AN INTEGRATIVE MODEL OF PERSONALITY DISORDER: PART 3: MECHANISM-BASED APPROACH TO THE PHARMACOTHERAPY OF PERSONALITY DISORDER: AN EMERGING CONCEPT

Dragan Svrakic1,2, Aaron Mofsen1, Ravikumar Chockalingam2, Mirjana Divac-Jovanovic3 & Charles F. Zorumski1

1Department of Psychiatry, Washington University in St. Louis School of Medicine, St. Louis, USA
2Veterans Affairs Medical Center, Saint Louis, Missouri, USA
3Faculty for Media and Communications, Department of Psychology, Singidunum University, Belgrade, Serbia

received: 19.6.2019; revised: 7.8.2019; accepted: 20.8.2019

SUMMARY
Temperament traits of Novelty Seeking, Harm Avoidance, Reward Dependence, and Persistence, are well defined in terms of their neural circuitry, neurochemical modulators, and patterns of associative learning. When heritably excessive, each of these traits may become a mechanistically fundamental biogenetic trait vulnerability for personality disorder. The other main risk factor for personality disorder is environmental, notably abuse, neglect, and psychological trauma. The emerging concept of mechanism-based pharmacotherapy aims to activate the brain's homeostasis as the only available delivery system to re-calibrate complex neuro-physiological participants in each of the temperament traits. In a positive feedback, a homeostasis-driven improvement of excessive temperament is expected to facilitate maturation of neocortical networks of cognition, most reliably in expert psychotherapy (Part I of this paper) and, ultimately, thereby improve top-down cortical control of subcortical affect reactivity. As an emerging concept informed by neuroscience and clinical research, mechanism-based pharmacotherapy has the potential to be superior to traditional symptom-based treatments. Such mechanism-based approach illustrates what the pharmacological treatment of Research Domain Criteria (RDoC) might look like.

Key words: personality disorder - temperament - pharmacotherapy - RDoC

The human brain and mind: the distinguishing features of the “wise man”

The human mind is best conceptualized as a progressively differentiated, but initially global, agency that mediates the brain’s interactive communication with internal and external environments (Proust 2009). Personality refers to the unique configuration of the mind in a given individual, which is the definition now supported by functional neuroimaging (Gratton et al. 2018). As a self-organizing complex adaptive system (Svrakic et al. 1996), the mind (personality) develops through punctuated homeostasis, with periods of relative stability (maintained by homeostasis) interspersed with more or less abrupt transitions (“optimization”) towards a higher level of organization and functioning (“development”). These transitions occur whenever the previous equilibrium of the mind organization becomes adaptively suboptimal, a natural result of the interrelated processes of biological maturation of the brain and increasingly complex environmental experiences.

The self-organizing optimization of the mind involves two parallel and inter-related developments: i) the first one is the development of mental faculties, notably cognition, emotion, and motivation, aka “mental trilogy” (LeDoux 2002). Basically, this means the gradual introduction of complex cognitive functions and cognitively processed affects (now conscious emotions1) into mental life. Mental faculties can be thought of as operating “tools” of the mind, rooted in functional brain networks; ii) The second aspect of the self-organizing optimization of the mind involves a sequential differentiation of its subsystems, e.g., temperament traits, the sense of self, identity, character, etc. The subsystems are best conceptualized as semi-stable coalitions of mental faculties (cognition, emotion, motivation); they canonically emerge at a specific phase of mental development, each carries a specific adaptive function, and each contributes to the functioning of the mind as a whole. The term “semi-stable” means that the subsystems are dynamic, can change via optimization processes with the changing conditions or demands. Progressive differentiation of personality follows a predictable sequence: the first to emerge are temperament traits (the main avenue for communication in the 1st year of life), followed by the sense of self (the associative glue of the mind via self-referenced cognition), then character traits (which maximize the fit between one’s needs and social norms - aka “functional adaptation”), and finally the

1 We reserve the term emotions for conscious experiences of affects, which (the affects) have a plethora of non-conscious physiological and behavioral correlates
subsystem of principled moral values (these are not only involved in social adaptation but also in self-evaluation).

In contrast to the modular organization of the brain, which involves hierarchical modules of information processing, the subsystems of the mind also show modularity, but not in the strictest sense, rather in a more modest sense of i) dissociability (each is functionally specialized and phenomenologically discernible) and ii) central inaccessibility (referring to the restricted flow of information out of a subsystem whose internal operations may be inaccessible to the remainder of cognition”) (Carruthers 2006, p. 12). The canonical emergence of the subsystems roughly coincides with the biological maturation of the brain, most notably with the gradual self-organization of large scale brain networks. In contrast to such gradual brain development, the mind differentiates following the dynamics of punctuated homeostasis, with relatively abrupt transitions towards higher level cognition in terms of Piagetian stages, and adaptive self-organization into subsystems, as discussed.

Revisiting the neuroscience and of psychoanalysis

In part 2 of this series (Svrakic et al. 2019) we discussed the neuroscience of mind. Here we revise the discussion based on our revised understanding of the subject, most notably the role of the prefrontal cortex (PFCx).

Object relations theory: a synopsis

According to the object relation theory, founded by Melanie Klein, the first inhabitants of the mind are achieved by the mechanism of projective identification in the form of “object relations”. Beginning with birth, the newborn’s interactions with caregivers can be gratifying (normally predominant) or frustrating (normally less frequent, but inevitable). These early interactions activate subcortical primal affects (as defined by Panksepp), here most relevant are pleasure/lust or fear and/or aggression. These activated affects are projected onto the object2 who is thus perceived as rewarding (in the case of a person, partner, child, work, a religion, an organization...) or frustrating (normally predominant) or frustrating (normally in the form of “object relations”. Beginning with birth,

Almost all our emotional responses are moderates and manageable. However, many people experience intense and overwhelming emotions that can interfere with their daily lives. These intense emotions can cause significant distress and can negatively impact personal relationships, work, and overall well-being. The use of a mood stabilizer drug can help moderate these intense emotions and improve the quality of life for those who suffer from severe mood fluctuations.

Object relations theory refers to a psychological framework developed by Melanie Klein, which emphasizes the role of early childhood experiences in shaping adult personality. According to Klein, the development of the mind and the formation of the self are closely intertwined with the relationship to the object, or the external world. The object, or the object relations, are formed through the process of projective identification and project strong affects (a person, partner, child, work, a religion, an organization...). The projective identification is a mechanism by which the child's own emotional states are projected onto the external world, creating an internal representation of the object that is based on the experiences with caregivers.

Projective identification is a crucial process in the early development of the self and the formation of object relationships. It involves the child's emotional states being transferred to specific objects, such as caregivers or other significant figures in their life. This process helps the child to distinguish between the self and the environment, and to develop a sense of self and identity. For example, a child who is feeling happy and elated may project this positive emotional state onto a nurturing caregiver, while a child who is feeling angry or frustrated may project negative emotional states onto a dismissive or rejecting caregiver.

The process of projective identification is an ongoing and dynamic process throughout life. As the child grows and develops, their experiences with different caregivers and objects will shape their object relationships. The process of projective identification can also be influenced by cultural and societal factors, as well as personal experiences and traumas. Understanding the role of projective identification in the development of object relationships is crucial for psychologists and therapists working with children and adults alike.
"consults" the amygdala for the affective valence of the memories of that or similar contexts; all this information is pulled into working memory which eventually executes adequate behaviour or conscious emotions in the circumstances.

These anatomically distinct groups of amygdala neurons are genetically programmed to memorize only one affective valence of experience, rewarding or frustrating. The groups are anti-correlated to an extent (processing reward suppresses processing in the fear and anger dedicated neurons and vice versa), and fixed in their function, i.e., cannot be re-trained to engramize affects not in accord with their program (Rodendo et al. 2014). In contrast, the hippocampal neurons that encode context can be re-trained to "switch" and engramize the opposite context. Moreover, the hippocampal neurons with a new, switched context are shown to strengthen their connections with the group of amygdala neurons which register affects that match the new experience, while weakening the original connection with now obsolete amygdala affective valence of the previous that context. Incidentally, such hippocampal plasticity might be the mechanism of corrective emotional experience in psychotherapy: a new, positive experience within the same context (here interpersonal) may reverse the early hippocampal engrams of the negative context with caregivers and connect the new context with amygdala neurons recording positive affects. Such reversal now informs the working memory (and thus the conscious sense of self) and functional networks involved in cognition, emotion, and motivation, making both more balanced. Psychotherapy based on corrective emotional experience is discussed in part I of this series (Divac-Jovanovic & Svrakic 2017).

The prefrontal cortex: where engrams become concepts

Early in life (the first 6 months), while the neocortex is still immature to abstract the separation of positive and negative affective valence of the same context, polarized self-reflections and object perceptions are a normal phase in mind development, called "partial objects" by M. Klein. In fact, a case could be made that this "deficit" is the first organizer of the mind, sorting out positive from negative experiences. Eventually, the anatomical separation ("splitting" in psychoanalytic terms) is functionally bridged by maturing cognition into a balanced understanding of one's self and the world, aka "whole object relations". Such cognitive integration is an emerging functional property, an abstraction rather than a structural change. The integration is believed to occur in the PFCx, one of the highest order brain's convergence zones, together with the posterior parietal and parahippocampal cortices. The PFCx is also home to episodic memory engrams, including early interactions with caregivers as first engramized contexts of interactive experience. A convergence zone is defined as a highly connected brain region where successive levels of convergence of outputs from lower level modules of the brain culminate in maximally integrative regions. Convergence zones might well be the regions where first concepts are formed from neural processing of percepts, i.e., where neural becomes mental ("cognitive synthesis") (LeDoux 2002). Importantly, these convergence zones serve as major hubs shared by multiple large scale functional networks, including networks involved in cognition, emotion, and motivation. As noted, the trilogy of mental faculties can be thought of as the "tools" of the mind to understand itself, the outer world, and to adapt to both. In other words, the brain's convergence zones, as shared hubs in the brains functional connectivity networks, inform the brain's functional networks with either balanced (normal development) or polarized (fragmented) object relations (personality disorder) rendering cognition, motivation, and emotions either balanced or polarized, respectively. Since about 6 months of age, object relations continue to mature until about 24-36 months of age together with the maturation of convergence zones, working memory, and attention. Object relations can be thought of as elementary particles of the mind (others have called them the "basic units of all experience"), as each contains an object perception, the related self-reflection, and the connecting affect, and thus is not divisible to lower order mental units, just to lower order neural processes. These elementary particles have a dual and indivisible role as both the substrate and the function of the mind: as the first-order concepts of the mind (the substrate) with lasting qualitative effects (the "how" of the mind) they instantaneously become its function, its rudimentary thought, emotion, and motivation.

Working memory: the sketch-pad of mental life

In addition to being one of the main convergence zones in the cortex and one of the most connected hubs in large scale brain networks, the PFCx is also providing the function of working memory, this in concert with other regions of the brain, perietal cortex in particular. Working memory is a function, not a storage of memories. It temporarily pulls in relevant information from distributed brain regions in order to inform reasoning and guide decision-making and behaviour. Simplified, the PFCx, with the dorsal frontoparietal system of attention7 as the main tool, mediates top-
Personality disorder: developing the mind with unfinished elementary particles

Personality disorder develops as personality ordinarily would, albeit here in the context of a heritably heightened negative affect reactivity rooted in excessive temperament (biogenetic “trait vulnerability”) and/or traumatizing or abusive environments (“bad-enough environments”). In such a treacherous situation, negative experiences (most notably frustration, fear, and stress) predominate as the memorized context (specific interactions with caregivers, such as stress, fear, discomfort, among many others) with connections to the corresponding amygdala memories of fear and anger, associated with that context. Such negative experiences can interfere with normal development of prefrontal cognitive networks, and can even fracture the already formed ones (Zorumski & Rubin 2011). This developmental deficit interferes with functional merging of anatomically separated polarized auto- and autonoetic (episodic) experiences, called splitting in psychoanalytic, hypofrontality in neuroscience. Consequently, extremized (polarized, “either-or”) building particles of the mind (“partial object relations”) persist passed early childhood. The fragmented experience of self and others is incorporated into developing networks of mental faculties, rendering these polarized as well (cognition is “black” or “white”, emotions either idealization of devaluation, motivation either enthusiasm or inertia and disinterest). Finally, such polarized mental faculties underlie polarized (fragmented) subsystems of the mind: turbulent temperament, inconsistent sense of self, unstable identity, character that subserves inner needs for organization at the expense of social adaptation, and opportunistic, “personal” morality. As such, these subsystems, and the mind as a whole, are functionally suboptimal, and give rise to a host of symptoms unique to personality disorder, as discussed in Part 2 of this series (Svrakic et al. 2019).

This core defect of fragmentation is attributed to either the splitting mechanism in the psychoanalytic interpretation (a nonconscious mental defense that protects the few positive units of experience, which are vital for mind building) or to a stress-induced hypofrontal PFCx in the neuroscientific interpretation (as noted, trauma and stress interfere with the maturation of cortical networks and can even fracture the already formed ones), or perhaps the former is merely the function of the latter. Ultimately, psychoanalytical theory and neuroscience are describing the same phenomenon; a deep and early fragmentation of the mind at its nonconscious core of internalized object relations is the central impairment shared by all clinical variants of personality disorder. Despite such a profound defect, as a phylogenetic given, a mind has to develop in a human being, even if the building units are less than perfect (here polarized). This is known as “developmental homeostasis”, a biological process in which all animals develop more or less normally within a range, despite defective genes and deficient environments. Here, homeostasis makes use of fantasy to create an uncontested virtual world of psychological power and safety.

Fantasy and imagination: very close, but not synonymous

Fantasy is defined as conscious mental activity of imagining things or events, especially those that are impossible or improbable, in response to psychological needs, typically emotional needs ("wishful unreality", "wishful thinking", or primary process in psychoanalysis10). Fantasy may include elements of deeper nonconscious phantasies (M. Klein’s concept of phantasy was defined earlier), and this is one of the differences between fantasy and imagination. Imagination is defined as conscious mental activity of forming new connections, or images or concepts of one self and external objects not present to the senses. The capacity to imagine emerges early in the course of childhood development ("pretend play" beginning cca. 3 years of age) and has an important role in children's developing conceptions of reality. A. Einstein stated that "Imagination is more important than knowledge. For knowledge is limited, whereas imagination embraces the entire world, stimulating progress, giving birth to evolution." It has been shown (Dias & Harris 1990) that children, 4-6 years of age, will accept premises that violate their empirical knowledge as a basis for reasoning so long as they are presented in a make-believe mode rather than a literal mode. There is a subtle but important difference between fantasy and imagination, although they may overlap: imagination refers to the ability to mentally create events or things that are not currently experienced. Fantasy is generally applied only to those products of our imagination that are impossible or unlikely to happen in reality. Imagination can be used, and frequently is, by creative or resourceful individuals. Fantasy is used, among other examples, by children experiencing early adversity for defensive purposes and, in adults, by individuals with the schizoid variant of personality disorder as a form of unilateral adaptation.

When defensive fantasy comes to rescue: creation of a unilateral “good” experience

As noted, in personality disorder, mental faculties are built with fragmented elementary particles, with the predominance of “bad” self- and object-images inside. Both can only be treated by the infant’s early mind as “Not I” (per M. Klein) and thus cannot represent the core around which the mind is organized. In fact, partial objects and persecutory anxieties render personality and its subsystems vulnerable to psychological annihilation (analogous to the physical fear of dying) and adaptively suboptimal, as discussed. Such polarized mind is unable to establish a realistic coupling with its environments, internal and external, but rather alternates between extreme experiences, typically persecutory anxieties and aggression predominate over occasional primitive idealizations. The syndrome of personality disorder arises as a homeostatic attempt of the early mind to organize the fragments through the power of defensive fantasy, such as narcissistic grandiosity, histrionic theatricality, schizoid withdrawal, paranoid suspiciousness, antisocial need for dominance, etc. The fantasized defensive image is then expected to be validated by others, leading to volatile interpersonal conflicts if the validation does not happen, even more if others criticize or devalue the individual.

As all fantasy is conscious11, homeostatic compensation through fantasy has to occur within working memory. It must involve attention-based access to long term memories and affects involving self and others. Once information (memory engrams, affects) is pulled into nonconscious working memory (the first step on the way to becoming fully conscious) it can be manipulated by emotional these needs, leading to primary process in cognition where reality testing and logic are suppressed by emotional needs. Thought cannot exist without associated emotions, but affects can function without cognition, in fact here they overpower cognition and logic. Within nonconscious working memory, emotional needs rearrange the reality so that the original negative interactive experience is turned fantasized into rewarding or revengeful, all in order to create a virtual episode in one’s experience culminating in comforting emotions (pleasure, sense of safety) and a sense of safer self. Such good experiences, albeit virtual and unilateral, now can become accepted as “Me”, albeit highly unrealistic, a positive core around which the mind continues to develop. To sum, defensive fantasy is a fabrication of the mind perhaps unique to humans to create an imaginary solution for real life psychological danger and suffering.

Different clinical variants of personality disorder merely express an impersonation of specific compensatory fantasized solutions to the central impairment. As early as 3–4 years of age, defensive fantasies and other symptoms of compensatory mental functioning are observable in children (reviewed in by Svrakic & Divac-Jovanovic 2019). The homeostatic compensation gives rise to a more stable albeit unrealistic self-image (narcissistic, antisocial, etc.). Such precarious self image continually needs re-affirmation and maintenance, which is accomplished internally (through fantasy) and interpersonally (through a lifetime need for selfobjects12).

---

10 In contrast, daydreaming associated with the default mode network relies on reality rather than on emotional needs, although it may also include imagination

11 M. Klein introduced the concept of “phantasy”, as a nonconscious, preconceptual image potential of the mind, a psychological facet of biological instincts that links affects to objects and represents the means by which infants make sense of the external world and relate to it through projection and introjection.

12 Kohut (1971) introduced the concept, referring to the infant’s relationships with caregivers whose psychological functions (such as approval, support, etc.) define and maintain the sense of the infant’s self; through the process of “transmuting internalization” (i.e., graded frustration), selfobjects and their functions are gradually replaced by self-regulation of self-image, self-esteem, mood, and goals.
The latter refers to people who are expected, frequently demanded to maintain the unrealistic self-image of the patient, hence the frequent interpersonal conflicts. The above described integrative model of personality disorder is discussed in detail in the Svrakic and Divac-Jovanovic book on “The Fragmented Personality” by Oxford University Press (2019).

**Treatment of personality disorder: top down and bottom up.**

As noted in Part 2 of this series (Svrakic et al. 2019), both brain and mind aberrances characterize personality disorder: i) the former refer to hypofunctional cognitive networks and hyper-reactive subcortical affects, the latter to ii) fragmented affects and self- and object-images (reviewed in Svrakic & Divac-Jovanovic 2019). It is widely accepted that optimal treatment strategy for individuals with personality disorder combines expert pharmacotherapy and in-depth, reconstructive psychotherapy. The latter uses corrective emotional experience within a safe therapeutic relationship (aka “re-parenting”) to help the patient merge the polarized mind into a realistic whole (the therapist is sometimes referred to as the “third parent”). In this sense, psychotherapy is a mechanism-based approach to treatment. Pharmacotherapy is typically either mechanism-based or symptom based, although the two may overlap. In psychiatry we are most familiar with using medications to control symptoms, which frequently does not lead to lasting and/or core improvement in patients with personality disorder. We argue here that the underlying neural mechanisms (here referring to subcortical circuitry of excessive temperamental traits), which are responsible not only for steering suboptimal brain and mental development but also of perpetuating symptoms of personality disorder once it has developed represent a better, mechanism-based pharmacologic target (discussed in more detail below).

Pharmacotherapy and psychotherapy appear to approach personality disorder from opposite directions: i) to establish a better cortical top down control of excessive subcorticality by psychotherapy, or ii) to target selected subcortical affects and impulses thereby freeing the cortex to self-organize into more mature cognitive networks with a more effective executive function. The latter can occur either serendipitously, through real life experiences\(^{13}\), or more reliably in reconstructive psychotherapy. As noted in Part II of this series (Svrakic et al. 2019), the two approaches will inevitably engage in a positive feedback where improved top down control perpetuates more effective cognitive functions and vice versa. Any pharmacotherapy, whether mechanism-based or symptom-domain based, is unlikely to be effective unless combined with psychotherapy, i.e., that a combined top-down and bottom-up approaches provide maximal effectiveness.

Improvement of the cognitive top-down control, presumably reflecting maturation of prefrontal networks, has been thus far the domain of psychotherapy. Recently, there has been a renewed interest in pharmacological modulation of cortical networks/functions, using agents classified as psychedelics (discussed in more detail at the end, as emerging options for top-down pharmacotherapy of personality disorder).

**Mechanism-Based Pharmacotherapy: An Emerging Concept**

Mechanism-based pharmacotherapy of personality disorder draws from the complex neurobiology of temperament traits which, when excessive, may act as risk factors to set in motion aberrant brain and mind development, resulting in personality disorder (Luby et al. 2013). Fortunately, this pathodevelopment is reversible and thus amenable to correction in therapy, both in terms of its neurological and mental aberrances, e.g., improved working memory and a more integrated sense of self (Krause-Utz et al. 2017). Temperament refers to the psychomotor characteristics and affective reactivity of the organism, such as susceptibility to fear, anger, and sociability. These are turned into enduring traits via associative conditioning early in life. Many of these vital biogenetic features are traceable to other animals, cobbled through evolution from simple instincts mediated by the ancient midbrain monoamines to more complex, contextualized behaviors that involve additional contributing neurocircuits, neuroactive molecules, and the neocortex. In humans, complex temperament traits originate in the midbrain but are comodulated by multiple, if not all neuroactive molecules and shaped by coordinated activities across multiple brain regions. For this reason, a precise widespread intervention is needed to address all participants. While we currently have neither the knowledge nor the technology necessary to design a multitarget-multistep yet functionally specific drug, we may instead be able to harness homeostasis to achieve the same end.

Homeostatic optimization can be set in motion by an abrupt disturbance of any complex biological equilibrium, whether pathological or healthy. The disturbance must be either mechanistically fundamental (“distal”) in order to engage and recalibrate other contributing (“proximal”) component processes of a targeted complex trait, or general enough to affect all components simultaneously. Examples of each approach

---

\(^{13}\) E.g., with a mature loving partner, who would represent figuratively a belated selfobject who mediates inner integration in the context of pharmacologically reduced trait vulnerability, a rare but fortunate milieu for the fragmented partner.
include midbrain monoamines (e.g., serotonin) and electroconvulsive therapy (ECT), respectively. Although the precise mechanisms of action of ECT are not known, it is plausible that a series of electrically-induced generalized seizures may activate the brain’s homeostasis and re-arrange the pre-existing neuro-pathological balance of depression into a healthier equilibrium. The class of SSRIs provide another example of homeostatic recalibration: here, the initial increase in extracellular serotonin, a mechanistically fundamental molecule in anxious temperaments, triggers weeks of homeostatic neuroadaptive processes which eventually downregulate the central serotonergic system. The so called “network scaling” of the serotonin system could be one of the mechanisms of recalibration. The “scaling” refers to the process whereby a hyperactive network, here further “agitated” by acute serotonin increase, decreases its activity by scaling down the activity of individual synapses while preserving the relative weight of each synapse in the network. As serotonergic neurons touch pretty much every other neuron in the brain, such scaling down will inevitably have widespread effect across the brain. In sum, such recalibration is believed to be responsible for the SSRI’s efficacy in depression and anxiety states, now both believed to be “hyper-serotonergic” (Andrews et al. 2015).

Temperament: the biogenetic backbone of personality in health and disorder

In the canonical course of mind development, temperament traits are the first motivated behaviors to emerge. They are observed as individual differences among infants in fearfulness, sociability, persistence, and impulsivity which are assessable by self-report around the age of five to six years (Luby et al. 1999). Subsequent cognitive development contextualizes the early temperament of childhood into adult temperament traits which influence the stylistic aspect of behavior (reminiscent of Galen’s choleric, sanguine, melancholic, or phlegmatic styles) as well as one’s phenomenally conscious preferences, such as motivational biases to initiate, stop, persist, or maintain behavior (Cloninger et al. 1993).

Cloninger (1987) pioneered a tridimensional model of temperament (Harm Avoidance - HA, Novelty Seeking - NS, Reward Dependence – RD), later adding Persistence (PE) based on empirical testing of the original model. Each of the four traits is postulated to arise from discernible neurophysiological systems of learning and memory. Table 1 presents a simplified summary of complex, functionally coregulated, and anatomically overlapping neurophysiological systems involved in temperament traits (modified from Cloninger & Svrakic 2017).

Human temperament traits are continuous with no clear natural breaks that separate trait normalcy from its pathological extremes. Livesley et al. (1998) showed the clinical construct of personality disorder arises from four underlying latent dimensions (a) emotional dysregulation, (b) antisocial and impulsive, (c) inhibition with emotional disinterest, and (d) compulsivity. These four latent dimensions of personality disorder fully correspond to the extreme ends of normal temperament, high HA, high NS, low RD, and high PE, respectively. Importantly, the latent trait structure was also found to closely reflect the underlying genetic architecture of personality disorder (Livesley et al. 1998). The “big four” appear to be the natural building blocks of a true biological organization of personality (more precisely of temperament) and personality disorder (excessive temperament traits as trait vulnerabilities) (Table 2).

### Table 1. Four dissociable brain systems underlying temperament

<table>
<thead>
<tr>
<th>Temperament Dimension (Brain System)</th>
<th>Fundamental (“root”) neurophysiological system</th>
<th>Relevant stimulus</th>
<th>Behavioral response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harm Avoidance (Inhibition of behavior)</td>
<td>GABA Serotonin (dorsal raphe)</td>
<td>Aversive conditioning (pairing CS and UCS)</td>
<td>Passive avoidance (shyness)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Extinction (fatigability)</td>
<td>Anticipatory worry (anxiety)</td>
</tr>
<tr>
<td>Novelty Seeking (Initiation of behavior)</td>
<td>Dopamine</td>
<td>Novelty CS of reward</td>
<td>Exploratory excitability</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CS/UCS of punishment</td>
<td>Impulsiveness</td>
</tr>
<tr>
<td>Social Reward Dependence (Activation of social behavior)</td>
<td>Norepinephrine Serotonin (median raphe)</td>
<td>Social reward (UCS)</td>
<td>Active Avoidance/Escape</td>
</tr>
<tr>
<td></td>
<td>Oxytocin Dopamine</td>
<td></td>
<td>Attachment</td>
</tr>
<tr>
<td></td>
<td>Dopamine Endogenous opiates</td>
<td></td>
<td>Sentimentiality</td>
</tr>
<tr>
<td>Persistence (Maintenance of behavior)</td>
<td>Glutamate Serotonin</td>
<td>Intermittent (partial) Reinforcement</td>
<td>Resistance to extinction despite</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Frustration and fatigue</td>
</tr>
</tbody>
</table>

GABA - gamma aminobutyric acid; UCS - unconditioned stimulus; CS - conditioned stimulus (modified from Cloninger 1987)
Table 2. Correspondence among neurobiological temperament traits and latent genotypic-phenotypic dimensional construct of personality disorder

<table>
<thead>
<tr>
<th>Neurobiological trait vulnerability (Cloninger 1987)</th>
<th>Latent dimensional genotypic-phenotypic architecture (Livesley et al. 1998)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harm Avoidance</td>
<td>Emotional Dysregulation</td>
</tr>
<tr>
<td>Worry and pessimism</td>
<td>Anxiousness</td>
</tr>
<tr>
<td>Fear of uncertainty</td>
<td>Submissiveness</td>
</tr>
<tr>
<td>Shyness</td>
<td>Identity problems</td>
</tr>
<tr>
<td>Fatigability</td>
<td>Social avoidance</td>
</tr>
<tr>
<td></td>
<td>Insecure attachment</td>
</tr>
<tr>
<td></td>
<td>Cognitive dysregulation</td>
</tr>
<tr>
<td></td>
<td>Affective lability</td>
</tr>
<tr>
<td></td>
<td>Oppositionality</td>
</tr>
<tr>
<td></td>
<td>Suspiciousness</td>
</tr>
<tr>
<td>Novelty Seeking</td>
<td>Dissocial</td>
</tr>
<tr>
<td>Exploratory excitability</td>
<td>Stimulus seeking</td>
</tr>
<tr>
<td>Impulsiveness</td>
<td>(impulsivity, recklessness, sensation seeking)</td>
</tr>
<tr>
<td>Extravagance</td>
<td>Callousness</td>
</tr>
<tr>
<td>Disorderliness</td>
<td>Interpersonal Rejection</td>
</tr>
<tr>
<td></td>
<td>(hostile, judgmental)</td>
</tr>
<tr>
<td></td>
<td>Conduct problems</td>
</tr>
<tr>
<td>Reward Dependence (Low)</td>
<td>Inhibition</td>
</tr>
<tr>
<td>Practical (non-sentimental)</td>
<td>Intimacy problems</td>
</tr>
<tr>
<td>Aloof</td>
<td>Restricted expression</td>
</tr>
<tr>
<td>Cold</td>
<td>(restricted affects, difficulty sharing)</td>
</tr>
<tr>
<td>Independent</td>
<td></td>
</tr>
<tr>
<td>Persistence</td>
<td>Compulsivity</td>
</tr>
<tr>
<td>Industriousness</td>
<td>Self-discipline</td>
</tr>
<tr>
<td>Determination</td>
<td>Dutifulness</td>
</tr>
<tr>
<td>Ambitiousness</td>
<td>Ambition</td>
</tr>
<tr>
<td>Perfectionism</td>
<td>Lack of spontaneity</td>
</tr>
</tbody>
</table>

Based on the high correspondence between heritable dispositions to temperament and latent genotypic and phenotypic trait architecture of personality disorder (Table 2) one may speculate their causative relationship in which excessive temperament limits mature outcomes in mental development, which is accord with our prior work (Svrakic et al. 1996). This makes pharmacotherapy of temperament a bona fide mechanism-based treatment.

**Excessive temperament: the main target for mechanism-based pharmacotherapy**

In humans, subcortical networks involved in instinctive social behaviors in animals (Newman 1999) usually work in conjunction with cortical networks, making temperament traits both phenomenally (passively) conscious, as the first person experience “I know how it feels to like what I like” as well as accessible to introspection (“I know what I like”). Temperament can also function independently from cognition, i.e., nonconsciously, bypassing the cortex, especially in emergency situations. Given its possible functional independence from the neocortex and the relatively well-defined subcortical circuits, temperament may be viewed as a relatively clean pharmacological target. For example, whole brain resting-state fMRI data was used to develop connectome-based neural models, involving cortex and subcortex, reliably predictive of the person’s temperament traits of NS, HA, RD, and PE (Jiang et al. 2018). The conceptual basis of mechanism-based pharmacotherapy is to trigger homeostatic optimization processes that may lead to a new and healthier equilibrium of an excessive complex trait. Because medications do not directly affect the fragmented personality core, pharmacotherapy is most effective if combined with expert psychotherapy akin to re-parenting, as discussed.

Table 3. Temperament profiles of personality disorder variants with major trait vulnerabilities, co-risk traits, and modifying traits

<table>
<thead>
<tr>
<th>Personality disorder variant</th>
<th>Harm Avoidance</th>
<th>Novelty Seeking</th>
<th>Reward Dependence</th>
<th>Persistence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antisocial</td>
<td>low** (modifying trait)</td>
<td>high* (main trait vulnerability)</td>
<td>low** (modifying trait)</td>
<td>variable</td>
</tr>
<tr>
<td>Histrionic</td>
<td>low** (modifying trait)</td>
<td>high* (main trait vulnerability)</td>
<td>high** (modifying trait)</td>
<td>low</td>
</tr>
<tr>
<td>Narcissistic</td>
<td>high* (main trait vulnerability)</td>
<td>variable (modifying trait)</td>
<td>high** (modifying trait)</td>
<td>high</td>
</tr>
<tr>
<td>Anankastic</td>
<td>high* (main trait vulnerability)</td>
<td>low (modifying trait)</td>
<td>moderate (co-risk factor)</td>
<td>high** (variable)</td>
</tr>
<tr>
<td>Negativistic-dependent</td>
<td>high* (main trait vulnerability)</td>
<td>low (modifying trait)</td>
<td>high** (modifying trait)</td>
<td>variable</td>
</tr>
<tr>
<td>Schizoid</td>
<td>low** (modifying trait)</td>
<td>low** (modifying trait)</td>
<td>low* (modifying trait)</td>
<td>low** (modifying trait)</td>
</tr>
<tr>
<td>Paranoid</td>
<td>high* (main trait vulnerability)</td>
<td>variable (modifying trait)</td>
<td>low** (modifying trait)</td>
<td>high** (modifying trait)</td>
</tr>
</tbody>
</table>

* - Main trait vulnerabilities; ** - Co-risk traits and modifying traits
The most frequent trait vulnerabilities are excessive HA (fearfulness), excessive NS (anger, impulsivity) and excessively low RD (detachment). While excessive PE in itself does not appear as trait vulnerability for personality disorder, it can, if coupled with excessive HA, become a co-risk factor for the anankastic variant, as discussed later. Table 3 illustrates configurations of temperament traits specific for individual variants of personality disorder, in which some traits act as primary vulnerabilities for a variant, while others act as modifying (shaping the clinical picture) and co-risk factors (only when coupled with other excessive dispositions) both contributing to the phenomenology of that particular subtype.

Below we review mechanism-based pharmacotherapy for excessive HA, NS, PE, and of excessively low RD (Table 4).

### Mechanism-based pharmacotherapy of trait vulnerabilities of personality disorder

This trait includes four facets involving passive avoidance of harm: fear of uncertainty, fatigability, social inhibition (shyness) and worry, and pessimism in anticipation of problems (Cloninger et al. 1993). As trait vulnerability, heritably excessive HA increases the risk for the majority of personality disorder variants (four out of seven: narcissistic, anankastic,14 negativistic-dependent15, paranoid). In many instances, an in-tuned, “good-enough” parenting especially when coupled with other protective factors (perhaps biological, like high RD) may foil the development of personality disorder in the context of excessive HA. In this case, this is observed as anxious temperament (children) or generalized anxiety disorder – GAD (adults), in both diagnoses with completed core integrations of the mind, i.e., without fragmentation of the mind typical of personality disorder.

Trait anxiety may be defined as mental and physiological preparedness for danger, with arousal, vigilance, negative subjective state, etc. (all also observed in fear) but here in the absence of an immediately threatening stimulus (which makes it different from fear). The psychobiology of HA is believed to be rooted in hypersensitive subcortical circuits involving the bed nucleus of the stria terminalis (BNST) as the main nexus. The BNST is considered an extended region of the amygdala, has a similar cellular organization and similar projections as the central nucleus of the amygdala (CNA), and is also capable of synaptic plasticity, i.e., can learn anxiety responses. However, the BNST may respond to more general cues of danger in comparison to the CNA, including non-imminent abstract threats16 (Nestler et al. 2012). Corticotropin releasing factor dependence and passive aggression regularly co-occur in this particular variant.

14 The term “anankastic” reduces possible confusion with OCD and captures the essence of this variant - from Greek "ἀνανακτικός" - anankastikos, for “compulsive” rigidity and rule orientation.

15 We combine the DSM passive aggressive and dependent subtypes into the negativistic- dependent variant of personality disorder, based on our clinical observations that excessive

### Table 4. Mechanism-based pharmacotherapy of trait vulnerabilities of personality disorder

<table>
<thead>
<tr>
<th>Trait Vulnerability</th>
<th>Fundamental neurophysiological mechanism</th>
<th>Mechanism-based treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Harm Avoidance</td>
<td>Overactive midbrain serotonergic system</td>
<td>Classic SSRIs (escitalopram, sertraline) “Expanded” SSRIs (vilazodone, vortioxetine)</td>
</tr>
<tr>
<td></td>
<td>Underactive GABA modulation of sub cortical &amp; midbrain anxiety signaling</td>
<td>GABA enhancers: tiagabine, vigabatrin, valproate, pregabalin, gabapentin</td>
</tr>
<tr>
<td>High Novelty Seeking</td>
<td>Overactive mesolimbic system</td>
<td>Partial D2/D3 agonists: aripiprazole, cariprazine</td>
</tr>
<tr>
<td>Low Social Reward</td>
<td>Hypoactive oxytocinergic system</td>
<td>Partial D4 agonists (flibanserin?)</td>
</tr>
<tr>
<td>Dependence</td>
<td>Hypoactive noradrenergic system</td>
<td>SNRIs: duloxetine, milnacipran, levo-milnacipran</td>
</tr>
<tr>
<td></td>
<td>Hypoactive noradrenergic system</td>
<td>MAOIs (phenelzine)</td>
</tr>
<tr>
<td></td>
<td>Hypoactive dopaminergic system</td>
<td>Bupropion (DA and NE)</td>
</tr>
<tr>
<td>High Persistence</td>
<td>Overactive glutamatergic “switch”</td>
<td>Partial µ-receptor agonists (buprenorphine)</td>
</tr>
<tr>
<td></td>
<td>(positive bias in reward perception)</td>
<td>Mild NMDA antagonists (ketamine)</td>
</tr>
</tbody>
</table>

GABA - gamma amino butyric acid; DA - dopamine; NE – nor epinephrine; MAOIs - monoamine oxidase inhibitors; SSRIs - selective serotonin reuptake inhibitors; NRI - norepinephrine reuptake inhibitors; SNRIs - serotonin norepinephrine reuptake inhibitors; D2/D3 - dopamine D2 and D3 type receptors; NMDA - N-methyl-D-aspartate
(CRF) receptor type 2 signaling in this region has been implicated in the regulation of anxiety and the stress response (Henckens et al. 2017). An overactive serotonergic modulatory system and/or underactive inhibitory gamma-aminobutyric acid (GABA) neurotransmission are thought to be mechanistically fundamental to define individual differences in HA. Mechanism-based pharmacotherapy targets either or both systems.

**Does homeostasis underlie the therapeutic effects of the SSRIs?**

Classical and “expanded” selective serotonergic reuptake inhibitors (SSRI), the latter referring to agents like vilazodone and vortioxetine (both have additional mechanisms in addition to SSRI), among others, are routinely used to treat trait biogenetic anxiety associated with excessive HA. The precise anxiolytic and antidepressant mechanisms of long-term SSRI administration are not fully understood. Both syndromes are “hyper-serotonergic” but also involve cortical networks and other neuromodulators (dopamine, norepinephrine) and neurotransmitters (glutamate, GABA) as contributing, executive, and control mechanisms, respectively. We submit here that serotonergic medications modulate HA, and other anxiety and depression syndromes, by means of homeostasis. SSRIs acutely, within hours, increase extracellular 5-HT by inhibiting its reuptake from the synaptic cleft. Despite this acute effect, SSRIs take weeks to impact both syndromes. In fact, SSRIs can worsen anxiety and depression in the first week of treatment, as would be expected in the context of an already overactive serotonergic system. One explanation for the 4-6 week therapeutic delay may involve the process of homeostatic optimization as an adaptive response: the acute increase (“disturbance of the pathological equilibrium”) in extracellular 5-HT leads to neuroadaptative desensitization (lower response threshold) and/or downregulation (decreased number) of 5-HT2A receptors and tonic activation of 5-HT1A/B somatodendritic autoreceptors on both DRN and postsynaptic inhibitory hetero-receptors in many brain regions. These adaptive responses coincide with adaptations in other circuits (Andrews et al. 2015). An interesting hypothesis points to energy homeostasis as the engine behind the widespread adaptive changes that are stirred by SSRI’s, improving in the process high-energy states of depression and anxiety (Andrews et al. 2015). Alternative explanations for the delayed response to SSRIs include neuroplastic and structural changes in neurons, such as but not limited to network scaling (as discussed), as well as neurogenesis in the dentate gyrus of the hippocampus that occur with most effective antidepressants.

**Underactive GABA underlies excessive HA: when the control mechanism fails**

GABA is the primary inhibitory “controller” of brain neurophysiology and, in general, suppresses affects and impulses, most notably anxiety and aggression. The use of GABAergic drugs to reduce excessive HA thus may be thought of as a “control-mechanism” based approach, because deficient GABA control may turn an otherwise normal neurobiology of HA into functionally excessive, perhaps to the same degree as heritably high HA. Of note, the same can be said for other affect-behavior states, notably anger-impulsivity associated with NS. However, GABA modulation of impulsivity occurs mostly in the anterior cingulate cortex, a contributing region involved in impulse control by performance monitoring and error detection. This process is “proximal” to the DA overactivity in the shell region of NAc (Jupp et al. 2013) which is the mechanistically fundamental mechanism of impulsivity.

Individuals with GAD17, have a downregulated GABA activity in key brain regions in the network of anxiety signaling, including the amygdala, hippocampus, and PFCx (Lydiard 2003) and lower levels of plasma GABA than healthy controls, at least in some studies (Roy-Byrne 2005). With respect to the pharmacotherapy of excessive HA, positive allosteric modulators of GABA receptors are rarely used as they carry abuse/addiction and other risks. More frequently used are GABAergic drugs which increase extracellular GABA, either by blocking its reuptake (tiagabine) or by inhibiting its degradation by GABA transaminase (vigabatrin) or through multiple or different mechanisms (valproate, pregabalin, gabapentin). Pregabalin, for example, has been approved in Europe for use in GAD/high HA individuals. In summary, low to moderate doses of GABAergic drugs effectively dis inhibit passive avoidance, reduce apprehension, modulate the ruminative thinking, and improve other facets of excessive HA, presumably by establishing a stricter control of the subcortical circuitry involved in HA.

**Mechanism-Based Pharmacotherapy of Trait Vulnerability Associated with High NS**

Human motivation is associated with seeking novelty as an independent motivator, with or without reward (Krebs & Davies 1997). In fact, both NS and social RD (another primary behavior motivator) are related components of the human motivational system. Both NS and RD are in large part modulated by the brain’s mesocorticolimbic DA system, albeit with some important differences as discussed below.

The complex psychobiology of NS is postulated to arise from a heritably strong disposition to an appetitive approach, which is conditioned during early childhood into a stable propensity to initiate behavior (called NS) in response to novelty, signals of reward, active avoidance of conditioned signals of punishment, and escape

---

17 The study included subjects with Generalized Anxiety Disorder (GAD), i.e., those with heritably high HA but without personality disorder, as discussed
from unconditioned punishment (Cloninger et al. 1993). Its facets are exploratory excitability, impulsivity, extravagance, and disorderliness. Of note, the trait combines desirable aspects (exploratory curiosity) and undesirable aspects (impulsivity). Curious and exploratory individuals are not necessarily impulsive, and can in fact be rather thoughtful and thorough. It is high impulsivity that acts as a trait vulnerability to the development of personality disorder, especially to the antisocial and histrionic variants (Table 3).

Functional MRI (fMRI) studies show increased striatal activity of (presumably) DA neurons triggered by novel stimuli regardless of whether they predict reward or not (Krebs & Davies 1997). High scores on NS are associated with reduced D2/D3 autoreceptor control of the midbrain dopaminergic nuclei (Zald et al. 2008), leading to accentuated dopaminergic responses to novelty. At least in some studies, the novelty-induced increased striatal DA release occurs in tandem with higher density of DA-reuptake transporters in the region and increased reuptake of DA at presynaptic terminals (Laine et al. 2001). In the context of high DA reuptake, frequent and novel sensorimotor stimulation is required to maintain optimal levels of postsynaptic stimulation by DA, which can be thought of as “stimulus hunger” to maintain optimal levels of postsynaptic stimulation. In the context of high DA reuptake, frequent and novel sensorimotor stimulation is required to maintain optimal levels of postsynaptic stimulation by DA, which can be thought of as “stimulus hunger” and a plausible explanation for impulsivity. At the risk of oversimplifying, it may be said that NS is primarily modulated by mesolimbic DA pathways originating in the VTA and involving the NAc shell (Jupp et al. 2013).

**Dopaminergic drugs modulate high NS**

At present, the first-line medications to modulate excessive NS are partial agonists at D2/D3 receptors, due to their unique ability to target the postulated DA overactivity without much disturbance of the DA-driven reward system. The candidate medications are aripiprazole and the related cariprazine, although neither has been tested for the indication. Aripiprazole has been referred to as a “biased ligand” and “smart drug” to illustrate its unique DA action profile: it can act as a full antagonist, partial antagonist, and partial agonist at D2/D3 receptors. The type of action is determined by local DA levels, cell type, and functional and signaling status of the target cells. In addition to the partial D2/D3 autoreceptor agonism, aripiprazole can act as a partial or full DA antagonist at other types of DA receptors in the abundant presence of DA, the latter a postulated neurophysiology of high NS, and this may reduce high NS in a mechanism-based manner as well. Finally, D4 receptor polymorphisms, notably the C allele which is associated with increased DA activity, have also been implicated in NS (Munafo et al. 2008). Flibanserin, a weak partial agonist at the D4 receptor (and thus a functional antagonist in the context of increased DA activity that underlies high NS), to our knowledge has not been tested to modulate high NS.

**Not all impulsivity is created equal: high NS vs Attention Deficit Hyperactivity Disorder (ADHD)**

The overactive DA neurotransmission in high novelty seekers not only makes people hungry for novelty (the limbic part) but also affects cognitive functioning (the cortical part). The effect on cognition is based on the inverse U-shape pattern of optimal levels of neocortical DA and NE, whereby small changes (up or down) can produce large effects (Arnsten & Rubia 2012). High NS individuals (with high basal DA activity) and those with ADHD (with low basal DA and NE activity) can both manifest ADHD-like features (impulsivity, distractibility). Psychostimulants increase central DA and NE and are effective in bona fide ADHD but may worsen impulsivity and/or distractibility in children with high NS (with an already high DA), perhaps contributing to the approximately 25% of ADHD-like children who do not respond to psychostimulants.

**Mechanism-Based Pharmacotherapy of Trait Vulnerability Associated with High PE**

The temperament trait of PE is defined as a conditioned bias in the maintenance of behavior despite failure, frustration, fatigue, and inconsistent reinforcement, with facets of industriousness, determination, ambition, and perfectionism (Cloninger et al. 1993). High PE by itself is unlikely activate excessive negative affective reactions or disturb early relationships with caregivers. Thus, this temperament trait does not appear as a major trait vulnerability for personality disorder, but rather a co-risk factor for the anankastic variant where it acts in concert with high HA. Here, HA is the main trait vulnerability, as discussed.

The trait of PE can be measured by the partial reinforcement extinction effect (PREE), a phenomenon in which partially reinforced operant responses persist longer in extinction than those acquired under continuous reinforcement. The integrity of the PREE depends on glutamatergic projections from the hippocampal subiculum to the NAc (Rawlins et al. 1989). High PE individuals manifest a positive bias in reward perception whereby neutral stimuli are perceived as more intrinsically motivating (Gusnard et al. 2003). The ability to modify internal motivation makes high PE a positive influence on other temperament traits, such as impulsivity (high NS) or passivity (high HA). With respect to personality disorder, excessive PE, notably the related compulsive perfectionism, may act to organize the fragmented core in the anankastic variant of personality disorder, for which the main risk factor is excessive HA. Here, high PE focuses the patient’s internal motivation on the external rather than...
on internal environment, defusing the problem in the process. The feeling of being in control of external circumstances soothes the catastrophic expectations associated with high HA.

**Pharmacological modulation of high PE with ketamine**

The literature on this subject is scarce at best. Medications with some anti-glutamatergic effects (riluzole, memantine, lamotrigine, ketamine) may help disrupt the hippocampal-accumbal glutamatergic switch which is involved in reframing signals of punishment into anticipations of reward. Treating with anti-glutamatergic agents would be a mechanism-based approach, albeit targeting the co-risk factor, thus far with no empirical evidence to support it. Of particular interest is ketamine, with some evidence of efficacy in OCD proper, perhaps by dampening obsessive ruminations, but no data for anankastic personality disorder. Interestingly, ketamine, a NMDA antagonist, leads to acute increased release of glutamate, enhancement of fast acting AMPA glutamate receptors, and a cascade of postsynaptic events, most notably the activation of brain-derived neurotrophic factor (BDNF) and mammalian target of rapamycin (mTOR), a protein kinase involved in a number of vital cellular functions. Low BDNF and low expression of mTOR have also been implicated in depression. Among other mechanisms, synaptogenesis in response to mTOR activation is hypothesized to contribute to ketamine's antidepressant effects (Aan Het Rot et al. 2012). It is thus plausible that anti-OCD effects of ketamine derive from antianxiety effects that are independent from NMDA antagonism. In turn such reduced anxiety may benefit compulsive perfectionism in OCD and anankastic personality disorder.

**Mechanism-Based Pharmacotherapy of Trait Vulnerability Associated with Low RD**

As a temperament trait, RD involves a heritable bias in favour of the maintenance of a behaviour in response to cues of social reward. Its facets are attachment, sociability, sentimentality, and dependence on other (Cloninger et al. 1993). Given the listed facets, RD is more accurately labelled as “social RD,” although the trait may facilitate other types of reward conditioning, including addictions. Social motivation for acceptance or approval by a group has played a major role in the evolutionary success of humans. The link with survival puts social reward in the class of unlearned primary positive reinforcers, together with food, mating, and water. Propensity for social interaction is observable early in development, even in newborns (Bicks et al. 2015).

Social RD is not a major trait vulnerability for personality disorder, excepting the schizoid variant where it is excessively low. In fact, high social RD is a protective trait against personality pathology, as sociable children tend to express and evoke positive social emotions. Social RD acts as a modifying trait insofar as it is low in antisocial, and paranoid types (“detached, callous”) and high in narcissistic, histrionic, and dependent variants (“highly engaged, needy”). (Table 3). Importantly, for most pharmacological treatments of low social RD, at least some baseline social interest and/or trust are needed to increase the chance of success. As social trust is invariably impaired in personality disorder, the first prosocial steps are usually achieved in psychotherapy which, here, sets the stage for pharmacotherapy, and not the other way around as is the case with other variants.

**The complex psychobiology of social RD**

In higher animals, social signals and cues are processed by two main subcortical networks: (i) the social behavior network (Newman 1999) and (ii) the mesolimbic dopaminergic reward pathway. In humans, subcortical social instincts are processed by the PFCx and other components of the general network of cognition. These cortical processes send top-down feedback to contextualize the subcortical perception of social information, together known as “social cognition”. With this in mind, social RD is the most complex temperament trait, with cortical and subcortical mechanisms co-participating in valuation, assessment, and execution of social behaviors. Hence its mechanism-based pharmacotherapy is complex as well.

**Subcortical oxytocin, dopamine, and endogenous opioids in human social behaviors**

Early activation of the oxytocin-vasopressin system, e.g., through breastfeeding and skin-to-skin contact, modulates development of brain networks involved in adult social behaviors. The organization of these networks is modified epigenetically through social experiences and circulating levels of modulating hormones and neuropeptides. The networks are open to lifelong plasticity based on actual social experience. Quantitative levels of oxytocin and vasopressin are heritable, which likely accounts for individual variability in social information processing and bonding affinity, both of which are pivotal for social RD. High scores on social RD are positively correlated (r=0.425) and explain about 17% of the variance in plasma oxytocin levels, making this neuropeptide the most powerful predictor of the trait (Bell et al. 2006). Endogenous oxytocin might be a fundamental mechanism in the development of social RD and perhaps a specific postfactum medication for socially detached individuals.
There is accumulating data to suggest that in healthy, trustful individuals, intranasal oxytocin increases prosocial while attenuating negative social behaviors. Oxytocin seems to reduce uncertainty and increase trust about the predictive value of a social stimulus, thereby inhibiting social defensiveness, promoting social cognition, and facilitating approach in social interactions (Heinrichs & Domes 2008). These prosocial effects vary depending on gender, context (whether people are familiar or unfamiliar), perceived threat, childhood experiences, and personality features, the most influential of which is quality of attachment: oxytocin increases sociability in securely attached, mature individuals, and worsens distrust in those with disorganized attachment. For the latter, at least some trust building in psychotherapy is likely to be needed to set the stage for oxytocin’s prosocial effects. Therefore, instead of being the universal “love” molecule, intranasal oxytocin may simply increase the salience of perceived social cues, whether positive or negative.

Pursuit of signals of reward is an integral part of NS, illustrating its close functional relatedness with social RD. Functional neuroimaging shows that social rewards (such as positive feedback from others) are processed by the same mesolimbic reward network as nonsocial reward and addictions (Caldwell & Albers 2016). Also, the mesolimbic DA system is thought to be involved in reward prediction in social interactions (aka “incentive salience”). The prediction is based on summation of an individual’s past subjective experience of reward delivery, making this mechanism especially relevant in personality disorder. A hypoactive mesolimbic DA system, postulated to underlie low NS, may contribute to a reduced incentive value of interpersonal contacts, resulting in social detachment/anhedonia. An increased D2 receptor density in the putamen, suggestive of decreased dopaminergic influences, was shown to correlate with the Karolinska scale of detachment, one of 15 scales in the test (Farde et al. 1997). Oxytocin, vasopressin, and their receptors are abundant in the DA reward network as well, perhaps providing the connection between the specific elements that are processed by each. For example, oxytocin augments the NAc response to viewing the faces of romantic partners as well as to the reciprocated cooperation from human partners (Caldwell & Albers 2016).

Pleasure is a component of most rewards, including social reward. Pleasure hotspots in the reward system that mediate the hedonic or the “liking” aspect of activity are subcortical (NAc shell, ventral pallidum) and cortical (ventromedial PFCx, insular cortex) and in addition to DA also involve mediation by endogenous opioids (Berridge & Kringelbach 2015), as discussed later.

**Low oxytocin in social distrust of personality disorder**

Low plasma oxytocin was found in callous-unemotional postpubescent boys, both key features of heritably low social RD. Low oxytocin reflected either an epigenetically silenced (hypermethylated) promoter for the oxytocin receptor gene (OXTR; Dadds et al. 2014a) or with OXTR mutation (Dadds et al. 2014b). These findings implicate the oxytocin receptor, either silenced epigenetically or mutant, in a host of structural and functional aberrances relevant to sociability, from differential amygdala reactivity (Dadds et al. 2014b) to aberrant brain networks of social cognition (Hammock 2015), the last of which becomes especially compromised in ambiguous situations per some reports.

Oxytocin has been studied mostly in borderline personality, which is characterized by rejection sensitivity and negative relationship representations. Overall, intranasal oxytocin tends to further impair an already low trust and cooperation in these individuals (Bartz et al. 2011a). The underlying negative bias tends to undermine the sociability effects of oxytocin pharmacotherapy, as discussed. Women with borderline personality have been shown to have low baseline oxytocin levels, which tend to correlate with the experience of trauma in childhood (Bertsch et al. 2013), pointing to epigenetic silencing of the OXTR in the context of early adversity. The silencing is believed to be maintained by low-circulating oxytocin and chronic interpersonal insecurities via negative feedback. Thus, early dysregulation of the oxytocin-vasopressin system could be instrumental in connecting childhood adversity and/or innate excessive affects, aberrant development of neural networks involved in social cognition and empathy, social distrust and hypersensitivity toward negative emotions, impaired social cognition, and poor top-down cortical control of behavior, all of which perpetuate each other in a vicious cycle.

**The Locus Coeruleus-Noradrenergic (LC-NA) system in social RD**

The LC-NA system appears to play a role in processing information concerning attentional orientation toward and incentive value of conditioned rewarding stimulus but not reward itself (Bouret & Sara 2004). The strength of social RD is inversely related to NE levels as measured by urine metabolites (Garvey et al. 1996). Reboxetine, a NE reuptake inhibitor - NRI, was found to selectively enhance detection of emotionally arousing stimuli, likely due to an increase in extracellular NE occurring in phasic bursts (De Martino et al. 2008). It appears that a constitutionally high baseline NE tone leaves little room for phasic bursts in NE.

---

20 We use the term “fragmented” instead of “borderline” personality, as the former highlights intrapsychic fragmentation as the essence of personality disorder. The term “borderline” has been overused and misused to the point of confusion, e.g., as a separate category of personality disorder, as a level of personality organization, and for a group of patients on the border between neuroses and psychoses.
which are needed to accentuate the salient stimuli of social reward, and this saturation flattens the need for social approval, perhaps manifested as low social RD ("detached and disinterested introverts"). Conversely, individuals high in social RD, and presumably low baseline NE activity, may be more reward vigilant, and thus prone to more readily engage in socially rewarding interactions ("people-oriented extraverts") as well as in other rewarding activities, including work and addictions ("task- or object-oriented extroverts").

### Pharmacological modulation of trait vulnerability associated with low social RD

With the above in mind, oxytocinergic, noradrenergic, and dopaminergic medications fit the concept of mechanism-based strategies for low social RD. None of the below cited drugs has been studied systematically or approved for social disinterest or distrust, but the clinician may use them off label at his or her discretion and based on risk/benefit analysis. Preparatory in-depth psychotherapy is indicated in most cases.

#### Intranasal oxytocin for low social RD and social distrust

At least some cases of schizoid personality disorder with low social RD are characterized by low plasma oxytocin (Bell et al. 2006). These individuals manifest low social RD (social disinterest) but not necessarily a pronounced interpersonal distrust or interpersonal conflicts, in contrast to all other variants. In this setting, oxytocin may be beneficial even without preparatory psychotherapy (Bakermans-Kranenburg & Van Ijzendoorn 2014). For other variants of personality disorder, in the context of high distrust yet a high need for approving social relations, the benefits of oxytocin are questionable, as discussed. Here, prosocial effects of oxytocin could be promoted by trust-building psychotherapy as noted. In short, oxytocin is a potentially promising "sociability" molecule but is bound to carry large individual differences in responses. The differences could reflect genetic and/or acquired variations in the oxytocin-vasopressin system, age, gender, context, pre-existing personality features, or specific environmental influences during sensitive periods in mental development. An interactionist approach in which stable individual differences and contextual factors are taken into account could maximize the desired effects of exogenous oxytocin in healthy and clinical populations (Bartz et al. 2011b).

#### Noradrenergic drugs may focus attention to positive social signals

As discussed, noradrenergic mechanisms seem to modulate the strength of social signals, an important mechanism in low social RD. As with oxytocin, at least some baseline prosociality appears to be needed to achieve the positive noradrenergic effect on social attention. Medications that enhance noradrenergic transmission include SNRIs (duloxetine, milnacipran, levo-milnacipran are most noradrenergic in the group), NRIs (reboxetine, atomoxetine), and bupropion (a weak NE and DA reuptake inhibitor). To our knowledge, noradrenergic medications have not been systematically studied in socially detached individuals with low social RD, with the exception of individuals with autism, an unrelated disorder which happens to be characterized by low social RD. With mixed success, atomoxetine improved social withdrawal in developmentally disabled young individuals, including those with autism who were characterized, among other features, by low social RD and social disinterest (Aman et al. 2014).

#### Dopaminergic drugs may increase the incentive salience of social contacts

Social detachment and disinterest characterize the schizoid, paranoid, and anankastic variants of personality disorder. There is some, albeit scarce, evidence that pharmacological enhancement of central DA activity may enhance the incentive salience and perhaps the hedonic aspects of social interactions. Psychostimulants and bupropion are thought to increase extracellular DA and NE via reuptake inhibition and possibly also via presynaptic release (bupropion). Figuratively, medications that enhance central DA and NE neurotransmission, may increase sociability by decreasing the averseness of social insecurities (a noise-reducing, incentivizing dopaminergic effect) while at the same time increasing the strength of positive social signals (a potentiating noradrenergic effect). Second generation partial DA agonists may also be of interest here, especially cariprazine with a strong affinity for the D3 receptor, abundant in the NAc. In humans, 3,4-methylenedioxy-methamphetamine (MDMA or "ecstasy") was shown to increase sociability and empathy by diminishing the amygdala responses to threatening stimuli (angry faces), and activating the NAc responses to rewarding social signals, such as happy faces (Bedi et al. 2009). Ecstasy, a complex stimulant, is mistakenly thought to be a psychedelic compound but is actually a substituted amphetamine with amphetamine-like monoaminergic targets. The ecstasy’s empathogenic effects have been in part attributed to its propensity to increase plasma levels of oxytocin in humans, per some reports in mesolimbic regions, thereby connecting the dopaminergic and oxytocinergic mechanisms of sociability, as discussed.

#### Opioids may increase hedonism and reduce frustration of social contacts

In addition to dopaminergic mechanisms, the hedonic aspect of social reward appears to be in part mediated by the endogenous opiate system, as discussed. In healthy individuals, high social RD, but not other temperament traits, was associated with...
increased density of opiate receptors in the NAc and ventral striatum, both major hubs in the brain’s reward system (Schreckenberger et al. 2008), which presumably mediates warm attachments to other people. Buprenorphine, a partial mu opiate receptor agonist with relatively lower risk of addiction may benefit individuals with chronic social detachment by activating the VTA-NAc pathway, although it has not been studied for this indication. Also, through its kappa receptor antagonism, buprenorphine may decrease dysphoria and improve frustration tolerance in social interactions.

**Future Directions: Psychedelics as a potential top-down mechanism for homeostatic recalibration of personality?**

The field of psychiatry has of late shown a renewed interest in psychedelic compounds such as LSD, psilocybin, mescaline among others. The majority of these compounds act as agonists at the 5-HT2A receptor, which is thought to mediate their unique effect on the brain (Glennon et al. 1984). The 5-HT2A receptors are densely expressed on the dendritic spines of layer 5 pyramidal neurons, which serve as the primary source of cortical output to the remainder of the brain (Spruston 2008, Weber & Andrade 2010). There is much we still do not know about the pharmacodynamics of psychedelics. The emerging view is that they can have a profound and sometimes lasting effects on the brain’s functional connectivity networks. These networks remain spatially segregated but temporally integrated even at rest, and these spontaneously formed patterns (“resting state” or “intrinsic” connectivity) are thought to be essential in maintaining a normally functioning brain and mind (Mitra et al. 2015). Given the high concentration of 5HT2A receptors in layer 5, psychedelics likely increase global entropy by disrupting the direction in which infra low frequency activity which proceeds from layer 5 and is thought to be crucial in maintaining the functional organization of brain networks, as it travels across the cortical layers (Mitra et al. 2018). LSD has been found to increase the global connectivity of normally distinct neural networks, which was found to correlate with subjective reports of ego dissolution following LSD administration (Tagliazucchi et al. 2016). Another study found decreased connectivity between the parahippocampus and retrosplenial cortex to correlate with the characteristic phenomenon of ego dissolution (Carhart-Harris et al. 2016). Decreased resting state functional connectivity within the default mode network (DMN) is another consistent finding with the administration of psychedelics, although this may merely reflect an increase in global brain entropy (Carhart-Harris et al. 2016). Another study found links between the above described global entropic effects and lasting personality changes in the measured trait openness (Lebedev et al. 2016). It is not uncommon for patients with personality disorder to exhibit a hypofrontal phenotype. Interestingly, the density of functional connectivity actually increases in the frontoparietal region with LSD administration (Tagliazucchi et al. 2016). This represents a potentially important difference from other altered states of consciousness such as procedural sedation that do just the opposite, which may explain the unique subjective experience following network reintegration in patients who have had exposure to psychedelic compounds. Additionally, it is interesting to note that psychedelics appear to have an effect first on the prefrontal cortical networks, which only later go on to influence subcortical affects downstream. This represents a unique top-down mechanism, akin to the effects of psychotherapy, which contrasts with the monoaminergic bottom-up approach commonly encountered in mechanism-based pharmacotherapy.

Whether a therapeutic benefit is to be found for psychedelics with respect to the treatment of personality disorder remains to be seen. In any case we would recommend these compounds be studied within a controlled clinical setting as acutely increasing global entropy of network connectivity is apt to elicit a variety of reactions depending on the individual and the setting, ranging anywhere from complete self-transcendence to unmitigated terror. Pre-treatment psychotherapy is thus mandated. Acute, judicious use of pimavanserin (an inverse agonist at the 5HT2A receptor), and benzodiazepines (which inhibit neuronal activity via GABA) may be useful to terminate the experience in the event of the latter.

**Limitations**

The proposed mechanism-based pharmacotherapy is an emergent and heuristic concept that is mounted on the idea put forth by mechanisms that are fundamentally empirical in nature. While they have significant underpinnings in neuroscience, the idea that pharmacological agents may be used to activate homeostasis and re-calibrate specific and complex phenotypic traits has not been previously studied within controlled clinical settings and therefore a priori evidence is sparse.

---

21 The term means “soul manifesting” in English, as these compounds cause an altered state of consciousness and sensory changes, produce experiences of emotional communion, oneness, relatedness, and emotional openness (empathogens or entactogens) and are believed to promote development of the unused potentials of the human mind (entheogens).

22 Subjects often report that their sense of being a self or ‘I’ distinct from the rest of the world has diminished or altogether dissolved.
Acknowledgements:

Some of this material was originally published in “The Fragmented Personality: An Integrative, Dynamic, and Personalized Approach to Personality Disorder” by Svrakic D & Divac-Jovanovic M, and has been used by permission of Oxford University Press https://global.oup.com/academic/product/the-fragmented-personality-9780198846117?q=the%20fragmented%20Personality&lang=en&cc=us

Conflict of interest: None to declare.

Contribution of individual authors:

All authors contributed equally to the organization and the write up of this paper

References


24. Farde L, Gustavsson JP, Jonsson E: D2 dopamine receptors and personality traits. Nature 1997; 385:590. doi:10.1038/385590a0


46. Miller EK, Lundqvist M, Bastos AM: Working Memory 2.0, Neuron 2018; 100:463-75


Correspondence:
Ravikumar Chockalingam, MD
Veterans Affairs Medical Center
Saint Louis, Missouri, USA
E-mail: chockalr30@gmail.com