsified if it provides a poor account of behavior or demonstrates flaws in generative simulations. Moreover, a model may describe behavior well, but its components may not be associated with neural activity (i.e., a failure to bridge levels of analysis).

The structure of formal models depends on the scientific question — indeed, a model describes information processing in the context of a specific task. For example, many models of value-based decision-making specify how individuals update their estimates of subjective value based on surprising outcomes (i.e., prediction errors; see Box 3), whereas models of category learning may focus on differentiating prototypes from exemplars [100]. For a thorough exposition of model-based cognitive neuroscience we refer readers to [19].

The growing field of *computational psychiatry* uses formal models and model-based neuroimaging analyses to characterize psychopathology in terms of latent cognitive processes that underlie symptoms and maladaptive behaviors [101]. Computational psychiatry may also elucidate meaningful clinical phenotypes that are based purely on theoretically motivated formal models rather than expert consensus [102]. Altogether, the marriage of formal models of behavior with functional neuroimaging enables more

specific tests of *cognitive representation*, thereby constraining and clarifying the interpretation of neural activity.



Box 2 Figure.

In traditional and model-based fMRI analyses, participants complete an experimental task designed to elicit a cognitive process and/or behavior of interest. In this example, on each trial, a dot revolves around a circle over the course of four seconds [103]. Responding at any point during the revolution terminates the trial and leads to a unique probabilistic reward displayed during a feedback phase lasting 900ms. Thus, participants must learn *when* to respond in order to obtain larger rewards.

In general linear model (GLM) analyses of fMRI data, the design matrix encodes the onset, duration, and magnitude of stimuli presented during the experiment. To estimate the predicted blood-oxygen-level-dependent (BOLD) response (i.e., the conventional index of neural activity in fMRI), each regressor in the design matrix is convolved with a canonical hemodynamic response function (*convolution with HRF)*.
In traditional analyses (left column), feedback events are binary (*design matrix*) and reflect whether participants were viewing the feedback stimulus at that time. The major difference in model-based fMRI analyses is that the design matrix represents a cognitive process estimated according to a formal model of behavior (*cognitive representation*). Here, we illustrate the magnitude of prediction errors in six trials from a single subject, fit according to a delta rule learning model (see Box 3). Thus, the brain activity associated with the model-based signal describes the response of functional regions to the *magnitude* of prediction errors, whereas regions associated with the feedback signal reflect overall activity without a specific hypothesis of what cognitive operations are involved in processing feedback.

Box 3

Reinforcement learning: A primer

Reinforcement learning (RL) is an algorithmic account of how *agents* learn to make decisions on the basis of rewards and punishments (to be brief, we refer to rewards below). Each decision point is a *state* marked by *cues* (stimuli) that may be predictive of rewards, as well as available *actions*. In Pavlovian learning (stimulus-outcome), if a stimulus (e.g., light) is present in a given state and a reward is subsequently delivered, the agent learns the *expected reward value* of the stimulus. In instrumental learning (action-outcome), by selecting an *action*, the agent can observe its *outcome* and update the *expected value* of that action. In both cases, agents seek to minimize the discrepancy between expected and obtained outcomes. This discrepancy is called the *prediction error* (PE), a key learning signal. Prediction errors are positive if the outcome is greater than expected and vice versa. Most RL models update expected value estimates on the basis of PEs according to a form of *delta rule* learning [21,24]:

$$V_{t+1} = V_t + \alpha(R_t - V_t)$$

where V is the expected value of a stimulus or action, R is reward, $(R_t - V_t)$ is the PE in the current learning episode, and is time (across learning episodes). The impact of each PE on the current value estimate is determined by the learning rate parameter α, which conventionally varies between zero and one. The gradual evolution of expected value across multiple learning episodes gives rise to a learning curve in which an agent's estimate converges toward the underlying (unknown) reward contingency. Human imaging studies map expected value representations to the *ventromedial prefrontal cortex* (vmPFC), whereas PE signals are found in the ventral and dorsal striatum, which receive projections from the dopaminergic midbrain, as well as the thalamus [see 104 for more on the correlates of RL in the human brain].

Some important behaviors fall outside of this simple stimulus-outcome or action-outcome learning framework. For example, in Pavlovian learning, although no proper actions are involved, the organism emits appetitive or aversive conditioned responses (CRs) in response to predictive cues (i.e., conditioned stimuli). From an evolutionary standpoint, CRs are genetically determined adaptive behaviors, such as freezing in response to a threat or licking in response to food cues [105]. In the case of instrumental learning, when the value of an action in a given state is over-trained, habits emerge, where cues signifying a specific state elicit an automatic action (i.e., stimulus-response pairings). For example, one may fail to exit the highway to meet friends at a restaurant because the exit falls along the usual route from home to work. Finally, as detailed in the main text, important interactions between these learning modes have been highlighted by recent research on Pavlovian-to-instrumental transfer (PIT; for a review, see [66]).

We argue that these limitations can be overcome by developing formal models of decision-making [18] in BPD and testing them using the tools of modern cognitive neuroscience [19,20] (Box 2). We focus specifically on reinforcement learning (Box 3), which has a rich tradition of developing algorithms *in silico* [21] and validating them at the level of behavior and neural representation [22–24]. The past two decades of behavioral and computational neuroscience research have provided overwhelming evidence that adaptive and maladaptive decisions can be understood as reflecting the contribution of at least three computational systems: goal-directed, habitual, and Pavlovian (see Box 3) [18]. Furthermore, a coarse anatomical distinction between the primitive limbic system and sophisticated neocortical networks has given way to the perspective that cortical-striatal-limbic networks [25] carry decision-relevant signals from each of these systems [26]. We propose that the interpersonal hallmarks of BPD reflect: 1) an abnormally heightened motivation for interpersonal affiliation combined with a vulnerability to shifts in motivation; and 2) a predominance of Pavlovian (i.e., stimulus-outcome) influences on behavior that dominate goal-directed decision-making in stressful and emotionally arousing contexts.

Disadvantageous Decisions in Borderline Personality

Clinicians in training are often vexed by the extreme and seemingly irrational decisions that patients with BPD make. For example, why would a patient who talked at length about hating his girlfriend attempt suicide a week later when she broke up with him? Clinical theories of BPD have offered psychodynamic and cognitive-behavioral explanations. For example, object relations theory frames contradictory behaviors in BPD in terms of split object representations in which attractive and aversive aspects of oneself and others remain unintegrated, with strong affect promoting transitions between extremes [27]. Conversely, dialectical behavior therapy emphasizes the functional role of extreme behaviors (e.g., rage), which may meet immediate interpersonal desires (e.g., emotional intimacy) at the expense of long-term relationships [28]. Several therapeutic approaches to BPD [27–29] involve improving metacognition (especially representation of others' states of mind) and providing direct feedback to patients about maladaptive decisions [30].

Disadvantageous decisions in BPD often occur in the context of strong negative emotions, a facet of impulsivity called negative urgency [31]. Such decisions may reflect an emotion-dependent preference for immediate outcomes over delayed outcomes [32]. For example, an individual coping with marital conflict may assign greater subjective value to substance use than problem-solving to preserve the relationship. However, the notion that negative emotions promote short-sighted decisions was challenged by a recent meta-analysis of decision-making tasks [33]: individuals with BPD had an exaggerated preference for immediate rewards, but this was not enhanced by stress [34] or social exclusion [35]. Moreover, this preference was associated with a lack of premeditation, but not negative urgency. Individuals with BPD also tend to make poor decisions on the Iowa Gambling Task (IGT), a widely used paradigm that involves learning to choose safer card decks over high risk-high reward decks. This deficit may reflect insensitivity to losses in BPD [36], but this literature is inconsistent [37] and the IGT is likely a sensitive, not specific, measure of decision deficits in psychopathology [38].

What is a Formal Model of Decision-Making?

Although clinical theories provide a framework for understanding and remediating decision-making in BPD, they do not offer a mathematical account of how decision processes result in maladaptive choices. Likewise, empirical research to date has largely focused on summary statistics of decision quality (e.g., proportion of risky bets on a game) [33]. Over the past few decades, scientists have developed formal models of decision-making that integrate behavioral economics [39], psychology [40], and artificial intelligence [21]. The purpose of a formal model is to test falsifiable hypotheses about *how* decisions are made, typically by observing an evolving series of choices in experiments that approximate real-world challenges. Using a set of related equations, formal models specify what information is tracked by the individual and how it is integrated to choose an action in a specific situation.

Most formal models of decision-making can be understood within a Bayesian decision theory framework [41], which has five major components: 1) a specific state of the environment; 2) a set of actions that can be chosen in the current state; 3) subjective value estimates for each action quantifying its net costs and benefits given the current state; 4) a policy for selecting an optimal action given current value estimates; and 5) a procedure for updating value estimates in light of additional experience or shifts in the environment [18]. Crucially, subjective representations of the environment incorporate objective factors, internal factors, and prior beliefs and experience [42]. For example, imagine a man waiting at a restaurant for his partner to arrive (objective). Knowing that the partner often runs late (prior experience), he is prepared and starts reading a novel. After about 30 minutes, however, he begins to feel lonely (an internal state in which emotional intimacy has high value). When the partner arrives, he immediately begins to complain about feeling unloved (greater response vigor in a ‘deprived’ state) and does not notice the partner’s agitation. Only after calming down does he learn that the partner was in a minor car accident and had to exchange insurance information before driving to dinner.

Decision makers rarely, if ever, have perfect knowledge about the environment, instead deciding based on heuristic estimates subject to cognitive constraints [43]. Formal models of decision-making offer the potential to identify whether maladaptive decisions in BPD reflect a short-sighted policy (e.g., acting to preserve immediate proximity to a partner at the cost of losing the relationship), imprecise or biased subjective value estimates (e.g., believing that emotional intimacy is valuable regardless of relationship context), or overgeneralized representations that fail to incorporate new evidence (e.g., expecting others to be hostile based on prior conflicts).

How can formal models of decision-making help to dissect clinical features of BPD?

Our approach is motivated by the observation that individuals with BPD struggle to develop long-term relationships characterized by reciprocity, trust, and secure attachment [29,44]. Instead, relationships are often intense and dysfunctional, which may reflect maladaptive responses to others' mental states [45,46], a poor understanding of appropriate behaviors across social domains [47], and unstable representations of others [27,48]. Indeed, ecological momentary assessment (EMA) studies have found increased positive feedback among interpersonal perception, affect, and behavior in BPD. For example, perceiving others as cold or quarrelsome triggers elevated negative affect and quarrelsome behavior in individuals with BPD, which elicits quarrelsome behavior from others [49]. Using more nuanced time-lagged analyses, Scott and colleagues [50] recently found that the tendency for perceived rejection to evoke aggression was exacerbated by negative emotions (especially anger) in young women with borderline symptoms.

Altogether, clinical theory and both cross-sectional and EMA studies provide three important leads that inform a decision neuroscience approach to BPD. First, individuals with BPD are often insensitive to adverse consequences of their own behaviors. For example, during social economic exchanges, individuals with BPD fail to cooperate to maximize their gains [51] and make inconsistent choices [52]. Second, coping strategies (e.g., self-injury) and responses to interpersonal discord (e.g., overt hostility) are often incongruent with individuals' longer-term goals [53]. Third, interpersonal perceptions and ensuing responses (e.g., perceived rejection eliciting aggression) are often rigidly coupled, especially during times of distress [50].

We propose that maladaptive, inflexible interpersonal behaviors reflect excessive influence of the Pavlovian system (stimulus-outcome) on social decisions. More specifically, conditioned responses (CRs) to social stimuli are likely a key component of disruptive interpersonal behaviors in BPD, from quarrelsomeness to self-destructive acts. The inflexibility of CRs and their sensitivity to internal state rather than consequences enables feedback loops, leading to escalating conflicts. Interpersonal problems are likely to emerge in situations where the motivational state (including emotional context) promotes actions with potentially aversive relationship consequences. For example, if one is angry with a seemingly distant partner, the congruent Pavlovian response is to start an argument (i.e., approach). This will often lead the partner to withdraw. Individuals with BPD may learn to

associate the partner (the conditioned stimulus) with the experience of abandonment (outcome) such that fights paradoxically escalate as the partner becomes more aloof in response (cf. [49,50]). Such dynamics are common early in therapy, when patients initially fail to understand how their actions affect others, focusing instead on how others affect them [29].

By contrast, adaptive social learning requires one to track others' internal states [46,54] and the results of one's actions in order to learn from interpersonal successes and mistakes. In the BPD literature, this has been called 'mentalizing' [29]; the learning theory homologue is goal-directed learning, which is defined as choosing an action by mentally traversing a transition structure to select a currently desired outcome (i.e., action-outcome) [55]. Returning to the example above, if one's goal is to maintain a mutually rewarding relationship, one could learn from repeated experience that fighting (goal-incongruent action) pushes the partner away (outcome) and restrain oneself (goal-congruent action). Such learning could involve emotion reappraisal (e.g., viewing the partner as preoccupied, but not emotionally distant) in order to promote goal-congruent decisions [56]. Goal-directed learners are sensitive to action-outcome contingencies [18,42] and update their preferences in light of specific feedback.

The role of Pavlovian-to-instrumental transfer (PIT) in BPD

Integrating clinical, experimental, and neuroscience evidence, we propose that PIT is a key mechanism through which the Pavlovian system dominates decisions in BPD, leading to maladaptive choices that are sensitive to one's motivational state and relatively insensitive to context and action-outcome contingencies. A comprehensive account of learning system alterations in BPD is beyond the scope of this brief review. We note, however, that stress disrupts goal-directed learning and enhances Pavlovian and habitual tendencies [57–61]. Furthermore, uncontrollable stress during development has disruptive long-term effects on stress reactivity and learning from new experience [62–64]. This accords with evidence that individuals with BPD experience relationships as stressful and unpredictable, which may relate to early experiences with caregivers who are erratic to the point of abusiveness [65].

Research over the past decade has revealed behavioral and biological distinctions between two forms of PIT, outcome-specific and general [66]. Outcome-specific PIT describes the tendency of a conditioned stimulus (S_1) associated with an appetitive outcome (O_1) to invigorate actions (A_1) that lead to that outcome. For example, the smell of freshly baked cookies may lead one to seek them out. Likewise, aversive cues can promote behaviors to avoid undesirable outcomes [67]. One account of specific PIT is that a Pavlovian cue activates the representation of its associated outcome (i.e., S_1-O_1), which in turn promotes outcome-congruent actions (A_1-O_1) [68]. In contrast, general PIT describes how a cue (S_1) can promote actions ($A_{2\ldots n}$) that are not specifically related to the cue and lead to a number of different outcomes ($O_{2\ldots n}$) [69]. For example, the presence of appealing foods at a party may elicit other reward-seeking behaviors such as drinking wine or socializing. Thus, outcome-specific PIT reflects the role of Pavlovian cues in promoting actions that are likely to achieve a particular goal. Conversely, in general PIT, cues signal the availability of

rewards and punishments in the broader environment, thereby altering the value of instrumental actions [66].

Crucially, general PIT depends on motivational state, whereas outcome-specific PIT does not [70]. Extending the concept of polarized affect states and representations of self and other from object relations theory [27], we view intense, changing emotions in BPD (i.e., affective instability [71]) as motivational shifts that amplify the perceived value of emotion-congruent outcomes [32,56] and enhance Pavlovian influences via general PIT. Accordingly, an individual who experiences a strong 'hunger' to affiliate with others may assign high value to various social outcomes. In BPD, affiliative motivation often results in decisions that may have aversive consequences (e.g., pursuing a sexual relationship with a therapist) or that are inappropriate to the context (e.g., expressing intense jealousy about a friend spending time with others) [47]. Such behaviors fail to satisfy the motivational demand and often lead to aversive outcomes (e.g., a therapist rejecting the patient's sexual advances).

Furthermore, the intensity of motivational states in BPD (i.e., states in which some outcomes have very high perceived value) likely has a general invigorating effect on behavior that may interfere with long-term goals [72]. Viewed through the lens of learning theory, this may reflect three related effects: 1) the enhancement of Pavlovian-congruent actions via general PIT, irrespective of goal; 2) a reduction in the overall threshold for executing actions [73]; and 3) the tendency to execute habitual actions more vigorously regardless of the value of the outcome [42]. For example, intense emotional distress following an argument may lead an individual with BPD to pursue casual sex or to have a drug relapse even if these are incongruent with a goal of preserving a long-term relationship. Altogether, the persistence of extreme motivational states in BPD suggests a difficulty mapping interpersonal contingencies in a goal-directed fashion or an overreliance Pavlovian influences that may not align with situational demands.

Neural substrates of Pavlovian-instrumental interactions

We predict that maladaptive behaviors in BPD involve specific alterations in the cortico-limbic circuitry underlying Pavlovian-instrumental interactions (detailed in Figure 1). We propose that the phylogenetically old pathway from the central nucleus of the amygdala (CNA) and ventral tegmental area (VTA) to the core of the nucleus accumbens (NAcc-C) underpins emotion-driven Pavlovian responses, as in general PIT [74]. Furthermore, general PIT tends to invigorate habitual responses [66], which are not sensitive to the value of the resulting outcome and thus may contribute to problems such as reactive aggression or substance abuse. Invigoration of goal-congruent behaviors by environmental stimuli, on the other hand, is likely mediated by the phylogenetically newer pathway from the basolateral amygdala (BLA) to the shell of the nucleus accumbens (NAcc-S) and by the recurrent projections of the ventromedial prefrontal (vmPFC) and medial orbitofrontal cortex (mOFC) into the NAcc-S, as in outcome-specific PIT [69,75]. Finally, we believe that projections from the medial prefrontal cortex (mPFC) to the striatum may facilitate the suppression of Pavlovian influences on instrumental behavior [76] and play a key role in the shift from reactive to deliberative behaviors in treatments for BPD.

Toward a decision neuroscience account of BPD

This new approach recasts a frontolimbic account of BPD in terms of the interplay among computational modes, and it is solidly grounded in experimental psychology, neuroscience, and computer science [21,26]. More than simply describing behavior, formal models of Pavlovian-instrumental interactions [77,78] enable scientists to understand individual differences in latent decision processes and their neural signatures. These models have yielded novel behavioral and neurobiological insights into alcohol dependence [79], schizophrenia [80], and depression [81]. Furthermore, recent advances in high-resolution functional neuroimaging [82] offer the potential to study subregions of the amygdala and striatum that are differentially implicated in general and outcome-specific PIT [83]. Thus, formal models of decision-making support theory-guided accounts of behavior during fMRI experiments, providing a crucial bridge between the complex clinical features of BPD and the latent neurobehavioral systems that underpin them [20,84]. We believe that a decision neuroscience approach that experimentally dissects Pavlovian-instrumental interactions in the context of social decisions [85] will advance our understanding of BPD.

Acknowledgments

This work was supported by the National Institutes of Health (R01 MH100095, R01 MH085651 to AYD; K01 MH090791 to MNH; R01 MH048463 to AYD and MNH).

References

* Of special interest

** Of outstanding interest

1. Skodol AE, Gunderson JG, Pfohl B, Widiger TA, Livesley WJ, Siever LJ. The borderline diagnosis I: Psychopathology, comorbidity, and personality structure. Biol Psychiatry. 2002; 51:936–950. [PubMed: 12062877]
2. Berenson KR, Gregory WE, Glaser E, Romirowsky A, Rafaeli E, Yang X, Downey G. Impulsivity, Rejection Sensitivity, and Reactions to Stressors in Borderline Personality Disorder. Cogn Ther Res. 2016; :1–12. DOI: 10.1007/s10608-015-9752-y
3. Brodsky BS, Groves SA, Oquendo MA, Mann JJ, Stanley B. Interpersonal precipitants and suicide attempts in borderline personality disorder. Suicide Life Threat Behav. 2006; 36:313–322. DOI: 10.1521/suli.2006.36.3.313 [PubMed: 16805659]
4. Gunderson JG, Lyons-Ruth K. BPD's Interpersonal Hypersensitivity Phenotype: A Gene-Environment-Developmental Model. J Personal Disord. 2008; 22:22–41. DOI: 10.1521/pedi.2008.22.1.22
5. Daros AR, Zakzanis KK, Ruocco AC. Facial emotion recognition in borderline personality disorder. Psychol Med. 2013; 43:1953–1963. DOI: 10.1017/S0033291712002607 [PubMed: 23149223]
6. Roepke S, Vater A, Preißler S, Heekeren HR, Dziobek I. Social cognition in borderline personality disorder. Front Neurosci. 2013; 6:195.doi: 10.3389/fnins.2012.00195 [PubMed: 23335877]
7. Minzenberg M, Fan J, New A, Tang C, Siever L. Fronto-limbic dysfunction in response to facial emotion in borderline personality disorder: An event-related fMRI study. Psychiatry Res Neuroimaging. 2007; 155:231–243. doi:http://dx.doi.org/10.1016/j.pscychresns.2007.03.006.
8. Koenigsberg HW, Denny BT, Fan J, Liu X, Guerreri S, Mayson SJ, Rimsky L, New AS, Goodman M, Siever LJ. The Neural Correlates of Anomalous Habituation to Negative Emotional Pictures in Borderline and Avoidant Personality Disorder Patients. Am J Psychiatry. 2014; 171:82–90. DOI: 10.1176/appi.ajp.2013.13070852 [PubMed: 24275960]

9. Domsalla M, Koppe G, Niedtfeld I, Vollstädt-Klein S, Schmahl C, Bohus M, Lis S. Cerebral processing of social rejection in patients with borderline personality disorder. Soc Cogn Affect Neurosci. 2014; 9:1789–1797. DOI: 10.1093/scan/nst176 [PubMed: 24273076]

10. Ongür D, Ferry AT, Price JL. Architectonic subdivision of the human orbital and medial prefrontal cortex. J Comp Neurol. 2003; 460:425–449. DOI: 10.1002/cne.10609 [PubMed: 12692859]

*11. Schulze L, Schmahl C, Niedtfeld I. Neural Correlates of Disturbed Emotion Processing in Borderline Personality Disorder: A Multimodal Meta-Analysis. Biol Psychiatry. 2016; 79:97–106. Meta-analysis of several functional imaging studies of emotion processing in BPD that finds hyperactivity of the amygdala and posterior cingulate cortex to negative emotional stimuli, but only in unmedicated samples. DOI: 10.1016/j.biopsych.2015.03.027 [PubMed: 25935068]

12. Silbersweig D, Clarkin JF, Goldstein M, Kernberg OF, Tuescher O, Levy KN, Brendel G, Pan H, Beutel M, Pavony MT, Epstein J, Lenzenweger MF, Thomas KM, Posner MI, Stern E. Failure of frontolimbic inhibitory function in the context of negative emotion in borderline personality disorder. Am J Psychiatry. 2007; 164:1832–41. DOI: 10.1176/appi.ajp.2007.06010126 [PubMed: 18056238]

13. Soloff PH, Abraham K, Burgess A, Ramaseshan K, Chowdury A, Diwadkar VA. Impulsivity and aggression mediate regional brain responses in Borderline Personality Disorder: An fMRI study. Psychiatry Res Neuroimaging. 2017; 260:76–85. DOI: 10.1016/j.pscychresns.2016.12.009 [PubMed: 28039797]

14. Soloff PH, Abraham K, Ramaseshan K, Burgess A, Diwadkar VA. Hyper-modulation of brain networks by the amygdala among women with Borderline Personality Disorder: Network signatures of affective interference during cognitive processing. J Psychiatr Res. 2017; 88:56–63. DOI: 10.1016/j.jpsychires.2016.12.016 [PubMed: 28086129]

15. Mancke F, Herpertz SC, Bertsch K. Aggression in borderline personality disorder: A multidimensional model. Personal Disord Theory Res Treat. 2015; 6:278–291. doi:http://dx.doi.org.ezaccess.libraries.psu.edu/10.1037/per0000098.

16. Berenson KR, Downey G, Rafaeli E, Coifman KG, Paquin NL. The rejection–rage contingency in borderline personality disorder. J Abnorm Psychol. 2011; 120:681–690. doi:http://dx.doi.org/10.1037/a0023335. [PubMed: 21500875]

17. Price JL, Drevets WC. Neural circuits underlying the pathophysiology of mood disorders. Trends Cogn Sci. 2012; 16:61–71. DOI: 10.1016/j.tics.2011.12.011 [PubMed: 22197477]

**18. Huys QJM, Guitart-Masip M, Dolan RJ, Dayan P. Decision-Theoretic Psychiatry. Clin Psychol Sci. 2015; 2167702614562040 Provides an outstanding introduction to the formalism of Bayesian decision theory and how it can help to dissect maladaptive decisions in psychopathology. doi: 10.1177/2167702614562040

*19. Turner BM, Forstmann BU, Love BC, Palmeri TJ, Van Maanen L. Approaches to analysis in model-based cognitive neuroscience. J Math Psychol. 2017; 76:65–79. An outstanding review that explores different approaches for synthesizing neural data with formal models of behavior. Covers both the theoretical differences among approaches and the empirical findings that each has yielded. DOI: 10.1016/j.jmp.2016.01.001

**20. Love BC. The Algorithmic Level Is the Bridge Between Computation and Brain. Top Cogn Sci. 2015; 7:230–242. A conceptual paper that argues for the importance of formal, algorithmic models in understanding how the brain implements specific cognitive processes that are expressed in behavior (i.e., bridging these levels of analysis). DOI: 10.1111/tops.12131 [PubMed: 25823496]

21. Sutton, RS., Barto, AG. Reinforcement learning: An introduction. MIT Press; Cambridge, MA: 1998.

22. Dayan P, Daw ND. Decision theory, reinforcement learning, and the brain. Cogn Affect Behav Neurosci. 2008; 8:429–453. DOI: 10.3758/CABN.8.4.429 [PubMed: 19033240]

23. Schultz W, Dayan P, Montague PR. A Neural Substrate of Prediction and Reward. Science. 1997; 275:1593–1599. DOI: 10.1126/science.275.5306.1593 [PubMed: 9054347]

24. Glimcher PW. Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. Proc Natl Acad Sci. 2011; 108:15647–15654. DOI: 10.1073/pnas.1014269108 [PubMed: 21389268]

25. Laird AR, Fox PM, Eickhoff SB, Turner JA, Ray KL, McKay DR, Glahn DC, Beckmann CF, Smith SM, Fox PT. Behavioral interpretations of intrinsic connectivity networks. J Cogn Neurosci. 2011; 23:4022–37. DOI: 10.1162/jocn_a_00077 [PubMed: 21671731]

26. Balleine BW, Delgado MR, Hikosaka O. The role of the dorsal striatum in reward and decision-making. J Neurosci. 2007; 27:8161–8165. DOI: 10.1523/JNEUROSCI.1554-07.2007 [PubMed: 17670959]

27. Kernberg OF. Borderline personality organization. J Am Psychoanal Assoc. 1967; 15:641–685. [PubMed: 4861171]

28. Linehan, MM. Cognitive-Behavioral Treatment of Borderline Personality Disorder. Guilford Press; 1993.

29. Bateman, A., Fonagy, P. Psychotherapy for Borderline Personality Disorder: Mentalization Based Treatment. Oxford University Press; New York: 2004.

30. Stone MH. Management of borderline personality disorder: a review of psychotherapeutic approaches. World Psychiatry Off J World Psychiatr Assoc WPA. 2006; 5:15–20.

31. Peters JR, Upton BT, Baer RA. Brief report: relationships between facets of impulsivity and borderline personality features. J Personal Disord. 2013; 27:547–552. DOI: 10.1521/pedi_2012_26_044

32. Phelps EA, Lempert KM, Sokol-Hessner P. Emotion and Decision Making: Multiple Modulatory Neural Circuits. Annu Rev Neurosci. 2014; 37:263–287. DOI: 10.1146/annurev-neuro-071013-014119 [PubMed: 24905597]

**33. Paret C, Jennen-Steinmetz C, Schmahl C. Disadvantageous decision-making in borderline personality disorder: Partial support from a meta-analytic review. Neurosci Biobehav Rev. 2017; 72:301–309. Meta-analysis of behavioral decision studies of BPD that describes key effects for delay discounting, reversal learning, and Iowa Gambling Task (IGT). Across studies, individuals with BPD displayed a preference for immediate rewards independent of acute stress, and they performed more poorly on the IGT. DOI: 10.1016/j.neubiorev.2016.11.019 [PubMed: 27914943]

*34. Krause-Utz A, Cackowski S, Daffner S, Sobanski E, Plichta MM, Bohus M, Ende G, Schmahl C. Delay discounting and response disinhibition under acute experimental stress in women with borderline personality disorder and adult attention deficit hyperactivity disorder. Psychol Med. 2016; 46:3137–3149. doi:http://dx.doi.org.ezaccess.libraries.psu.edu/10.1017/S0033291716001677 Describes the effect of acute stress on several behavioral measures of impulsivity in BPD. The BPD group exhibited difficulty withholding in appropriate responses on a speeded choice task under stress. [PubMed: 27572473]

35. Lawrence KA, Allen JS, Chanen AM. Impulsivity in Borderline Personality Disorder: Reward-Based Decision-Making and its Relationship to Emotional Distress. J Personal Disord. 2010; 24:785–799. DOI: 10.1521/pedi.2010.24.6.785

36. Saunders KEA, Goodwin GM, Rogers RD. Insensitivity to the Magnitude of Potential Gains or Losses When Making Risky Choices: Women With Borderline Personality Disorder Compared With Bipolar Disorder and Controls. J Personal Disord. 2015; 30:530–544. DOI: 10.1521/pedi_2015_29_216

37. LeGris J, Toplak M, Links PS. Affective Decision Making in Women With Borderline Personality Disorder. J Personal Disord. 2014; 28:698–719. DOI: 10.1521/pedi_2014_28_140

*38. Mukherjee D, Kable JW. Value-Based Decision Making in Mental Illness A Meta-Analysis. Clin Psychol Sci. 2014; 2167702614531580 A meta-analysis of value-based decision tasks in psychopathology. Focuses especially on the Iowa Gambling Task, finding that decision deficits are common across mental disorders. Underscores the importance of employing decision battery assessments that account for specific cognitive processes in order to differentiate among disorders. doi: 10.1177/2167702614531580

39. Loewenstein, G. Exotic Preferences: Behavioral Economics and Human Motivation. Oxford University Press; New York: 2008.

40. Hastie, R., Dawes, RM. Rational Choice in an Uncertain World: The Psychology of Judgment and Decision Making. Sage Publications; Los Angeles: 2009.

41. Berger, JO. Statistical Decision Theory and Bayesian Analysis. 2. Springer; New York: 1993.

42. Niv Y, Joel D, Dayan P. A normative perspective on motivation. Trends Cogn Sci. 2006; 10:375–381. DOI: 10.1016/j.tics.2006.06.010 [PubMed: 16843041]

43. Simon HA. Rational choice and the structure of the environment. Psychol Rev. 1956; 63:129–38. [PubMed: 13310708]

44. Levy KN. The Implications of attachment theory and research for understanding Borderline Personality Disorder. Dev Psychopathol. 2005; 17:959–986. DOI: 10.1017/S0954579405050455 [PubMed: 16613426]

45. Beeney JE, Stepp SD, Hallquist MN, Scott LN, Wright AGC, Ellison WD, Nolf KA, Pilkonis PA. Attachment and social cognition in borderline personality disorder: Specificity in relation to antisocial and avoidant personality disorders. Personal Disord Theory Res Treat. 2015; 6:207–215. doi:http://dx.doi.org.ezaccess.libraries.psu.edu/10.1037/per0000110.

*46. Miano A, Dziobek I, Roepke S. Understanding Interpersonal Dysfunction in Borderline Personality Disorder: A Naturalistic Dyadic Study Reveals Absence of Relationship-Protective Empathic Inaccuracy. Clin Psychol Sci. 2017; 5:355–366. An observational study of conversation dynamics in romantic couples that found greater empathic accuracy to relationship-threatening content in BPD, suggesting a promising latent process underlying disturbed interpersonal functioning. DOI: 10.1177/2167702616683505

47. Hill J, Pilkonis P, Morse J, Feske U, Reynolds S, Hope H, Charest C, Broyden N. Social domain dysfunction and disorganization in borderline personality disorder. Psychol Med. 2008; 38:135–146. DOI: 10.1017/S0033291707001626 [PubMed: 17892627]

*48. Beeney JE, Hallquist MN, Ellison WD, Levy KN. Self–other disturbance in borderline personality disorder: Neural, self-report, and performance-based evidence. Personal Disord Theory Res Treat. 2016; 7:28–39. doi:http://dx.doi.org.ezaccess.libraries.psu.edu/10.1037/per0000127synthesizes functional MRI and behavioral data to demonstrate that inconsistency in representations of self and others mediates hyperactivity of social cognitive regions such as mPFC in BPD.

49. Sadikaj G, DS, Russell JJ, Zuroff DC, Paris J. Quarrelsome behavior in borderline personality disorder: Influence of behavioral and affective reactivity to perceptions of others. J Abnorm Psychol. 2013; 122:195–207. DOI: 10.1037/a0030871 [PubMed: 23231460]

50. Scott LN, GC, Beeney JE, Lazarus SA, Pilkonis PA, Stepp SD. Borderline personality disorder symptoms and aggression: A within-person process model. J Abnorm Psychol. 2017; 126:429–440. DOI: 10.1037/abn0000272 [PubMed: 28383936]

51. King-Casas B, Sharp C, Lomax-Bream L, Lohrenz T, Fonagy P, Montague P. The rupture and repair of cooperation in borderline personality disorder. Science. 2008; 2008:806–810.

52. Preuss N, Brändle LS, Hager OM, Haynes M, Fischbacher U, Hasler G. Inconsistency and social decision making in patients with Borderline Personality Disorder. Psychiatry Res. 2016; 243:115–122. DOI: 10.1016/j.psychres.2016.06.017 [PubMed: 27380424]

53. Carpenter RW, Trull TJ. Components of Emotion Dysregulation in Borderline Personality Disorder: A Review. Curr Psychiatry Rep. 2013; 15:335.doi: 10.1007/s11920-012-0335-2 [PubMed: 23250816]

54. Apps MAJ, Rushworth MFS, Chang SWC. The Anterior Cingulate Gyrus and Social Cognition: Tracking the Motivation of Others. Neuron. 2016; 90:692–707. DOI: 10.1016/j.neuron.2016.04.018 [PubMed: 27196973]

55. Dickinson A, Balleine B. Motivational control of goal-directed action. Anim Learn Behav. 1994; 22:1–18. DOI: 10.3758/BF03199951

56. Huys QJM, Renz D. A formal valuation framework for emotions and their control. Biol Psychiatry. 2017; doi: 10.1016/j.biopsych.2017.07.003

57. Schwabe L, Tegenthoff M, Höffken O, Wolf OT. Simultaneous Glucocorticoid and Noradrenergic Activity Disrupts the Neural Basis of Goal-Directed Action in the Human Brain. J Neurosci. 2012; 32:10146–10155. DOI: 10.1523/JNEUROSCI.1304-12.2012 [PubMed: 22836250]

58. Schwabe L, Wolf OT. Stress Prompts Habit Behavior in Humans. J Neurosci. 2009; 29:7191–7198. DOI: 10.1523/JNEUROSCI.0979-09.2009 [PubMed: 19494141]

59. Minor TR, Jackson RL, Maier SF. Effects of task-irrelevant cues and reinforcement delay on choice-escape learning following inescapable shock: Evidence for a deficit in selective attention. J Exp Psychol Anim Behav Process. 1984; 10:543. [PubMed: 6491612]

*60. de Berker AO, Tirole M, Rutledge RB, Cross GF, Dolan RJ, Bestmann S. Acute stress selectively impairs learning to act. Sci Rep. 2016; 6:srep29816. The authors demonstrate that an acute stressor (socially evaluated cold pressor task) impairs the ability to learn an instrumental action, regardless of the valence of the outcome. This is consistent with the possibility that aversive Pavlovian influences impair instrumental approach, likely a form of Pavlovian-to-instrumental transfer. doi: 10.1038/srep29816

61. Otto AR, Raio CM, Chiang A, Phelps EA, Daw ND. Working-memory capacity protects model-based learning from stress. Proc Natl Acad Sci. 2013; 110:20941–20946. DOI: 10.1073/pnas.1312011110 [PubMed: 24324166]

62. Maier SF, Watkins LR. Role of the medial prefrontal cortex in coping and resilience. Brain Res. 2010; 1355:52–60. DOI: 10.1016/j.brainres.2010.08.039 [PubMed: 20727864]

63. Arnsten AFT. Stress signalling pathways that impair prefrontal cortex structure and function. Nat Rev Neurosci. 2009; 10:410–422. DOI: 10.1038/nrn2648 [PubMed: 19455173]

64. Huys QJM, Daw ND, Dayan P. Depression: A Decision-Theoretic Analysis. Annu Rev Neurosci. 2015; 38:1–23. DOI: 10.1146/annurev-neuro-071714-033928 [PubMed: 25705929]

65. Crowell SE, Beauchaine TP, Linehan MM. A biosocial developmental model of borderline personality: Elaborating and extending Linehan's theory. Psychol Bull. 2009; 135:495–510. DOI: 10.1037/a0015616 [PubMed: 19379027]

**66. Cartoni E, Balleine B, Baldassarre G. Appetitive Pavlovian-instrumental Transfer: A review. Neurosci Biobehav Rev. 2016; 71:829–848. Provides an outstanding review of theoretical and empirical research on Pavlovian-to-instrumental transfer. Summarizes both computational accounts and neural substrates in animal and human studies. DOI: 10.1016/j.neubiorev.2016.09.020 [PubMed: 27693227]

67. Alarcón DE, Bonardi C, Delamater AR. Associative Mechanisms Involved in Specific Pavlovian-to-instrumental transfer (PIT) in Human Learning Tasks. Q J Exp Psychol. 2017; :1–55. DOI: 10.1080/17470218.2017.1342671

68. Alarcón D, Bonardi C. The effect of conditioned inhibition on the specific Pavlovian-instrumental transfer effect. J Exp Psychol Anim Learn Cogn. 2016; 42:82–94. DOI: 10.1037/xan0000087 [PubMed: 26569018]

69. Corbit LH, Fischbach SC, Janak PH. Nucleus accumbens core and shell are differentially involved in general and outcome-specific forms of Pavlovian-instrumental transfer with alcohol and sucrose rewards. Eur J Neurosci. 2016; 43:1229–1236. DOI: 10.1111/ejn.13235 [PubMed: 26970240]

70. Corbit LH, Janak PH, Balleine BW. General and outcome-specific forms of Pavlovian-instrumental transfer: the effect of shifts in motivational state and inactivation of the ventral tegmental area. Eur J Neurosci. 2007; 26:3141–3149. DOI: 10.1111/j.1460-9568.2007.05934.x [PubMed: 18005062]

71. Ebner-Priemer UW, Houben M, Santangelo P, Kleindienst N, Tuerlinckx F, Oravecz Z, Verleysen G, Van Deun K, Bohus M, Kuppens P. Unraveling affective dysregulation in borderline personality disorder: A theoretical model and empirical evidence. J Abnorm Psychol. 2015; 124:186–198. DOI: 10.1037/abn0000021 [PubMed: 25603359]

72. Berridge KC. Motivation concepts in behavioral neuroscience. Physiol Behav. 2004; 81:179–209. DOI: 10.1016/j.physbeh.2004.02.004 [PubMed: 15159167]

73. Kim S, Lee D. Prefrontal Cortex and Impulsive Decision Making. Biol Psychiatry. 2011; 69:1140–1146. DOI: 10.1016/j.biopsych.2010.07.005 [PubMed: 20728878]

74. Corbit LH, Balleine BW. Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer. J Neurosci. 2005; 25:962–70. DOI: 10.1523/JNEUROSCI.4507-04.2005 [PubMed: 15673677]

75. Balleine BW, Killcross S. Parallel incentive processing: an integrated view of amygdala function. Trends Neurosci. 2006; 29:272–279. DOI: 10.1016/j.tins.2006.03.002 [PubMed: 16545468]

76. Cavanagh JF, Eisenberg I, Guitart-Masip M, Huys Q, Frank MJ. Frontal Theta Overrides Pavlovian Learning Biases. J Neurosci. 2013; 33:8541–8548. DOI: 10.1523/JNEUROSCI.5754-12.2013 [PubMed: 23658191]

77. Guitart-Masip M, Huys QJM, Fuentemilla L, Dayan P, Duzel E, Dolan RJ. Go and no-go learning in reward and punishment: Interactions between affect and effect. NeuroImage. 2012; 62:154–166. DOI: 10.1016/j.neuroimage.2012.04.024 [PubMed: 22548809]

78. Huys QJM, Cools R, Gölzer M, Friedel E, Heinz A, Dolan RJ, Dayan P. Disentangling the Roles of Approach, Activation and Valence in Instrumental and Pavlovian Responding. PLoS Comput Biol. 2011; 7:e1002028.doi: 10.1371/journal.pcbi.1002028 [PubMed: 21556131]

79. Garbusow M, Schad DJ, Sebold M, Friedel E, Bernhardt N, Koch SP, Steinacher B, Kathmann N, Geurts DEM, Sommer C, Müller DK, Nebe S, Paul S, Wittchen H-U, Zimmermann US, Walter H, Smolka MN, Sterzer P, Rapp MA, Huys QJM, Schlagenhauf F, Heinz A. Pavlovian-to-instrumental transfer effects in the nucleus accumbens relate to relapse in alcohol dependence. Addict Biol. 2016; 21:719–731. DOI: 10.1111/adb.12243 [PubMed: 25828702]

**80. Morris RW, Quail S, Griffiths KR, Green MJ, Balleine BW. Corticostriatal Control of Goal-Directed Action Is Impaired in Schizophrenia. Biol Psychiatry. 2015; 77:187–195. An empirical study of Pavlovian-to-instrumental transfer and goal-directed learning in schizophrenia. Patients persisted in choosing a devalued outcome (food) despite rating it more negatively, suggesting problems updating action-outcome representations. This deficit was associated with reduce caudate activity during decisions. DOI: 10.1016/j.biopsych.2014.06.005 [PubMed: 25062683]

**81. Huys QJM, Gölzer M, Friedel E, Heinz A, Cools R, Dayan P, Dolan RJ. The specificity of Pavlovian regulation is associated with recovery from depression. Psychol Med. 2016; 46:1027–1035. An empirical study of Pavlovian-to-instrumental transfer that identified reduced ability of Pavlovian cues to invigorate congruent actions (e.g., positive cue–approach behavior) in currently depressed individuals. Greater congruence of PIT effects predicted the degree of symptom recovery at 4–6-month follow-up. DOI: 10.1017/S0033291715002597 [PubMed: 26841896]

82. Xu J, Moeller S, Auerbach EJ, Strupp J, Smith SM, Feinberg DA, Yacoub E, Uğurbil K. Evaluation of slice accelerations using multiband echo planar imaging at 3T. NeuroImage. 2013; 83:991–1001. DOI: 10.1016/j.neuroimage.2013.07.055 [PubMed: 23899722]

83. Prévost C, Liljeholm M, Tyszka JM, O'Doherty JP. Neural Correlates of Specific and General Pavlovian-to-Instrumental Transfer within Human Amygdalar Subregions: A High-Resolution fMRI Study. J Neurosci. 2012; 32:8383–8390. DOI: 10.1523/JNEUROSCI.6237-11.2012 [PubMed: 22699918]

84. Depue, RA., Lenzenweger, MF. Toward a Developmental Psychopathology of Personality Disturbance: A Neurobehavioral Dimensional Model. In: Cicchetti, D., Cohen, DJ., editors. Dev Psychopathol. 2. John Wiley & Sons, Inc; 2015. p. 762-796.http://onlinelibrary.wiley.com/doi/10.1002/9780470939390.ch20/summary [accessed August 23, 2016]

85. Ruff CC, Fehr E. The neurobiology of rewards and values in social decision making. Nat Rev Neurosci. 2014; 15:549–562. DOI: 10.1038/nrn3776 [PubMed: 24986556]

86. Feighner JP, Robins E, Guze SB, Woodruff RA, Winokur G, Munoz R. Diagnostic criteria for use in psychiatric research. Arch Gen Psychiatry. 1972; 26:57–63. [PubMed: 5009428]

87. Gunderson JG, Singer MT. Defining borderline patients: An overview. Am J Psychiatry. 1975; 132:1–10. [PubMed: 802958]

88. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders: DSM-5. 5. American Psychiatric Publishing; Washington, D.C: 2013.

89. Widiger TA, Trull TJ. Plate tectonics in the classification of personality disorder: Shifting to a dimensional model. Am Psychol. 2007; 62:71–83. [PubMed: 17324033]

90. Krueger RF, Hopwood CJ, Wright AG, Markon KE. DSM-5 and the path toward empirically based and clinically useful conceptualization of personality and psychopathology. Clin Psychol Sci Pract. 2014; 21:245–261.

91. Clark LA. Assessment and Diagnosis of Personality Disorder: Perennial Issues and an Emerging Reconceptualization. Annu Rev Psychol. 2007; 58:227–257. DOI: 10.1146/annurev.psych.57.102904.190200 [PubMed: 16903806]

92. Eaton NR, Krueger RF, Keyes KM, Skodol AE, Markon KE, Grant BF, Hasin DS. Borderline personality disorder co-morbidity: relationship to the internalizing–externalizing structure of common mental disorders. Psychol Med. 2011; 41:1041–1050. DOI: 10.1017/S0033291710001662 [PubMed: 20836905]

93. Krueger RF. The Structure of Common Mental Disorders. Arch Gen Psychiatry. 1999; 56:921–926. DOI: 10.1001/archpsyc.56.10.921 [PubMed: 10530634]

94. Caspi A, Houts RM, Belsky DW, Goldman-Mellor SJ, Harrington H, Israel S, Meier MH, Ramrakha S, Shalev I, Poulton R, Moffitt TE. The p Factor: One General Psychopathology Factor in the Structure of Psychiatric Disorders? Clin Psychol Sci J Assoc Psychol Sci. 2014; 2:119–137. DOI: 10.1177/2167702613497473

95. Sharp C, Wright AGC, Fowler JC, Frueh BC, Allen JG, Oldham J, Clark LA. The structure of personality pathology: Both general (‘g’) and specific (‘s’) factors? J Abnorm Psychol. 2015; 124:387–398. doi:http://dx.doi.org.ezaccess.libraries.psu.edu/10.1037/abn0000033. [PubMed: 25730515]

96. Abram SV, DeYoung CG. Using personality neuroscience to study personality disorder. Personal Disord Theory Res Treat. 2017; 8:2–13.

97. Yarkoni, T. Neurobiological substrates of personality: A critical overview. In: Mikulincer, M., Shaver, PR., editors. APA Handb Personal Soc Psychol. American Psychological Association; Washington, D.C: 2014.

98. Krakauer JW, Ghazanfar AA, Gomez-Marin A, MacIver MA, Poeppel D. Neuroscience Needs Behavior: Correcting a Reductionist Bias. Neuron. 2017; 93:480–490. DOI: 10.1016/j.neuron.2016.12.041 [PubMed: 28182904]

99. Palmeri TJ, Love BC, Turner BM. Model-based cognitive neuroscience. J Math Psychol. 2017; 76:59–64. DOI: 10.1016/j.jmp.2016.10.010

100. Mack ML, Preston AR, Love BC. Decoding the Brain’s Algorithm for Categorization from Its Neural Implementation. Curr Biol. 2013; 23:2023–2027. DOI: 10.1016/j.cub.2013.08.035 [PubMed: 24094852]

101. Huys QJM, Maia TV, Frank MJ. Computational psychiatry as a bridge from neuroscience to clinical applications. Nat Neurosci. 2016; 19:404–413. DOI: 10.1038/nn.4238 [PubMed: 26906507]

102. Wiecki TV, Poland J, Frank MJ. Model-Based Cognitive Neuroscience Approaches to Computational Psychiatry Clustering and Classification. Clin Psychol Sci. 2015; 2167702614565359. doi: 10.1177/2167702614565359

103. Moustafa AA, Cohen MX, Sherman SJ, Frank MJ. A role for dopamine in temporal decision making and reward maximization in Parkinsonism. J Neurosci. 2008; 28:12294–12304. DOI: 10.1523/JNEUROSCI.3116-08.2008 [PubMed: 19020023]

*104. Chase HW, Kumar P, Eickhoff SB, Dombrovski AY. Reinforcement learning models and their neural correlates: An activation likelihood estimation meta-analysis. Cogn Affect Behav Neurosci. 2015; Meta-analysis of human imaging studies that mapped prediction error and expected value signals using reinforcement learning, accompanied by a brief overview of the implementation of reinforcement learning in functional imaging. doi: 10.3758/s13415-015-0338-7

105. Domjan M. Pavlovian Conditioning: A Functional Perspective. Annu Rev Psychol. 2005; 56:179–206. DOI: 10.1146/annurev.psych.55.090902.141409 [PubMed: 15709933]

106. Hawrylycz MJ, Lein ES, Guillozet-Bongaarts AL, Shen EH, Ng L, Miller JA, van de Lagemaat LN, Smith KA, Ebbert A, Riley ZL, Abajian C, Beckmann CF, Bernard A, Bertagnolli D, Boe AF, Cartagena PM, Chakravarty MM, Chapin M, Chong J, Dalley RA, Daly BD, Dang C, Datta S, Dee N, Dolbeare TA, Faber V, Feng D, Fowler DR, Goldy J, Gregor BW, Haradon Z, Haynor DR, Hohmann JG, Horvath S, Howard RE, Jeromin A, Jochim JM, Kinnunen M, Lau C, Lazarz ET, Lee C, Lemon TA, Li L, Li Y, Morris JA, Overly CC, Parker PD, Parry SE, Reding M, Royall JJ, Schulkin J, Sequeira PA, Slaughterbeck CR, Smith SC, Sodt AJ, Sunkin SM, Swanson BE, Vawter MP, Williams D, Wohnoutka P, Zielke HR, Geschwind DH, Hof PR, Smith SM, Koch C, Grant SGN, Jones AR. An anatomically comprehensive atlas of the adult human brain transcriptome. Nature. 2012; 489:391–399. DOI: 10.1038/nature11405 [PubMed: 22996553]

- Disadvantageous decisions are a key element of borderline personality
- Pavlovian, habitual, and goal-directed computational systems influence decisions
- Formal models test how these systems are instantiated in behavior and neural activity
- We propose that Pavlovian-to-instrumental transfer (PIT) plays a crucial role in BPD
- Interpersonal dysfunction may reflect the influence of neural circuits involved in PIT

Clinical phenomenon in BPD	Learning Theory Construct	Neural substrates
Top-down control over maladaptive behavior, possibly including shifts from reactive to deliberative decisions following treatment	Resistance to Pavlovian influences	Medial prefrontal cortex
Difficulty learning from negative consequences of interpersonal mistakes Limited mentalizing	Goal-directed learning	Dorsolateral prefrontal cortex Dorsomedial striatum Dorsolateral striatum
Goal-incongruent emotion-driven behaviors: - Interpersonal aggression - Self-injury - Impulsive substance use	Non-specific invigoration of instrumental responses by Pavlovian values (general PIT)	Nucleus accumbens core Central nucleus of amygdala Ventral tegmental area Dopaminergic midbrain
Goal-congruent emotion-driven behaviors: - Interpersonal manipulation	Enhancement of goal-congruent instrumental responses by Pavlovian values (specific PIT)	Nucleus accumbens shell Basolateral amygdala Medial orbitofrontal cortex Ventromedial prefrontal cortex

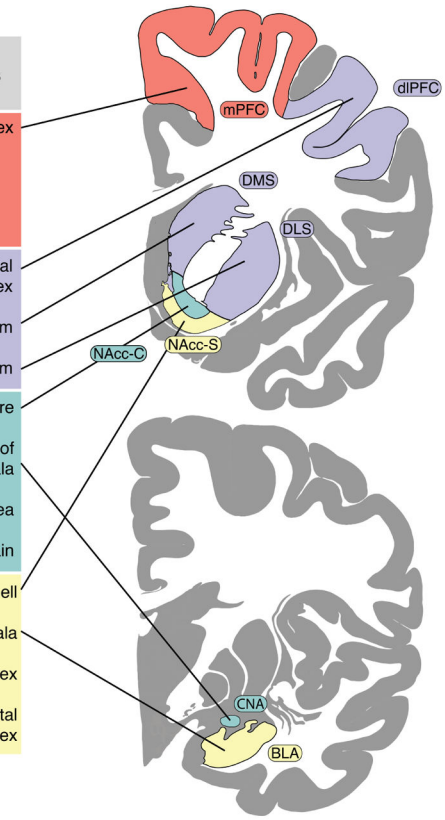


Figure 1.

Mapping clinical manifestations of borderline personality in terms of learning theory and neural systems

Brain images were adapted from the Allen Institute Human Brain atlas (http://atlas.brain-map.org/atlas?atlas=265297126) [106].