Changing character: A narrative review of personality change in psychotherapies for personality disorder

Article *in* Psychotherapy Research · January 2018

DOI: 10.1080/10503307.2018.1425930

CITATIONS
2

READS
176

2 authors:

Jack R Keefe
Weill Cornell Medical College
44 PUBLICATIONS **537** CITATIONS

SEE PROFILE

Robert J. DeRubeis
University of Pennsylvania
199 PUBLICATIONS **14,133** CITATIONS

SEE PROFILE

Psychotherapy Research

ISSN: 1050-3307 (Print) 1468-4381 (Online) Journal homepage: http://www.tandfonline.com/loi/tpsr20

Changing character: A narrative review of personality change in psychotherapies for personality disorder

John R. Keefe & Robert J. DeRubeis

To cite this article: John R. Keefe & Robert J. DeRubeis (2018): Changing character: A narrative review of personality change in psychotherapies for personality disorder, Psychotherapy Research, DOI: 10.1080/10503307.2018.1425930

To link to this article: https://doi.org/10.1080/10503307.2018.1425930

Published online: 19 Jan 2018.

Psychotherapy Research, 2018
https://doi.org/10.1080/10503307.2018.1425930

METHOD PAPER

Changing character: A narrative review of personality change in psychotherapies for personality disorder

JOHN R. KEEFE & ROBERT J. DERUBEIS

Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA

(Received 15 April 2017; revised 31 December 2017; accepted 2 January 2018)

Abstract

Objective: Personality disorder (PD) is a negative prognostic indicator for treatment, and absolute improvements in functioning among these patients are often modest. This may be because personality features that give rise to dysfunction in PD are not targeted optimally during most treatments. **Method:** Attachment, mentalization, core beliefs, and personality organization/defense use were identified as personality constructs that have been pursued in treatment studies and that are proposed to underlie PD. **Results:** All constructs correlate with psychiatric symptoms, PD diagnosis, and functioning. Defense mechanisms and core beliefs further distinguish specific PDs, whereas personality organization separates more versus less severe PDs. Evidence from treatment and naturalistic studies indicate that maturation of defense mechanisms temporally precedes improvements in symptoms and functioning. Changes in attachment and mentalization correlate with some outcomes, but mediation of improvement has not been established. In psychodynamic therapy, transference interpretations may promote amelioration of personality dysfunction. With the exception of attachment, the experimental literature is lacking that could explicate the mechanisms by which these personality constructs maintain psychosocial dysfunction. **Conclusions:** Future research should aim to identify changes in these mechanisms that mediate positive outcomes in PD, as well as the specific therapeutic procedures that best promote positive change in PD.

Keywords: Long-term psychotherapy; outcome research; personality disorders; process research; attachment; psychoanalytic/psychodynamic therapy; cognitive behavior therapy

Clinical or methodological significance of this article: Patients with personality disorder (PD) suffer from high psychosocial impairment and are treatment resistant, which may be instantiated on the level of personality by dysfunctions in attachment, mentalization, core beliefs, and personality organization/defense use. Specialized treatments for PD appear to promote improvement in most of these constructs. Improvement in in-session use of and insight into defense mechanisms signal later functional gains among PD patients, and may be especially important to monitor and target in-treatment.

A personality disorder (PD) is an enduring disturbance in how an individual experiences and interprets themselves, others, and the world, understood as an exaggeration, rigidity, or breakdown of normal personality processes that consequently promotes dysfunctional behaviors (APA, 2013; Hopwood, Zimmermann, Pincus, & Krueger, 2015). A core phenomenological component of PD is that features that others see as difficult are often understood by the patient as essential to their way of being, rather than reflecting a personality feature that they come into treatment wanting to change (Oltmanns, Rodrigues, Weinstein, & Gleason, 2014). For example, patients with a narcissistic PD rarely enter treatment desiring to work on their heightened grandiosity; rather, they might enter treatment suffering from a major failure, or to complain about the inability of others to treat them with the admiration and respect they deserve (Caligor, Levy, & Yeomans, 2015). Accordingly, patients with PD have idiosyncratic

Correspondence concerning this article should be addressed to John R. Keefe, Department of Psychology, University of Pennsylvania, 425 S. University Avenue, Levin Building, 4th Floor, Philadelphia, PA 19104, USA. Email: jack.keefe@gmail.com and jkeefe@sas.upenn.edu

and unsatisfying functioning within interpersonal relationships, often through utilizing self-defeating or non-optimal strategies for managing their emotional responses and getting their needs met (Sadikaj, Moskowitz, Russell, Zuroff, & Paris, 2013).

PDs are highly prevalent psychiatric conditions, with estimates of 6.1–9.1% in United States and international samples (Huang et al., 2009; Lenzenweger, Lane, Loranger, & Kessler, 2007) and a 35–50% comorbidity rate with mood, anxiety, and eating disorders (Friborg, Martinsen, et al., 2014; Friborg, Martinussen, et al., 2014; Friborg, Martinussen, Kaiser, Øvergård, & Rosenvinge, 2013). PD is often considered a marker of clinical severity and treatment resistance for these disorders (Ansell et al., 2011; Skodol, Geier, Grant, & Hasin, 2014). For example, in the treatment of major depressive disorder (MDD), PD comorbidity is a negative prognostic indicator, with such patients being less likely to attain clinically significant depression change with psychotherapy, medication, or their combination (Newton-Howes et al., 2014), and a faster time to relapse after remission (Grilo et al., 2010).

Furthermore, over and above comorbid "Axis-I" conditions such as MDD, diagnosis with a PD is a predictor of diminished functioning and impaired, dissatisfying interpersonal relationships (Ansell, Sanislow, McGlashan, & Grilo, 2007; Hill et al., 2008; Skodol et al., 2005). Moreover, improvements in functioning longitudinally are often modest, especially for more severe PD. Only about 40% of BPD patients in a treatment-seeking sample were found to experience at least one 8-year period of functional recovery within a 16-year window (Zanarini, Frankenburg, Reich, & Fitzmaurice, 2012). The Collaborative Longitudinal Personality Disorder Study concluded that while BPD patients can, over the course of 10 years, fall under the DSM-IV diagnostic threshold for the disorder and maintain a symptomatic remission, they retain a "severe and persistent impairment in social functioning" even after remitting (Gunderson et al., 2011). Functional recovery rates for other PDs have been reported to be in between those found in BPD patients and non-PD patients with MDD (Gunderson et al., 2011; Zanarini et al., 2012).

Personality Change and Improvements in Functioning

An assumption behind many tested treatments for PD is that change in *personality* is necessary for a successful treatment (Beck, Freeman, & Davis, 2004; Yeomans, Clarkin, & Kernberg, 2015), both to normalize the patient's functioning in social, cognitive, and affective domains, and to prevent the return of acute symptomatology by ameliorating personality-related vulnerabilities. For example, a patient with MDD and a comorbid obsessive-compulsive PD (OCPD) whose mood symptoms remit may be expected to experience improvements in romantic relationship functioning (e.g., the stress of having a depressed partner is no longer present), yet will still encounter significant interpersonal difficulties related to their persistent OCPD character traits (e.g., tendency to become angry with their romantic partner when they "violate" minor household rules). These interpersonal difficulties, therefore, remain a source of vulnerability for future episodes of mood disturbance. Indeed, patients who attain depression remission who do not simultaneously fall under the diagnostic threshold for their PD are also more likely to experience substantially worse functioning across a 2-year follow-up (Markowitz et al., 2007).

PD diagnoses comprise acute behavioral symptoms in addition to longstanding personality patterns that are less prone to change (Morey & Hopwood, 2013), such that, for example, change in latent personality traits has been observed to be a better predictor of change in psychosocial functioning than change in DSM-IV PD criteria (Wright et al., 2015). Thus, in the prevailing diagnostic systems, remission from PD does not necessarily indicate that personality vulnerabilities have been substantially reduced. These pathological personality patterns can then generate adverse experiences even after acute symptoms remit, increasing the likelihood of symptom exacerbation. It is unclear to what extent evidence-based treatments for PD address the personality problems that sustain functional impairments.

Theories of Personality Disorder

A comprehensive, empirically grounded theory of personality pathology and its remediation through a linked psychotherapy will have, at minimum, the following features (cf., Kazdin, 2007): (1) A clear definition of the nature of the pathology (the mechanism) that accounts for the behaviors and experiences that define a given PD; (2) Delineation of how particular therapeutic procedures and in-session processes may promote changes in these mechanisms; and (3) Conceptualization of how, when change in the mechanism occurs, improvements in functioning accrue to the patient.

Such a theory would be supported by the existence of reliable methods of assessing these mechanisms that can identify patients with PD, experimental evidence that demonstrates how the proposed

mechanisms instantiate pathology and dysfunction, and mediational work relating particular intervention use to change in the mechanism and then to improvements in symptoms and functioning. We will use this framework to guide an explication of four common approaches to the understanding and treatment of personality pathology, which have been examined in randomized trials for PD: adult attachment representations (Gillath, Karantzas, & Fraley, 2016), mentalization/reflective functioning (RF) (Fonagy, Gergely, Jurist, & Target, 2002), core beliefs (Beck et al., 2004), and psychodynamic personality organization and its attendant psychological defenses (Vaillant, Bond, & Vaillant, 1986; Yeomans et al., 2015).[1] Generally, improvements in these constructs are conceived of as change mechanisms promoting shifts in cognition, experiencing, and behavior that are observable outside of the therapy room, rather than in-session client change processes that are in response to therapeutic interventions (Doss, 2004).

Four Personality Mechanisms of Pathology and Change in PD

Attachment Representations

Attachment refers to a characteristic style of relating to others, such as parents, children, or romantic partners, in intimate care-giving and care-receiving situations (Bowlby, 1973). The attachment system is thought to be activated under stress, and particular styles describe how and why an individual proceeds to act once the system is activated. Attachment styles are thought to develop first from the character of the affective bond between child and caregiver, which forms an individual's initial "working model" of relationships and of his or her self-role vis-à-vis those relationships (Bowlby, 1973). Throughout development, an individual may have life experiences that are assimilated into their current working model (e.g., betrayal enhancing extant attachment insecurity), as well as experiences that require the working model to accommodate those experiences (e.g., relationship safety and intimacy ameliorating attachment insecurity) (Fraley, Roisman, Booth-LaForce, Owen, & Holland, 2013).

Attachment is thought to be composed of "unconscious" or implicit components of working models of relationships and conscious or cognitive components representing appraisals of relationships and one's role in them (Roisman et al., 2007; Yaseen, Zhang, Muran, Winston, & Galynker, 2016). These are, respectively, assessed by interviews assessing attachment states of mind (e.g., the Adult Attachment Interview; AAI; (George, Kaplan, & Main, 1985)) and self-report measures (e.g., the Experiences in Close Relationships questionnaire; (Fraley, Waller, & Brennan, 2000). Although in this review we will discuss attachment from both perspectives—as results relating attachment to PD diagnosis specifically have been relatively similar across modes of assessment—we note that there exists a substantive, active debate in the attachment literature as to the motivational, cognitive, and behavioral correlates of each method of assessing attachment styles (Roisman et al., 2007). A commonality of secure attachment in both perspectives is the capacity to predict the ability of an individual to use attachment figures for support, and to be relied upon by others as an attachment figure (Shaver, Belsky, & Brennan, 2000). Interestingly, experimental investigations of attachment utilizing self-report assessments have nevertheless produced results bespeaking contributions of not only conscious attributions but also implicit/unconscious processes to self-reported attachment (Shaver & Mikulincer, 2002). By contrast, relatively less experimental work based on attachment representations has involved observer-rated attachment styles (Gillath et al., 2016), although investigations involving the AAI and relationships sometimes find that it relates more to observer-rated as compared to self-reported relationship behaviors (Jacobvitz, Curran, & Moller, 2002).

The most frequently employed systems for classifying attachment styles make the following distinctions: (a) Secure attachment (realistic expectations of being able to rely on attachment figures during times of trouble; comfort with normative dependency; sense of self as basically competent) versus insecure attachment (significant disruption of any or all of these capacities); (b) Among insecure attachments, anxious/preoccupied attachment (overwhelmed by consciously accessible anxiety concerning access to attachment figures; believes self to be unable to function without attachment figures) versus avoidant/dismissive attachment (tendency to distance the self from attachment figures in an effort to avoid or escape the experience of painful emotions that are aroused by attachment needs); and (c) On the AAI, classifiable attachment (relatively consistent patterns of attachment) versus two special categorization codes—unclassifiable (e.g., unusual variability and instability of expression; no organized attachment response) or trauma-unresolved attachment (e.g., conceptions of past trauma overwhelm attachment narratives).

Dysfunctional intimate relationships are endemic among PD patients, and insecure attachment has been viewed as a baseline dysfunction interacting with other distinctive features of PDs to promote maladjustment (Levy, Johnson, Clouthier, Scala, & Temes, 2015; Meyer & Pilkonis, 2005). Insecure

self-reported and observer-rated attachment is much more common among patients with PD, relative to non-clinical controls and patients with mental disorder diagnoses but without PD (Bouchard et al., 2008; Brennan & Shaver, 1998; Crawford et al., 2006; Crawford et al., 2007; MacDonald, Berlow, & Thomas, 2013; Olssøn & Dahl, 2014). Olssøn and Dahl (2014) estimated the rate of self-reported insecure attachment to be 90% among patients with PD, compared to 35% in community controls. Patients with BPD, in particular, have been found to have especially high rates of insecure attachment (Choi-Kain, Fitzmaurice, Zanarini, Laverdiere, & Gunderson, 2009), in addition to increased rates of unclassifiable or trauma-unresolved adult attachment, which are observed infrequently in the general population (Bakermans-Kranenburg & van IJzendoorn, 2009). It should be noted, however, that whereas a very high percentage of PD patients exhibit an insecure attachment style, most individuals with an insecure attachment style do not qualify for a PD (Bakermans-Kranenburg & van IJzendoorn, 2009).

There has been less research regarding specific attachment styles as they relate to specific PDs other than BPD. As described in a review by Levy (2005), self-reported anxious attachment was associated with histrionic, dependent, and avoidant PDs, avoidant attachment was associated with paranoid, narcissistic, antisocial, and schizoid PDs, and dual elevations in avoidant and anxious attachment were common in schizotypal, paranoid, avoidant, obsessive-compulsive, narcissistic, and borderline PDs (Levy, 2005). However, a meta-analytic review of this literature would be a useful contribution to understanding non-replications of specific attachment-PD associations between studies. To this general point, significant differences in observer-rated attachment have been observed between the styles of BPD patients with versus without a comorbid narcissistic PD (NPD), such that most "pure" BPD patients exhibit primary attachment anxiety and trauma-unresolved attachment, whereas BPD patients who additionally have NPD exhibit more attachment avoidance and unclassifiable attachment (Diamond et al., 2014).

More than any other personality construct impaired in PD, findings in the attachment literature provide experimental empirical grounding for understanding the mechanisms of seemingly strange, dysfunctional interpersonal patterns prevalent in PD (Shaver & Mikulincer, 2005). For example, avoidantly attached individuals engage in "attachment-deactivating" relational strategies (Gillath et al., 2016). Attachment theory holds that this is in part because avoidant individuals believe that there is something dangerous about intimacy (e.g., that they will not get what they want; that they will surely be hurt; that they will get lost in an intimate relationship) and are sensitive to attachment stressors (Bartholomew, 1990).

Indeed, more avoidantly attached adults as assessed by both self- and observer-report have been observed to exhibit increased psychophysiological electrodermal activity in response to questions about potential abandonment or rejection in past close relationships (Diamond, Hicks, & Otter-Henderson, 2006; Dozier & Kobak, 1992). Curiously, physiological data suggest markedly heightened inflammatory reactivity among avoidantly attached individuals, despite their self-reported disavowal of attachment stresses (Gouin et al., 2009). Notably, attachment avoidance attenuates the correlation between self-reported and physiologically recorded arousal in response to an interpersonal stressor, suggesting a disconnect between what is reported (or, perhaps, even reportable) by the individual, and some facet of their "objective" arousal (Diamond et al., 2006). Cognitively, more avoidantly attached individuals (via self-report) show diminished working memory capacity for attachment-related stimuli (Edelstein, 2006), memory biases toward recalling negative information about past relationships (Haggerty, Siefert, & Weinberger, 2010), difficulties suppressing separation-related cognitions under stress but greater suppression under normal conditions (Mikulincer, Dolev, & Shaver, 2004), and pre-emptive and post-emptive biases against processing of emotional information (Andriopoulos & Kafetsios, 2015; Edelstein & Gillath, 2008), among other altered processing (Dykas & Cassidy, 2011).

These findings are consistent with the theory that avoidantly attached individuals *do* care about attachment relationships, but may experience considerable distress and inhibition in situations in which they may need to rely on attachment figures for social and emotional support. This conflict represents a maladaptive admixture of motivational states (i.e., to both attach and avoid attachment), potentiating seemingly incoherent behavioral strategies that have been observed experimentally, such as to reject an exciting potential romantic partner toward whom one feels *too much* interest (Spielmann, Maxwell, Macdonald, & Baratta, 2013). These avoidant behaviors resemble interpersonal problems and patterns characteristic of particular PD styles—for example, the propensity of individuals with OCPD to understand themselves as unemotional, without relational needs, and counter-dependent within relationships.

This clinically and empirically rich conceptualization of avoidant attachment could, in turn, inform goals in psychotherapy with avoidantly attached

individuals with PD regardless of the specific PD diagnosis. These goals could include helping these individuals experience, recognize, and accept their own affects and strivings related to attachment and dependency more directly, and to help them be intimate toward others without feeling unsafe or mentally disorganized (Mikulincer, Shaver, & Berant, 2013).

Attachment change in PD treatment. Despite considerable theoretical interest in attachment as a construct linking clinically meaningful concepts of self and other relationships to an empirically rich social-developmental literature (Blatt & Levy, 2003; Levy et al., 2015), it has been investigated only a handful of times in PD treatment. Among BPD patients, AAI-rated, categorical attachment classifications have been observed to improve significantly in transference-focused psychotherapy (TFP) (Yeomans et al., 2015) as compared to dialectical-behavioral therapy (DBT) (Linehan, 1993) and dynamic-supportive therapy (Levy et al., 2006) (converted $d = 0.82$). TFP is theorized to repair dysfunctional attachment by helping patients become aware of their working models of relationships, to help them reconcile and integrate incoherent parts of these working models, and to build capacities to compensate for their attachment weaknesses (e.g., to encourage patients with attachment anxiety to come up with solutions with the therapist's guidance but not direction; to help avoidant patients express their normative needs for closeness). This finding has been replicated in a sample of patients receiving TFP as compared to enhanced treatment as usual (TAU) by community experts, with significantly more TFP patients moving both from AAI-rated insecure to secure attachments styles (converted $d = 0.71$), and from unclassifiable to classifiable styles (converted $d = 0.50$) (Buchheim et al., 2017; Buchheim, Horz, Rentrop, Doering, & Fischer-Kern, 2012). In both studies, attachment did not show significant change from over treatment among the therapies that were compared to TFP (i.e., among DBT, dynamic-supportive therapy, or enhanced TAU).

Improvements in self-reported (ECR) attachment avoidance specifically have also been observed to occur during mentalization-based therapy (MBT) (Bateman & Fonagy, 2006) for adolescents with self-harm (the majority of whom had a BPD diagnosis), but not in TAU ($d = 0.42$) (Rossouw & Fonagy, 2012). In addition, the advantage of MBT over TAU in improving self-harm was significantly attenuated when controlling for the superiority of MBT in improving attachment avoidance (r of avoidance and self-harm change = 0.55) (Rossouw & Fonagy, 2012), suggesting a possible mediational role for attachment change, to be investigated in future studies.

Summary—attachment. Change in attachment status is in and of itself a notable outcome, given the maladaptive behaviors and ways of being that are characteristic of insecure attachment and the established relationship between insecure attachment and observable interpersonal dysfunction. Improvements in attachment status may reflect a meaningful reorganization of a patient's relationships to self and other, which could lead to stable, self-perpetuating gains in functioning. As such, future research on attachment in PD treatment should seek to relate change in attachment to improvements in PD symptomatology, interpersonal problems, and psychosocial functioning, over the course of treatment and during follow-up. Furthermore, a synthesis between self-report and observer-rated attachment perspectives would be desirable.

Mentalization/Reflective Functioning

RF—or one's degree of mentalization—has been defined by developmental and personality researchers as a capacity to understand and interpret one's own and others' behaviors as expressions of mental states, including feelings, thoughts, fantasies, beliefs, and desires (Fonagy et al., 2002). This capacity is hypothesized to be especially impaired when an individual is acutely stressed, and especially when the attachment system is activated. Individuals with poor mentalization report experiences that suggest a limited ability to distinguish between their own emotions, thoughts, and intentions and those of other, and/or impoverished conceptions of their own and others' mental states (Fonagy, Target, Steele, & Steele, 1998). Impairments in mentalization include both so-called *hypo*-mentalization (lack of interest and capacity in mentalization) and *hyper*-mentalization (over-focus on others' motives).

Most commonly, RF scores are calculated from the AAI (George et al., 1985), in which individuals are asked to describe and elaborate upon various aspects and scenarios involving their childhood relationships with their parents. Many questions inquire as to why their parents acted in certain ways, and what one's own responses and reactions were to those actions. Global RF scores have been found to be assessed with good interrater reliability and with stability across points of measurement, and confirmatory factor analysis suggests that global RF indicators may most parsimoniously reflect a single latent construct (Taubner et al., 2013). Normatively, an RF score of 5 or above indicates

that an individual has a mature, coherent model of the mind that they have the capacity to apply even in emotionally charged or attachment-activating situations (Fonagy et al., 1998). There is some evidence that RF as typically measured taps more into the ability to identify intentions and higher-order mental states rather than alexithymia *per se* or an ability to engage in affective empathy, and may be more akin to conceptions of cognitive empathy (Falkenström et al., 2014).

RF as general marker of PD. Across nearly all studies investigating RF and PD, RF has been found to distinguish PD from other types of psychopathology (Katznelson, 2014). RF disruptions may thus be a general dysfunction common to all PD, a mechanism underlying the marked interpersonal difficulties and relational distress that PD patients experience.

In the first clinical investigation of RF, patients with PD were found to have lower RF scores than a non-psychiatric control group, with BPD patients having the lowest mean RF (Fonagy et al., 1996). Since this original investigation, deficits in RF have been further characterized among patients with BPD, with and without comorbid NPD (Diamond et al., 2014; Fischer-Kern et al., 2015). RF, above indexes of attachment style and general psychiatric symptomatology, has been observed to be positively related to PD diagnoses and severity (Bouchard et al., 2008; Fischer-Kern et al., 2010; Nazzaro et al., 2017), and among PD patients to predict symptomatology, interpersonal problems, and psychosocial functioning (Antonsen, Johansen, Rø, Kvarstein, & Wilberg, 2016). There is evidence suggesting that impaired RF may be associated with Cluster A and B PD traits but not Cluster C traits (Nazzaro et al., 2017).

RF change in treatments for PD. In contrast with the primary research literature on RF, which has employed standardized observer ratings, MBT treatment research has used recently developed self-report RF questionnaires (Fonagy et al., 2016; Sharp et al., 2009). Patients are asked to report on their experiences of difficulties in mentalization, with questions such as "I sometimes do things without really knowing why" and "I frequently feel that my mind is empty." In an RCT of MBT for self-harming adolescents, 76% of whom had a BPD diagnosis, RF improved significantly more in MBT compared to TAU (Rossouw & Fonagy, 2012) (d = 0.38). Moreover, the advantage of MBT over TAU in preventing self-harm in this trial was accounted for by improvements in RF (r = 0.48), such that differences between the two treatments were significantly attenuated when improvements in RF were controlled for, although the temporal precedence of increases in RF, relative to reductions in self-harm, was not established (Rossouw & Fonagy, 2012). In a well-powered (n = 175) open trial of psychodynamic hospitalization incorporating BPD techniques, improvements in symptoms across assessment points strongly tracked concurrent self-reported decreases in mentalizing impairments (r = 0.89) (De Meulemeester, Vansteelandt, Luyten, & Lowyck, 2017). A study of adolescent BPD inpatients receiving MBT reported significant pre-to-post increases in self-reported RF (Bo et al., 2017).

RF has also been studied, with AAI-based assessments of RF, in trials of TFP for BPD. RF is theorized to improve in TFP in large part due to the therapist's work helping the patient understand and distinguish between representations of the self and therapist as they emerge in therapy, and the affects and mental states linked to these representations. RF has been observed to improve more over the course of a year for BPD patients receiving TFP as compared to either DBT or a manualized dynamic-supportive therapy (Levy et al., 2006) (d vs. DBT = 0.56, d vs. supportive = 0.85). In fact, in neither of the two comparison treatments did RF reliably change during treatment. In another trial, patients randomized to TFP exhibited significantly greater improvements in RF as compared to patients in an enhanced TAU from expert community providers (d = 0.45), among whom RF did not significantly increase (Fischer-Kern et al., 2015). Notably, improvements in RF were found to covary with improvements in personality organization (r = 0.31). However, its association with functioning *per se* was not examined.

Summary—RF. Impaired RF appears to be a marker of personality and interpersonal dysfunction that improves in specialized treatments for PD. Improvements in RF among patients with a PD may indicate an increasing capacity to accurately and actively assess the motives of one's self and others, and ergo an increasing capacity for interpersonal functioning—a hypothesis that should be tested in future studies. Furthermore, it would be valuable to better understand how RF scores relate to performance differences in social cognition (e.g., cognitive empathy; attention to facial cues) and interactions in social situations.

Cognitive–affective Schemas

Cognitive frameworks for understanding PDs primarily explain PD as a function of the presence and

penetrance of maladaptive cognitive–affective schemas or core beliefs (Beck et al., 2004). Schemas have also received attention from psychodynamic clinical theorists, who in different language emphasize the relevance of the *content* of unconscious schemata, such as different self-object dyads as conceptualized in TFP (Yeomans et al., 2015).

Beck's schema theory. From the perspective of Beck's cognitive theory, patients with particular personality disorders hold characteristic core beliefs that underlay those specific PDs (Beck et al., 2004). For example, a patient with a dependent PD might strongly believe that "I am only safe if there is someone to take care of me" and that "I am needy and weak," and a patient with a paranoid PD might strongly believe that "I can't trust anyone" and that "others will try to use and manipulate me if I don't watch out." In this way, Beck's theory is foremost a theory of how the *content* of mental life creates personality pathology. While all individuals are purported to hold particular core beliefs that form personality styles, patients with PD are held to distinguish themselves from less pathological individuals by the rigidity and intensity with which they hold these beliefs, rendering them as having a PD. These rigid, intense beliefs are posed to strongly bias information processing about the self and others in a manner that leads to the manifestation of particular PD symptoms (e.g., perfectionistic beliefs leading to a sense that nothing one or others does is "good enough" unless immaculately performed, as can be seen in OCPD). While Beck's therapeutic formulation of CT for PD is specifically designed to challenge the core beliefs proposed to underlay various PDs (Beck et al., 2004), many contemporary CT researchers would suggest that multiple procedures —not simply CT-specific procedures—could lead to meaningful revisions to a patient's core beliefs (Lorenzo-Luaces, German, & DeRubeis, 2014). Generally, more strongly endorsing PD-related core beliefs exhibit higher concurrent depression and anxiety, in addition to diminished self-esteem, lower self-reported social support, and lower psychosocial functioning (Butler, Beck, & Cohen, 2007).

On the whole, there is support for the proposition that differences in core beliefs between PD patients correlate with their specific PD diagnosis in a fashion corresponding to the content of those beliefs. The Personality Beliefs Questionnaire (PBQ) (Beck & Beck, 1991; Butler et al., 2007), a self-report measure of putative core beliefs, has been found to distinguish between patients with versus those without a PD (Bhar, Beck, & Butler, 2012; Fournier, DeRubeis, & Beck, 2012). The PBQ has also been found to partially distinguish between specific PDs on the basis of endorsed beliefs, such that beliefs hypothesized to characterize particular PDs both form specific replicated factors and are more present among patients with that PD diagnosis compared to other PDs (Beck & Beck, 1991; Butler et al., 2007; Fournier et al., 2012). The correspondence between hypothesized belief endorsements and the factor structures of the measure is relatively strong and replicated for SCID-II diagnosed avoidant/dependent PD, OCPD, paranoid PD, schizoid PD, and NPD. By contrast, BPD does not appear to be well-distinguished by belief endorsement in the PBQ (Fournier et al., 2012). In addition, theorized avoidant and dependent PD-specific beliefs appear instead to represent a shared factor entailing beliefs about dependency and fear of interpersonal rejection and its consequences (Fournier et al., 2012). These findings converge with those obtained in a study finding specific correlations between five types of self-reported PD on the Millon Clinical Multiaxial Inventory and their corresponding beliefs on the PBQ (avoidant, dependent, passive-aggressive, schizoid, and borderline) (Jones, Burrell-Hodgson, & Tate, 2007).

Core belief change in treatments for PD. Where the literature is significantly lacking in terms of Beck's conception of schema are data on how these beliefs change during treatment for PD, and whether change in these beliefs drive symptom change or predict further gains or resistance to relapse after the end of treatment. To our knowledge, there is only a single, small study using the PBQ in the context of PD treatment, in which avoidant PD beliefs were observed to change over the course of manualized CT for AVDP, as compared to a wait-list and an unmanualized psychodynamic therapy that was intended as a control condition (Emmelkamp et al., 2006) (d vs. wait-list = 1.02, d vs. dynamic control = 0.50). The relationship between change in avoidant beliefs and core PD symptomatology was not reported. In a process study comparing CT for MDD patients with and without PD, early therapist focus on core beliefs was associated with subsequent depression and pre-to-post PD symptom change among patients with PD, but not in those without PD (Keefe, Webb, & DeRubeis, 2016) (sr = 0.46 for depression; sr = 0.29 for PD symptoms), although core beliefs were not directly measured in this trial. Overall, while PDs differ in terms of the core beliefs prototypically associated with each disorder, we have relatively minimal empirical data on the relevance of these beliefs to treatment course.

Core beliefs—summary. Self-report measures of maladaptive schema content are hypothesized to

reflect interpretive frameworks or mental content that play important roles in the maladaptive behaviors and painful experiences of patients with PDs. The PDQ, a measure that draws on Beck's conceptualization of core beliefs can, in addition, distinguish among patients with different PDs. However, changes on the PBQ during PD treatment have yet to be robustly linked to changes in symptoms or functioning. Moreover, specific contentions attached to the theory, for example, that activation of core beliefs should prime or bias toward/against particular interpretations, have yet to be tested experimentally.

Personality Organization and Defense Mechanisms

Whereas cognitive theories of PD emphasize the maladaptive content and rigid use of schemas, psychodynamic theories focus more so on personality organization and the psychological defenses one performs to maintain one's sense of self and of one's relationship to others. Personality organization refers not to the presence of particular kinds of mental content (i.e., schema) *per se*, but rather the characteristic ways in which an individual reconciles mental contents. Strategies that maintain psychological homeostasis (i.e., maintaining a personality organization) and deal with anxiety-provoking thoughts, feelings, and desires are known as psychological defenses (Freud, 1937/1966; Vaillant, 1994).

Personality organization. Kernberg's (1984) structural theory of personality organization proposes that non-psychotic personality organization is arrayed on a spectrum, from healthy to neurotic to borderline. All individuals are assumed to employ a plethora of "self-object" dyads, each of which comprises a representation of the other and its attitudes toward the self (e.g., the other as strong and protective), the self and its positionality toward the other (e.g., the self as safe and deserving of protection), and an affect that binds the two (e.g., love). He posits that neurotic individuals, because their representations of themselves and others are relatively consistent, use "mature" mechanisms of defense more frequently to deal with conflict relating to those representations of self and other. The problems that arise at this level are typified by *ambivalence* and *conflict* between opposing desires or representations. For example, the person may wish to express anger but avoids doing so because it would signify that one is acting like one's abandoning father.

Individuals who operate on a borderline level of functioning are said to evidence instead an inability to integrate pieces of a complex image of themselves and of others. This limits the patient's ability to employ more mature or neurotic defenses, leaving only the more "primitive" or borderline defense mechanisms, which lack nuance and thereby overemphasize one mental experience or representation over another at a given time. Arguably, the most well-known example of this is the tendency to act toward the same person at different times as though the person is all-good or all-bad (i.e., splitting). Personality problems at this level are conceived of as emerging from *splitting* or *dissociation between* opposing representations. Overall, personality organization is arrayed on a continuum of mental integration (Kernberg, 1984; Stern et al., 2010). In Kernberg's conception, emotional and behavioral patterns associated with a given PD emerge from an interaction of the patient's level of personality organization and the particular representations of self and others held by the patient.

The Structured Interview of Personality Organization (STIPO) (Stern et al., 2010) and self-report Inventory of Personality Organization (Lenzenweger, Clarkin, Kernberg, & Foelsch, 2001) have been used to assess personality organization per Kernberg's conceptualization, and have exhibited good psychometric properties (Doering et al., 2013; Ellison & Levy, 2012; Lenzenweger et al., 2001; Preti et al., 2015; Stern et al., 2010). They each yield scores for three factors thought to reflect level of personality organization: use of primitive defense mechanisms (e.g., splitting), the quality and depth of relationships (e.g., extent to which the patient describes people in their lives in a stereotyped manner), and instability in conceptions of self and others, goals, and behavior. For example, an item from the STIPO regarding the instability of representations is "Would you say that you feel different about yourself, about who you are as a person, across different situations or depending on who you're with? Would you say that you come across to the same person in very different ways at different times?"

Over and above Axis-I pathology, lower levels of personality organization have been found to predict higher PD traits and likelihood of being diagnosed with a PD, decrements in psychosocial functioning, and decreased relationship satisfaction by both members of a couple (Doering et al., 2013; Ellison & Levy, 2012; Fischer-Kern et al., 2010; Fischer-Kern et al., 2011; Lenzenweger et al., 2001; Stern et al., 2010; Verreault, Sabourin, Lussier, Normandin, & Clarkin, 2013). Individuals with lower levels of personality organization have also been found to have higher rates of suicide attempts and non-suicidal self-injury (Baus et al., 2014). In addition, patients scoring in the "borderline" range of the STIPO tend to qualify for primary Cluster A and B PD. By

contrast, patients with Cluster C diagnoses tend to score in the "neurotic" personality band, whereas patients with no PD score better still (Doering et al., 2013; Stern et al., 2010).

Lastly, a recent large-scale bifactor modeling study found that BPD criteria—most of which bespeak instabilities in conceptions of self and others—may form a common "backbone" of PD severity shared by different PDs, from the more functional (e.g., OCPD) to the less functional (e.g., paranoid PD) (Sharp et al., 2015). At the same time, several PD-specific factors were identified, which varied alongside the general PD factor. These findings suggest that difficulties in integrating representations of the self and others are a core feature of all PDs, which in turn have unique difficulties and features. The authors interpreted this bifactor model as being supportive of Kernberg's theory of personality organization, in that personality organization is proposed also to reflect a continuum of mental integration that combines with other, more specific personality features to produce what we understand to be a given PD. From the perspective that personality organization may be a common dimension of all PD, is unclear to what extent personality organization is more similar to a trait, in that it is predominantly descriptive and reflects the combined operations of other psychological mechanisms (e.g., borderline defense mechanisms), or to what extent it can be conceived of as a mechanism.

Personality organization and change in psychotherapy. Kernberg's theory of personality organization informs the theoretical basis of TFP, which has been tested in studies of treatment for BPD (Clarkin, Levy, Lenzenweger, & Kernberg, 2007; Doering et al., 2010) but is hypothesized to be efficacious for all PDs existing at the "borderline" level of organization (e.g., NPD, paranoid PD). In the only such study measuring personality organization as an outcome (Doering et al., 2010), borderline PD patients receiving a year of TFP experienced significantly greater improvements in personality organization (as measured by the STIPO) compared to patients treated by community experts ($d = 0.65$). Kernberg hypothesizes that improvements in personality organization form the most meaningful signal of eventual health and functioning in PD patients, and that these improvements would be more likely to occur in psychodynamic treatment frames, but these questions have yet to be tested empirically.

Defense mechanisms. A psychological defense is a particular way of coping with a distressing, uncomfortable, or undesired mental state. The individual need not be aware of the nature of a given mechanism, or the link between the event or mental state and the application of a defense. Defenses can be arrayed in a hierarchy of more to less mature and adaptive defenses (Bond, Gardner, Christian, & Sigal, 1983; Vaillant, 1994; Vaillant et al., 1986). To illustrate this: suppose that a man is told by his romantic partner that they wish to have dinner alone with a former lover, and that this understandably makes the man feel angry. Mature defenses would be engaged if the man could assert his feelings to his partner directly for their consideration (i.e., self-assertion), or consciously postpone dealing with his distress while at work (i.e., suppression). Alternatively, the engagement of a less mature, "neurotic" defense mechanism would lead to the *avoidance* of experiencing the anger, through repression, transformation, or displacement of the feeling. Such defenses could include the man discussing the inciting situation with a friend without an angry tone (i.e., isolation of affect), or expressing frustration with that same friend rather than with his partner (i.e., displacement). Application of an even less mature, "borderline" or "primitive" defense would result in distortions that affect internal representations of the self and others and/or external reality to eliminate anxiety about expressing the anger, or to express it in an exaggerated, pathological form. "Borderline" defenses the man could use in response to this situation include treating his partner's behaviors as if they were meant to hurt him. This then could lead to the man's sense that the anger is completely justifiable as the partner's intentions are purely loathsome, and perhaps have been all along (i.e., splitting). Alternatively, the man might retaliate impulsively by contacting and scheduling a drink with a former love interest with the hope of provoking his partner to feel the jealousy and anger he is currently undergoing (i.e., acting out and projective identification).

Use of a particular defense in a given situation is hypothesized to emerge from an interaction of the person's dispositional personality patterns, the ways in which he or she has learned to use defenses in the past, the person's general capacity to hold in mind uncomfortable states, and the situational context. Individuals who exhibit a more mature level of defense have been shown in numerous longitudinal studies to exhibit better mental health and functioning outcomes over time (Bond, 2004).

For individuals with a PD, defenses may often serve as the symptoms in and of themselves (e.g., splitting in BPD). Defense usage may also provide barriers to self-improvement by diverting a patient's awareness away from aspects of stress and conflict

in their lives; avoidance of affect and of stressful topics renders it difficult to alter one's behaviors and ways of experiencing the world. Defensive functioning, as assessed with both observer and self-report measures, has been found repeatedly to be lower among patients with PD, compared to nonpsychiatric controls and psychiatric patients without PD (Perry & Bond, 2005).

Importantly, psychodynamic theories of PD posit that particular patterns of personality pathology result from the use of specific psychological defenses. For example, patients with NPD and BPD exhibit elevations in defenses of splitting representations of the self (i.e., thinking of the self as all-good or all-bad), and of projecting an air of omnipotence to the external world (Perry, Presniak, & Olson, 2013). However, BPD patients are discriminated from NPD patients by their more frequent use of splitting of representations of others, projective identification, dissociation, and acting out (Kramer, de Roten, Perry, & Despland, 2013; Perry et al., 2013; Zanarini, Weingeroff, & Frankenburg, 2009). By contrast, patients with NPD, relative to those with BPD, more often devalue the therapist and their social relations, and habitually rationalize their own behavior both in the therapeutic relationship and in their outside social contacts (Perry et al., 2013). Distinctive patterns of defense use have also been identified for antisocial and schizotypal PDs (Perry et al., 2013). By comparison, specific defense use has been little examined for Cluster C PDs. Although qualitative clinical literatures make strong predictions, such as that patients with OCPD engage in increased intellectualization and isolation of affect (Summers & Barber, 2010), these proposals have yet to be tested empirically.

In addition to distinguishing the types of PD, defense mechanisms in PD help explain complex or seemingly paradoxical features of PD. For example, among BPD patients, relatively higher narcissistic defense use—idealization of the self, devaluation of others, and attempts to control the therapist—correlates with *lower* self-reported symptomatology (Kramer et al., 2013). Yet narcissistic defenses are associated with PD diagnosis rather than health (Kramer et al., 2013; Perry et al., 2013; Zanarini et al., 2009) and narcissistic pathology with higher severity of both symptoms and interpersonal problems, in addition to poor functioning and treatment dropout (Ellison, Levy, Cain, Ansell, & Pincus, 2013; Ogrodniczuk, Piper, Joyce, Steinberg, & Duggal, 2009). This pattern of findings is consistent with an account that narcissistic defense use enhances self-esteem temporarily, providing an "illusion of health" that renders the patient less interested in engaging in the work of therapy (Kramer et al., 2013).

Optimal match between defenses and therapeutic procedures. Given the identification of specific defense mechanisms with individual PDs, a therapeutic procedure that targets a PD-concordant defensive style might be expected to have more impact than a procedure that addresses a defensive style characteristic of a different PD. TFP for BPD instantiates this principle in that it focuses on how the patient's use of "borderline" defenses distorts their experience of the therapist and of themselves in relation to the therapist. A goal of TFP is to render conscious the patient's tendency to experience the therapist and themselves in binary, shifting black-and-white terms (i.e., engage in splitting) as well as to engage in efforts to provoke the therapist to experience particular feelings or act out particular roles (i.e., projective identification). Increased maturity of defense use in session is hypothesized to generalize to interpersonal relationships outside of therapy. Transference interpretation—or moment-to-moment analysis of the way in which the patient's impressions of and feelings toward the therapist come from their own schemas and personality dynamics—is thought to be key.

In terms of promoting improved attachment and mentalization, TFP has shown an advantage in clinical trials over psychodynamic-supportive therapy, DBT, and TAU. These findings have been interpreted as supporting the unique value of the technique of transference interpretation for this patient group (Buchheim et al., 2012; Fischer-Kern et al., 2015; Levy et al., 2006). More direct evidence comes from a randomized clinical trial of psychodynamic psychotherapy (PDT) for patients with mixed psychiatric disorders, in which PD patients randomized to receive transference interpretations in their therapy evidenced superior gains in functioning during treatment and in the long-term follow-up, compared to PD patients who did not receive transference interpretations (Hoglend, Dahl, Hersoug, Lorentzen, & Perry, 2011; Høglend et al., 2008).

Change in defenses for mixed PD. Perry and Bond (2012) found, in a predominantly PD sample (76% with a full diagnosis; all with elevated PD traits), that patients who experienced greater improvement in observer-rated defensive functioning over the course of 2.5 years of long-term PDT experienced superior further improvements in symptoms ($r = 0.58$) and functioning ($r = 0.60$) over the next 2.5 years of follow-up. Improvements in defensive functioning reflected diminished use of defenses theorized to be less mature and adaptive (e.g., projective identification; $d = -0.67$) as well as increases in the use of adaptive defenses (e.g., humor; $d = 0.80$) (Perry & Bond, 2012). In secondary analyses of data from

the Høglend et al. (2008) trial, the advantage of transference interpretations for patients with low quality of object relations (predominantly PD patients) through a 3-year follow-up ($d = 0.40$ at termination, 0.52 at 1-year, 0.32 at 3-years) was mediated by the technique promoting increases in the patient's awareness of their use of defenses (Johansson et al., 2010). In total, 60% of the intervention effect was explained by differential improvements in insight.

Change in defenses for cluster-C PD. In a randomized comparison of affect-focused PDT for PD versus Beck's CT for PD (Svartberg, Stiles, & Seltzer, 2004), Cluster-C patients receiving PDT who were rated by observers as exhibiting an increased awareness of how they used defenses at a late stage of treatment experienced more improvements in general symptoms ($r = 0.18$) and self-reported interpersonal problems ($r = 0.18$) from treatment termination to a 2-year follow-up (Kallestad et al., 2010). Furthermore, in both the PDT and CT groups, decreases in the observed use of these defense mechanisms in-session from intake to a late session predicted greater improvements in general symptoms from pretreatment to a 2-year follow-up ($r = 0.34$) (Johansen, Krebs, Svartberg, Stiles, & Holen, 2011).

Change in defenses for BPD. Finally, in a naturalistic study of treatment-seeking patients with BPD, improvements in self-report measures of four defense mechanisms (humor, projection, help-seeking complaining, and acting out) were linked to subsequent changes in interpersonal functioning across a 16-year follow-up period. Changes in these indices at one-time point predicted faster time to attain interpersonal and occupational recovery, with hazard ratios ranging from 0.82–0.64 for immature defense change and 1.18 for humor (Zanarini, Frankenburg, & Fitzmaurice, 2013).

Summary—change in defenses. In the five studies reviewed, increased use of mature defenses or increased awareness of the use of defenses predicted improvements on measures of symptoms and functioning. In some cases, the benefits of these changes appeared to extend beyond the end of treatment. This is consistent with the notion that changes in defense use signal reconfigurations in the way patients relate to troubling feelings, thoughts, and desires. Relative to the other reviewed constructs, change in defense mechanisms has been more clearly linked to subsequent improvements in functioning, and to the use of a particular therapeutic procedure (interpretations of transference). However, while there have been efforts to link or interpret phenomena and findings in other fields of psychology as indicating the use of particular psychological defenses (Baumeister, Dale, & Sommer, 1998), there is little work being done examining manifestations of defense use outside of the therapy room.

Discussion

We have reviewed four personality constructs that have been linked, through theoretical and empirical work, to the maladaptive functioning observed in persons with PD: attachment representations, mentalization/RF, core beliefs, and personality organization/defense use. So far, to our knowledge only studies that have focused on defense mechanisms have been designed to identify the temporal precedence of personality and symptom change, and indeed the findings indicate that improvements in defense mechanisms precede and act as a mediator of gains in functioning. Treatment-related improvements in attachment and mentalization have been demonstrated to correlate, pre- to post-treatment, with improvements, but tests of mediation have not been performed for these constructs. Such personality changes should also be investigated for their ability to predict the maintenance, enhancement, or loss of gains after treatment, as has been demonstrated for defense mechanisms.

The practical value of an understanding of the relation of these constructs to improvement in functioning would be the translation of these findings into recommendations for therapists as to how they should tailor their treatment to best target the dysfunctional personality of each specific patient. In PDT, use of transference interpretations has been demonstrated to promote improvements in attachment, mentalization, and awareness of defense mechanisms among patients with a PD, and to be superior to PDT without transference interpretations among these patients. However, not all studies have suggested a positive relationship between transference work and improvements in PD (Ryum, Stiles, Svartberg, & McCullough, 2010), and further research is needed to determine for what PDs and treatment contexts transference work is especially valuable (e.g., Cluster B). Moreover, randomized findings regarding techniques other than transference interpretations in PDT are necessary. For example, although within CBT greater symptomatic improvement in patients with comorbid depression and PD was found to be associated with a more frequent use of core belief-focused techniques (Keefe et al., 2016), there has been no test of the hypothesis that an emphasis on such techniques leads to better outcomes for PD patients.

Alone among the constructs we reviewed, attachment has a robust experimental literature that elucidates when, why, and how disrupted attachment creates dysfunction in an individual's life. The experimental attachment literature covers the gamut from developmental antecedents of attachment, to motivational/desire states characteristic of particular styles, to implicit/unconscious representations and cognitive/attentional processes activated in attachment situations, to emotion-regulation strategies under attachment stress, to observed normative and problem-solving behaviors in romantic relationships (Gillath et al., 2016). However, many of these findings relate to self-reported attachment, and more experimental studies regarding observer-reported attachment would be welcome. There are few-to-no experiments examining how individual differences in mentalization, core beliefs/schemas, and psychodynamic personality organized/defense mechanisms translate into differences in motivation/desire, cognition, and behavior. We believe that the study of these other personality constructs would be buoyed substantially by an appropriate program of study modeled after the attachment literature, to understand how these constructs instantiate observed deficits in functioning. This work may also provide insights as to how these processes can be altered through psychotherapy.

Limitations

A major theory of personality not covered in this review is the Five-Factor Model (FFM) of personality traits, most prominently specified by the DSM-V alternative diagnostic system for PD (APA, 2013) and the Personality Inventory for the DSM-5 (Al-Dajani, Gralnick, & Bagby, 2016). This is because, although there is ample work indicating that different PDs have distinguishing FFM patterns (Hopwood, Thomas, Markon, Wright, & Krueger, 2012), this literature is not associated with any empirically supported therapy for the treatment of PD. Nor, to our knowledge, has change in FFM-based traits ever been examined or even intended to be targeted in a PD treatment trial. In addition, the FFM perspective is agnostic as to how traits come to be constituted and expressed, or what may change a trait. We would argue that although the FFM perspective highlights domains of dysfunctions that are located at the extremes of normal personality dimensions and that these may characterize particular problem areas, the mechanisms whereby these dysfunctions arise and recede has not received attention from FFM theorists or researchers. Therefore, the particular clinical routes toward improving a given PD or moving FFM scores away from the extremes have not been specified. If the theory behind the FFM is to inform and be informed by treatment literatures, it will be necessary to develop an understanding of the convergence between FFM-derived measures of PD and personality constructs that have been of greater interest to treatment researchers, such as between core beliefs and correlated FFM-derived traits (Hopwood, Schade, Krueger, Wright, & Markon, 2013).

Another unresolved issue concerns the best means of assessing the relevant personality constructs, especially in the context of treatment trials. Self-report measures are relatively easy to administer and therefore lend themselves to repeated administrations. However, ratings from trained observer may provide more valid information. Mentalization, which describes a propensity and capacity to engage in a particular process, may be less amenable to assessment via self-report, due to its focus on performance (i.e., of mentalization) and the fact that poor mentalization might by its nature be unavailable to introspection. The associations between the self-report and observer-rated measures of RF have not, to our knowledge, been reported (Fonagy et al., 2016).

Some of the PD constructs that are assessed via self-report constructs, such as core beliefs in Beck's PD theory, re-capitulate DSM criteria. For example, one of the DSM criteria for OCPD, "shows perfectionism that interferes with task completion," is the objective statement of the PBQ-self-report that "It is important to do a perfect job on everything" (Fournier et al., 2012). Beck might argue that such connections between core beliefs and PD symptoms are meaningful insofar as it is the mental content (the core belief) that gives rise to the symptom. Importantly, it is the contention of Beck that changes in the consciously accessible mental states associated with PD symptomatology should lead to improvements in the relevant symptomatology. Ultimately, high-quality evidence on the temporal relationships between belief change and symptom/functioning improvements, and demonstrating that core belief activation biases cognition, would disentangle the extent to which core beliefs underlay PD.

Future Directions

Few empirical tests of the temporal precedence of personality change *vis á vis* symptom reduction or functional improvements have included multiple measurements of both types of changes over the course of treatment (for an exception, see Hoffart, Versland, & Sexton, 2002). Observations of repeated within-person relationships between personality change and symptomatic/functional improvements, controlling for the autocorrelations in each type of measurement, would constitute superior evidence

for a mechanistic role for personality change on the level of the individual (Curran & Bauer, 2011). Attaining an empirical understanding of how these personality processes play out for individual patients (i.e., a person's unique personality dynamics) may enable matching of therapeutic procedures to specific personality features contributing to dysfunction (Fisher & Boswell, 2016).

Another major question facing the field of PD research is how to integrate understanding of these constructs. Each has been used to describe and explain specific aspects of the experience of PD, but it is unclear to what extent they represent uniquely informative perspectives on PD, or instead are different ways of describing similar core processes. Some writers have theorized specific inter-relations and hierarchies among the conceptualizations. For example, defenses help maintain a particular organization (Yeomans et al., 2015). Another example is that mentalization is assumed to be impaired when the attachment system is activated (Fonagy et al., 2002), and has been observed to be more impaired among BPD patients with increasingly worse personality organization (Fischer-Kern et al., 2010). However, it is also possible that constructs from differing traditions may both overlap conceptually and interact meaningfully with one another. Attachment, when assessed by self-report, might to an extent reflect specific, special types of core beliefs. On the other hand, observer-rated attachment, derived from structured relational narratives, may also be codeable for defense mechanism use and personality organization during the narratives, which may be found to relate to particular attachment styles.

It is lamentably rare for two or more of the PD constructs we have reviewed to be assessed simultaneously in the same study. Integration of what appear to be complementary conceptualizations (e.g., personality organization and defense mechanisms) as well as more divergent conceptions of personality is crucial to advancing knowledge about the psychopathology and treatment of PD. In treatment studies, assessments of multiple personality constructs at several different points during therapy could help discriminate between these different causal hypotheses of personality change.

Summary and Conclusion

At present, we know relatively little about what needs to change to disrupt pathological personality, and how clinicians can best facilitate such change. However, research has identified several personality constructs that predict PD diagnoses and could explain a host of behaviors and experiences common to PD, such as conflicts and disrupted strivings regarding intimacy (attachment), difficulties in understanding the motivations of others in context of one's own motivations (mentalization), inflexible and idiosyncratic ways of interpreting experiences (core beliefs), inconsistency and lack of nuance/integration in conceptions of self and other (personality organization), and avoidance of core emotional and interpersonal problems in day-to-day life (defense mechanisms). These constructs work to clarify PD categories as not merely reflecting exaggerations of normative personality, but as constituting several, potentially overlapping domains of inter- and intrapersonal dysfunction. Encouragingly, many of these personality constructs have been observed to improve over the course of psychotherapies specialized to treat PD, but not in TAU or control therapies. This suggests that these specialized therapies may be on the right track toward ameliorating pathological personality.

The next generation of PD trials should work to identify these constructs as they change in response to specific psychotherapies, and how they covary over time with symptomatic and functional improvements and with each other. Future research into the processes that promote and the effects that result from personality change may provide better clues for clinicians who wish to assist their patients with PD to achieve "a life worth living" (Linehan, 1993) or, as has been attributed to Freud, a meaningful capacity "to work and to love" (Erikson, 1950).

Acknowledgments

The authors wish to thank Rachel Schwartz and Drs. Kelly Allred, Dianne Chambless, Sara Jaffee, Sigal Zilcha-Mano, Suzanne van Bronswijk, who provided helpful comments on different versions of this manuscript.

Note

[1] Due to changing space limitations of the journal requiring substantive cuts to our review, we removed a section discussing Jeffrey Young's conception of schema and the associated evidence behind it. We encourage interested readers to consult reviews by Taylor, Bee, & Haddock (2017) and by Jacob & Arntz (2013).

ORCID

John R. Keefe http://orcid.org/0000-0001-7250-8960

References

Al-Dajani, N., Gralnick, T. M., & Bagby, R. M. (2016). A psychometric review of the personality inventory for DSM-5 (PID-5): current status and future directions. *Journal of Personality Assessment*, *98*(1), 62–81. doi:10.1080/00223891.2015.1107572

Andriopoulos, P., & Kafetsios, K. (2015). Avoidant attachment and the processing of emotion information: Selective attention or cognitive avoidance? *Journal of Relationships Research*, *6*, e6 (12 pages). doi:10.1017/jrr.2015.2

Ansell, E. B., Pinto, A., Edelen, M. O., Markowitz, J. C., Sanislow, C. A., Yen, S., … Grilo, C. M. (2011). The association of personality disorders with the prospective 7-year course of anxiety disorders. *Psychological Medicine*, *41*(5), 1019–1028. doi:10.1017/s0033291710001777

Ansell, E. B., Sanislow, C. A., McGlashan, T. H., & Grilo, C. M. (2007). Psychosocial impairment and treatment utilization by patients with borderline personality disorder, other personality disorders, mood and anxiety disorders, and a healthy comparison group. *Comprehensive Psychiatry*, *48*(4), 329–336. doi:10.1016/j.comppsych.2007.02.001

Antonsen, B. T., Johansen, M. S., Rø, F. G., Kvarstein, E. H., & Wilberg, T. (2016). Is reflective functioning associated with clinical symptoms and long-term course in patients with personality disorders? *Comprehensive Psychiatry*, *64*, 46–58. doi:http://doi.org/10.1016/j.comppsych.2015.05.016

APA. (2013). *Diagnostic and statistical manual* (5th ed.). DSM-5. Washington, DC: American Psychiatric Publishing.

Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2009). The first 10,000 adult attachment interviews: Distributions of adult attachment representations in clinical and non-clinical groups. *Attachment & Human Development*, *11*(3), 223–263. doi:10.1080/14616730902814762

Bartholomew, K. (1990). Avoidance of intimacy: An attachment perspective. *Journal of Social and Personal Relationships*, *7*(2), 147–178. doi:10.1177/0265407590072001

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

Baumeister, R. F., Dale, K., & Sommer, K. L. (1998). Freudian defense mechanisms and empirical findings in modern social psychology: Reaction formation, projection, displacement, undoing, isolation, sublimation, and denial. *Journal of Personality*, *66*(6), 1081–1124. doi:10.1111/1467-6494.00043

Baus, N., Fischer-Kern, M., Naderer, A., Klein, J., Doering, S., Pastner, B., … Kapusta, N. D. (2014). Personality organization in borderline patients with a history of suicide attempts. *Psychiatry Research*, *218*(1–2), 129–133. doi:http://doi.org/10.1016/j.psychres.2014.03.048

Beck, A. T., & Beck, J. S. (1991). *The personality belief questionnaire*. Unpublished assessment instrument. Bala Cynwyd, PA: The Beck Institute for Cognitive Therapy and Research.

Beck, A. T., Freeman, A., & Davis, D. D. (2004). *Cognitive therapy of personality disorders* (2nd ed.). New York, NY: Guilford.

Bhar, S. S., Beck, A. T., & Butler, A. C. (2012). Beliefs and personality disorders: An overview of the personality beliefs questionnaire. *Journal of Clinical Psychology*, *68*, 88–100.

Blatt, S. J., & Levy, K. N. (2003). Attachment theory, psychoanalysis, personality development, and psychopathology. *Psychoanalytic Inquiry*, *23*(1), 102–150. doi:10.1080/07351692309349028

Bo, S., Sharp, C., Beck, E., Pedersen, J., Gondan, M., & Simonsen, E. (2017). First empirical evaluation of outcomes for mentalization-based group therapy for adolescents with BPD. *Personality Disorders*, 8, 396–401. doi:10.1037/per0000210

Bond, M. (2004). Empirical studies of defense style: Relationships with psychopathology and change. *Harvard Review of Psychiatry*, *12*(5), 263–278. doi:10.1080/10673220490886167

Bond, M., Gardner, S. T., Christian, J., & Sigal, J. J. (1983). Empirical study of self-rated defense styles. *Archives of General Psychiatry*, *40*(3), 333–338.

Bouchard, M.-A., Target, M., Lecours, S., Fonagy, P., Tremblay, L.-M., Schachter, A., & Stein, H. (2008). Mentalization in adult attachment narratives: Reflective functioning, mental states, and affect elaboration compared. *Psychoanalytic Psychology*, *25*(1), 47–66. doi:10.1037/0736-9735.25.1.47

Bowlby, J. (1973). *Attachment and loss: Separation* (Vol. 2). New York, NY: Basic Books.

Brennan, K. A., & Shaver, P. R. (1998). Attachment styles and personality disorders: Their connections to each other and to parental divorce, parental death, and perceptions of parental caregiving. *Journal of Personality*, *66*(5), 835–878.

Buchheim, A., Horz, S., Rentrop, M., Doering, S., & Fischer-Kern, M. (2012). *Attachment status before and after one year of transference focused psychotherapy (TFP) versus therapy as usual (TAU) in patients with borderline personality disorder*. Paper presented at the 2nd international congress on borderline personality disorders and allied disorders, Amsterdam, The Netherlands.

Buchheim, A., Hörz-Sagstetter, S., Doering, S., Rentrop, M., Schuster, P., Buchheim, P., … Fischer-Kern, M. (2017). Change of unresolved attachment in borderline personality disorder: RCT study of transference-focused psychotherapy. *Psychotherapy and Psychosomatics*, *86*(5), 314–316.

Butler, A. C., Beck, A. T., & Cohen, L. H. (2007). The personality belief questionnaire-short form: Development and preliminary findings. *Cognitive Therapy and Research*, *31*, 357–370.

Caligor, E., Levy, K. N., & Yeomans, F. E. (2015). Narcissistic personality disorder: Diagnostic and clinical challenges. *American Journal of Psychiatry*, *172*(5), 415–422. doi:10.1176/appi.ajp.2014.14060723

Choi-Kain, L. W., Fitzmaurice, G. M., Zanarini, M. C., Laverdiere, O., & Gunderson, J. G. (2009). The relationship between self-reported attachment styles, interpersonal dysfunction, and borderline personality disorder. *Journal of Nervous and Mental Disease*, *197*(11), 816–821. doi:10.1097/NMD.0b013e3181bea56e

Clarkin, J. F., Levy, K. N., Lenzenweger, M. F., & Kernberg, O. F. (2007). Evaluating three treatments for borderline personality disorder: A multiwave study. *American Journal of Psychiatry*, *164*(6), 922–928. doi:10.1176/ajp.2007.164.6.922

Crawford, T. N., John Livesley, W., Jang, K. L., Shaver, P. R., Cohen, P., & Ganiban, J. (2007). Insecure attachment and personality disorder: A twin study of adults. *European Journal of Personality*, *21*(2), 191–208. doi:10.1002/per.602

Crawford, T. N., Shaver, P. R., Cohen, P., Pilkonis, P. A., Gillath, O., & Kasen, S. (2006). Self-reported attachment, interpersonal aggression, and personality disorder in a prospective community sample of adolescents and adults. *Journal of Personality Disorders*, *20*(4), 331–351. doi:10.1521/pedi.2006.20.4.331

Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology*, *62*, 583–619. doi:10.1146/annurev.psych.093008.100356

De Meulemeester, C., Vansteelandt, K., Luyten, P., & Lowyck, B. (2017). Mentalizing as a mechanism of change in the treatment of patients with borderline personality disorder: A parallel process growth modeling approach. *Personality Disorders*. doi:10.1037/per0000256

Diamond, L. M., Hicks, A. M., & Otter-Henderson, K. (2006). Physiological evidence for repressive coping among avoidantly

attached adults. *Journal of Social and Personal Relationships*, *23* (2), 205–229. doi:10.1177/0265407506062470

Diamond, D., Levy, K. N., Clarkin, J. F., Fischer-Kern, M., Cain, N. M., Doering, S., … Buchheim, A. (2014). Attachment and mentalization in female patients with comorbid narcissistic and borderline personality disorder. *Personality Disorders*, *5*(4), 428–433. doi:10.1037/per0000065

Doering, S., Burgmer, M., Heuft, G., Menke, D., Bäumer, B., Lübking, M., … Schneider, G. (2013). Reliability and validity of the German version of the structured interview of personality organization (STIPO). *BMC Psychiatry*, *13*(1), 1–13. doi:10.1186/1471-244x-13-210

Doering, S., Horz, S., Rentrop, M., Fischer-Kern, M., Schuster, P., Benecke, C., … Buchheim, P. (2010). Transference-focused psychotherapy v. Treatment by community psychotherapists for borderline personality disorder: Randomised controlled trial. *British Journal of Psychiatry*, *196*(5), 389–395. doi:10.1192/bjp.bp.109.070177

Doss, B. D. (2004). Changing the way we study change in psychotherapy. *Clinical Psychology: Science and Practice*, *11*(4), 368–386. doi:10.1093/clipsy.bph094

Dozier, M., & Kobak, R. R. (1992). Psychophysiology in attachment interviews: Converging evidence for deactivating strategies. *Child Development*, *63*(6), 1473–1480. doi:10.1111/j.1467-8624.1992.tb01708.x

Dykas, M. J., & Cassidy, J. (2011). Attachment and the processing of social information across the life span: Theory and evidence. *Psychological Bulletin*, *137*(1), 19–46. doi:10.1037/a0021367

Edelstein, R. S. (2006). Attachment and emotional memory: Investigating the source and extent of avoidant memory impairments. *Emotion*, *6*(2), 340–345. doi:10.1037/1528-3542.6.2.340

Edelstein, R. S., & Gillath, O. (2008). Avoiding interference: Adult attachment and emotional processing biases. *Personality and Social Psychology Bulletin*, *34*(2), 171–181. doi:10.1177/0146167207310024

Ellison, W. D., & Levy, K. N. (2012). Factor structure of the primary scales of the Inventory of Personality Organization in a nonclinical sample using exploratory structural equation modeling. *Psychological Assessment*, *24*(2), 503–517. doi:10.1037/a0026264

Ellison, W. D., Levy, K. N., Cain, N. M., Ansell, E. B., & Pincus, A. L. (2013). The impact of pathological narcissism on psychotherapy utilization, initial symptom severity, and early-treatment symptom change: A naturalistic investigation. *Journal of Personality Assessment*, *95*(3), 291–300. doi:10.1080/00223891.2012.742904

Emmelkamp, P. M., Benner, A., Kuipers, A., Feiertag, G. A., Koster, H. C., & van Apeldoorn, F. J. (2006). Comparison of brief dynamic and cognitive-behavioural therapies in avoidant personality disorder. *British Journal of Psychiatry*, *189*, 60–64. doi:10.1192/bjp.bp.105.012153

Erikson, E. H. (1950). *Childhood and society*. New York, NY: Norton.

Falkenström, F., Solbakken, O. A., Möller, C., Lech, B., Sandell, R., & Holmqvist, R. (2014). Reflective functioning, affect consciousness, and mindfulness: Are these different functions? *Psychoanalytic Psychology*, *31*(1), 26–40. doi:10.1037/a0034049

Fischer-Kern, M., Buchheim, A., Hörz, S., Schuster, P., Doering, S., Kapusta, N. D., … Fonagy, P. (2010). The relationship between personality organization, reflective functioning, and psychiatric classification in borderline personality disorder. *Psychoanalytic Psychology*, *27*(4), 395–409. doi:10.1037/a0020862

Fischer-Kern, M., Doering, S., Taubner, S., Horz, S., Zimmermann, J., Rentrop, M., … Buchheim, A. (2015). Transference-focused psychotherapy for borderline personality disorder: Change in reflective function. *British Journal of Psychiatry*, doi:10.1192/bjp.bp.113.143842

Fischer-Kern, M., Kapusta, N. D., Doering, S., Horz, S., Mikutta, C., & Aigner, M. (2011). The relationship between personality organization and psychiatric classification in chronic pain patients. *Psychopathology*, *44*(1), 21–26. doi:10.1159/000317271

Fisher, A. J., & Boswell, J. F. (2016). Enhancing the personalization of psychotherapy with dynamic assessment and modeling. *Assessment*, *23*(4), 496–506. doi:10.1177/1073191116638735

Fonagy, P., Gergely, G., Jurist, E. L., & Target, M. (2002). *Affect regulation, mentalization, and the development of the self*. New York, NY: Other Press.

Fonagy, P., Leigh, T., Steele, M., Steele, H., Kennedy, R., Mattoon, G., … Gerber, A. (1996). The relation of attachment status, psychiatric classification, and response to psychotherapy. *Journal of Consulting and Clinical Psychology*, *64*(1), 22–31.

Fonagy, P., Luyten, P., Moulton-Perkins, A., Lee, Y.-W., Warren, F., Howard, S., … Lowyck, B. (2016). Development and validation of a self-report measure of mentalizing: The reflective functioning questionnaire. *PLoS One*, *11*(7), e0158678. doi:10.1371/journal.pone.0158678

Fonagy, P., Target, M., Steele, H., & Steele, M. (1998). *Reflecting functioning manual, version 5, for application to adult attachment interviews*. London: Unpublished manuscript.

Fournier, J. C., DeRubeis, R. J., & Beck, A. T. (2012). Dysfunctional cognitions in personality pathology: The structure and validity of the personality belief questionnaire. *Psychological Medicine*, *42*, 795–805.

Fraley, R. C., Roisman, G. I., Booth-LaForce, C., Owen, M. T., & Holland, A. S. (2013). Interpersonal and genetic origins of adult attachment styles: A longitudinal study from infancy to early adulthood. *Journal of Personality and Social Psychology*, *104*(5), 817–838. doi:10.1037/a0031435

Fraley, R. C., Waller, N. G., & Brennan, K. A. (2000). An item response theory analysis of self-report measures of adult attachment. *Journal of Personality and Social Psychology*, *78*(2), 350–365.

Freud, A. (1937/1966). *Ego and the mechanisms of defense*. London: Karnac Books.

Friborg, O., Martinsen, E. W., Martinussen, M., Kaiser, S., Overgard, K. T., & Rosenvinge, J. H. (2014). Comorbidity of personality disorders in mood disorders: A meta-analytic review of 122 studies from 1988 to 2010. *Journal of Affective Disorders*, *152-154*, 1–11. doi:10.1016/j.jad.2013.08.023

Friborg, O., Martinussen, M., Kaiser, S., Overgard, K. T., Martinsen, E. W., Schmierer, P., & Rosenvinge, J. H. (2014). Personality disorders in eating disorder not otherwise specified and binge eating disorder: A meta-analysis of comorbidity studies. *Journal of Nervous and Mental Disease*, *202*(2), 119–125. doi:10.1097/nmd.0000000000000080

Friborg, O., Martinussen, M., Kaiser, S., Øvergård, K. T., & Rosenvinge, J. H. (2013). Comorbidity of personality disorders in anxiety disorders: A meta-analysis of 30 years of research. *Journal of Affective Disorders*, *145*(2), 143–155. doi:10.1016/j.jad.2012.07.004

George, C., Kaplan, N., & Main, M. (1985). *The Berkeley adult attachment interview*. Department of Psychology. Berkeley: University of California. Retrieved from http://www.psychology.sunysb.edu/attachment/measures/content/aai_interview.pdf

Gillath, O., Karantzas, G. C., & Fraley, R. C. (2016). *Adult attachment: A concise and introduction to theory and research*. San Diego, CA: Academic Press/Elsevier.

Gouin, J.-P., Glaser, R., Loving, T. J., Malarkey, W. B., Stowell, J., Houts, C., & Kiecolt-Glaser, J. K. (2009). Attachment avoidance predicts inflammatory responses to marital conflict.

Brain, Behavior, and Immunity, *23*(7), 898–904. doi:10.1016/j.bbi.2008.09.016

Grilo, C. M., Stout, R. L., Markowitz, J. C., Sanislow, C. A., Ansell, E. B., Skodol, A. E., & McGlashan, T. H. (2010). Personality disorders predict relapse after remission from an episode of major depressive disorder: A 6-year prospective study. *Journal of Clinical Psychiatry*, *71*(12), 1629–1635. doi:10.4088/JCP.08m04200gre

Gunderson, J. G., Stout, R. L., McGlashan, T. H., Shea, M. T., Morey, L. C., Grilo, C. M., … Skodol, A. E. (2011). Ten-year course of borderline personality disorder: Psychopathology and function from the Collaborative Longitudinal Personality Disorders study. *Archives of General Psychiatry*, *68*(8), 827–837. doi:10.1001/archgenpsychiatry.2011.37

Haggerty, G. D., Siefert, C. J., & Weinberger, J. (2010). Examining the relationship between current attachment status and freely recalled autobiographical memories of childhood. *Psychoanalytic Psychology*, *27*(1), 27–41. doi:10.1037/a0018638

Hill, J., Pilkonis, P., Morse, J., Feske, U., Reynolds, S., Hope, H., … Broyden, N. (2008). Social domain dysfunction and disorganization in borderline personality disorder. *Psychological Medicine*, *38*(1), 135–146. doi:10.1017/s0033291707001626

Hoffart, A., Versland, S., & Sexton, H. (2002). Self-Understanding, empathy, guided discovery, and schema belief in schema-focused cognitive therapy of personality problems: A process–outcome study. *Cognitive Therapy and Research*, *26*(2), 199–219. doi:10.1023/a:1014521819858

Høglend, P., Bøgwald, K. P., Amlo, S., Marble, A., Ulberg, R., Sjaastad, M. C., … Johansson, P. (2008). Transference interpretations in dynamic psychotherapy: Do they really yield sustained effects? *American Journal of Psychiatry*, *165*(6), 763–771. doi:10.1176/appi.ajp.2008.07061028

Hoglend, P., Dahl, H. S., Hersoug, A. G., Lorentzen, S., & Perry, J. C. (2011). Long-term effects of transference interpretation in dynamic psychotherapy of personality disorders. *European Psychiatry*, *26*(7), 419–424. doi:10.1016/j.eurpsy.2010.05.006

Hopwood, C. J., Schade, N., Krueger, R. F., Wright, A. G., & Markon, K. E. (2013). Connecting DSM-5 personality traits and pathological beliefs: Toward a unifying model. *Journal of Psychopathology and Behavioral Assessment*, *35*(2), 162–172. doi:10.1007/s10862-012-9332-3

Hopwood, C. J., Thomas, K. M., Markon, K. E., Wright, A. G. C., & Krueger, R. F. (2012). DSM-5 personality traits and DSM-IV personality disorders. *Journal of Abnormal Psychology*, *121*(2), 424–432. doi:10.1037/a0026656

Hopwood, C. J., Zimmermann, J., Pincus, A. L., & Krueger, R. F. (2015). Connecting personality structure and dynamics: Towards a more evidence-based and clinically useful diagnostic scheme. *Journal of Personality Disorders*, *29*(4), 431–448. doi:10.1521/pedi.2015.29.4.431

Huang, Y., Kotov, R., de Girolamo, G., Preti, A., Angermeyer, M., Benjet, C., … Kessler, R. C. (2009). DSM-IV personality disorders in the WHO world mental health surveys. *British Journal of Psychiatry*, *195*(1), 46–53. doi:10.1192/bjp.bp.108.058552

Jacob, G. A., & Arntz, A. (2013). Schema Therapy for Personality Disorders—A Review. *International Journal of Cognitive Therapy*, *6*(2), 171–185. doi:10.1521/ijct.2013.6.2.171

Jacobvitz, D., Curran, M., & Moller, N. (2002). Measurement of adult attachment: The place of self-report and interview methodologies. *Attachment & Human Development*, *4*(2), 207–215. doi:10.1080/14616730210154225

Johansen, P. O., Krebs, T. S., Svartberg, M., Stiles, T. C., & Holen, A. (2011). Change in defense mechanisms during short-term dynamic and cognitive therapy in patients with cluster C personality disorders. *Journal of Nervous and Mental Disease*, *199*(9), 712–715. doi:10.1097/NMD.0b013e318229d6a7

Johansson, P., Hoglend, P., Ulberg, R., Amlo, S., Marble, A., Bogwald, K. P., … Heyerdahl, O. (2010). The mediating role of insight for long-term improvements in psychodynamic therapy. *Journal of Consulting and Clinical Psychology*, *78*(3), 438–448. doi:10.1037/a0019245

Jones, S. H., Burrell-Hodgson, G., & Tate, G. (2007). Relationships between the personality beliefs questionnaire and self-rated personality disorders. *The British Journal of Clinical Psychology*, *46*(Pt 2), 247–251. doi:10.1348/014466506x164791

Kallestad, H., Valen, J., McCullough, L., Svartberg, M., Høglend, P., & Stiles, T. C. (2010). The relationship between insight gained during therapy and long-term outcome in short-term dynamic psychotherapy and cognitive therapy for cluster C personality disorders. *Psychotherapy Research*, *20*(5), 526–534. doi:10.1080/10503307.2010.492807

Katznelson, H. (2014). Reflective functioning: A review. *Clinical Psychology Review*, *34*(2), 107–117. doi:http://doi.org/10.1016/j.cpr.2013.12.003

Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology*, *3*(1), 1–27. doi:10.1146/annurev.clinpsy.3.022806.091432

Keefe, J. R., Webb, C. A., & DeRubeis, R. J. (2016). In cognitive therapy for depression, early focus on maladaptive beliefs may be especially efficacious for patients with personality disorders. *Journal of Consulting and Clinical Psychology*, *84*(4), 353–364. doi:10.1037/ccp0000071

Kernberg, O. F. (1984). *Severe personality disorders: Psychotherapeutic strategies*. New Haven, CT: Yale University Press.

Kramer, U., de Roten, Y., Perry, J. C., & Despland, J.-N. (2013). Beyond splitting: Observer-rated defense mechanisms in borderline personality disorder. *Psychoanalytic Psychology*, *30*(1), 3–15. doi:10.1037/a0029463

Lenzenweger, M. F., Clarkin, J. F., Kernberg, O. F., & Foelsch, P. A. (2001). The inventory of personality organization: Psychometric properties, factorial composition, and criterion relations with affect, aggressive dyscontrol, psychosis proneness, and self-domains in a nonclinical sample. *Psychological Assessment*, *13*, 577–591. doi:10.1037/1040-3590.13.4.577

Lenzenweger, M. F., Lane, M. C., Loranger, A. W., & Kessler, R. C. (2007). DSM-IV personality disorders in the national comorbidity survey replication. *Biological Psychiatry*, *62*(6), 553–564. doi:10.1016/j.biopsych.2006.09.019

Levy, K. N. (2005). The implications of attachment theory and research for understanding borderline personality disorder. *Dev Psychopathol*, *17*(4), 959–986.

Levy, K. N., Johnson, B. N., Clouthier, T. L., Scala, J. W., & Temes, C. M. (2015). An attachment theoretical framework for personality disorders. *Canadian Psychology/Psychologie Canadienne*, *56*(2), 197–207. doi:10.1037/cap0000025

Levy, K. N., Meehan, K. B., Kelly, K. M., Reynoso, J. S., Weber, M., Clarkin, J. F., & Kernberg, O. F. (2006). Change in attachment patterns and reflective function in a randomized control trial of transference-focused psychotherapy for borderline personality disorder. *Journal of Consulting and Clinical Psychology*, *74*(6), 1027–1040. doi:10.1037/0022-006X.74.6.1027

Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder*. New York, NY: Guilford.

Lorenzo-Luaces, L., German, R. E., & DeRubeis, R. J. (2014). It's complicated: The relation between cognitive change procedures, cognitive change, and symptom change in cognitive therapy for depression. *Clinical Psychology Review*, 41, 3–15. doi:10.1016/j.cpr.2014.12.003

MacDonald, K., Berlow, R., & Thomas, M. L. (2013). Attachment, affective temperament, and personality disorders: A study of their relationships in psychiatric outpatients. *Journal of Affective Disorders*, *151*(3), 932–941. doi:http://doi.org/10.1016/j.jad.2013.07.040

Markowitz, J. C., Skodol, A. E., Petkova, E., Cheng, J., Sanislow, C. A., Grilo, C. M., … McGlashan, T. H. (2007). Longitudinal effects of personality disorders on psychosocial functioning of patients with major depressive disorder. *Journal of Clinical Psychiatry*, *68*(2), 186–193.

Meyer, B., & Pilkonis, P. A. (2005). An attachment model of personality disorders. In M. F. Lenzenweger, & J. F. Clarkin (Eds.), *Major theories of personality disorder* (2nd ed., pp. 231–281). New York, NY: Guilford.

Mikulincer, M., Dolev, T., & Shaver, P. R. (2004). Attachment-related strategies during thought suppression: Ironic rebounds and vulnerable self-representations. *Journal of Personality and Social Psychology*, *87*(6), 940–956. doi:10.1037/0022-3514.87.6.940

Mikulincer, M., Shaver, P. R., & Berant, E. (2013). An attachment perspective on therapeutic processes and outcomes. *Journal of Personality*, *81*(6), 606–616. doi:10.1111/j.1467-6494.2012.00806.x

Morey, L. C., & Hopwood, C. J. (2013). Stability and change in personality disorders. *Annual Review of Clinical Psychology*, *9*(1), 499–528. doi:doi:10.1146/annurev-clinpsy-050212-185637

Nazzaro, M. P., Boldrini, T., Tanzilli, A., Muzi, L., Giovanardi, G., & Lingiardi, V. (2017). Does reflective functioning mediate the relationship between attachment and personality? *Psychiatry Research*, *256*, 169–175. doi:http://doi.org/10.1016/j.psychres.2017.06.045

Newton-Howes, G., Tyrer, P., Johnson, T., Mulder, R., Kool, S., Dekker, J., & Schoevers, R. (2014). Influence of personality on the outcome of treatment in depression: Systematic review and meta-analysis. *Journal of Personality Disorders*, *28*(4), 577–593. doi:10.1521/pedi_2013_27_070

Ogrodniczuk, J. S., Piper, W. E., Joyce, A. S., Steinberg, P. I., & Duggal, S. (2009). Interpersonal problems associated with narcissism among psychiatric outpatients. *Journal of Psychiatric Research*, *43*(9), 837–842. doi:http://doi.org/10.1016/j.jpsychires.2008.12.005

Olssøn, I., & Dahl, A. A. (2014). Comparison of patterns and strength of adult attachment in patients with axis I, axis I + II disorders and a community sample. *Nordic Journal of Psychiatry*, *68*(7), 464–471. doi:10.3109/08039488.2013.859298

Oltmanns, T. F., Rodrigues, M. M., Weinstein, Y., & Gleason, M. E. (2014). Prevalence of personality disorders at midlife in a community sample: Disorders and symptoms reflected in interview, self, and informant reports. *Journal of Psychopathology and Behavioral Assessment*, *36*(2), 177–188. doi:10.1007/s10862-013-9389-7

Perry, J. C., & Bond, M. (2005). Defensive functioning [in personality disorders]. In J. Oldham, A. E. Skodol, & D. Bender (Eds.), *The American psychiatric publishing textbook of personality disorders* (pp. 589–609). Washington, DC: American Psychiatric Press.

Perry, J. C., & Bond, M. (2012). Change in defense mechanisms during long-term dynamic psychotherapy and five-year outcome. *American Journal of Psychiatry*, *169*(9), 916–925. doi:10.1176/appi.ajp.2012.11091403

Perry, J. C., Presniak, M. D., & Olson, T. R. (2013). Defense mechanisms in schizotypal, borderline, antisocial, and narcissistic personality disorders. *Psychiatry*, *76*(1), 32–52. doi:10.1521/psyc.2013.76.1.32

Preti, E., Prunas, A., De Panfilis, C., Marchesi, C., Madeddu, F., & Clarkin, J. F. (2015). The facets of identity: Personality pathology assessment through the inventory of personality organization. *Personality Disorders*, *6*(2), 129–140. doi:10.1037/per0000119

Roisman, G. I., Holland, A., Fortuna, K., Fraley, R. C., Clausell, E., & Clarke, A. (2007). The adult attachment interview and self-reports of attachment style: An empirical rapprochement. *Journal of Personality and Social Psychology*, *92*(4), 678–697. doi:10.1037/0022-3514.92.4.678

Rossouw, T. I., & Fonagy, P. (2012). Mentalization-Based treatment for self-harm in adolescents: A randomized controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(12), 1304–1313.e1303. doi:http://doi.org/10.1016/j.jaac.2012.09.018

Ryum, T., Stiles, T. C., Svartberg, M., & McCullough, L. (2010). The role of transference work, the therapeutic alliance, and their interaction in reducing interpersonal problems among psychotherapy patients with cluster C personality disorders. *Psychotherapy: Theory, Research, Practice, Training*, *47*(4), 442–453. doi:10.1037/a0021183

Sadikaj, G., Moskowitz, D. S., Russell, J. J., Zuroff, D. C., & Paris, J. (2013). Quarrelsome behavior in borderline personality disorder: Influence of behavioral and affective reactivity to perceptions of others. *Journal of Abnormal Psychology*, *122*(1), 195–207. doi:10.1037/a0030871

Sharp, C., Williams, L. L., Ha, C., Baumgardner, J., Michonski, J., Seals, R., … Fonagy, P. (2009). The development of a mentalization-based outcomes and research protocol for an adolescent inpatient unit. *Bulletin of the Menninger Clinic*, *73*(4), 311–338. doi:10.1521/bumc.2009.73.4.311

Sharp, C., Wright, A. G. C., Fowler, J. C., Frueh, B. C., Allen, J. G., Oldham, J., & Clark, L. A. (2015). The structure of personality pathology: Both general ('g') and specific ('s') factors? *Journal of Abnormal Psychology*, *124*(2), 387–398. doi:10.1037/abn0000033

Shaver, P. R., Belsky, J. A. Y., & Brennan, K. A. (2000). The adult attachment interview and self-reports of romantic attachment: Associations across domains and methods. *Personal Relationships*, *7*(1), 25–43. doi:10.1111/j.1475-6811.2000.tb00002.x

Shaver, P. R., & Mikulincer, M. (2002). Attachment-related psychodynamics. *Attachment & Human Development*, *4*(2), 133–161. doi:10.1080/14616730210154171

Shaver, P. R., & Mikulincer, M. (2005). Attachment theory and research: Resurrection of the psychodynamic approach to personality. *Journal of Research in Personality*, *39*(1), 22–45. doi:http://doi.org/10.1016/j.jrp.2004.09.002

Skodol, A. E., Geier, T., Grant, B. F., & Hasin, D. S. (2014). Personality disorders and the persistence of anxiety disorders in a nationally representative sample. *Depression and Anxiety*, *31*(9), 721–728. doi:10.1002/da.22287

Skodol, A. E., Grilo, C. M., Pagano, M. E., Bender, D. S., Gunderson, J. G., Shea, M. T., … McGlashan, T. H. (2005). Effects of personality disorders on functioning and well-being in major depressive disorder. *Journal of Psychiatric Practice*, *11*(6), 363–368.

Spielmann, S. S., Maxwell, J. A., Macdonald, G., & Baratta, P. L. (2013). Don't get your hopes up: Avoidantly attached individuals perceive lower social reward when there is potential for intimacy. *Personality and Social Psychology Bulletin*, *39*(2), 219–236. doi:10.1177/0146167212472541

Stern, B. L., Caligor, E., Clarkin, J. F., Critchfield, K. L., Horz, S., MacCornack, V., … Kernberg, O. F. (2010). Structured interview of personality organization (STIPO): preliminary psychometrics in a clinical sample. *Journal of Personality Assessment*, *92*, 35–44. doi:10.1080/00223890903379308

Summers, R. F., & Barber, J. P. (2010). *Psychodynamic therapy: A guide to evidence-based practice*. New York, NY: Guilford.

Svartberg, M., Stiles, T. C., & Seltzer, M. H. (2004). Randomized, controlled trial of the effectiveness of short-term dynamic psychotherapy and cognitive therapy for cluster C personality disorders. *American Journal of Psychiatry*, *161*(5), 810–817.

Taubner, S, Hörz, S, Fischer-Kern, M, Doering, S, Buchheim, A, & Zimmermann, J. (2013). Internal structure of the Reflective Functioning Scale. *Psychological Assessment*, *25*(1), 127–135. doi:10.1037/a0029138

Taylor, C. D. J., Bee, P., & Haddock, G. (2017). Does schema therapy change schemas and symptoms? A systematic review across mental health disorders. *Psychology and Psychotherapy: Theory, Research and Practice*, *90*(3), 456–479. doi:10.1111/papt.2017.90.issue-3

Vaillant, G. E. (1994). Ego mechanisms of defense and personality psychopathology. *Journal of Abnormal Psychology*, *103*(1), 44–50.

Vaillant, G. E., Bond, M., & Vaillant, C. O. (1986). An empirically validated hierarchy of defense mechanisms. *Archives of General Psychiatry*, *43*(8), 786–794.

Verreault, M., Sabourin, S., Lussier, Y., Normandin, L., & Clarkin, J. F. (2013). Assessment of personality organization in couple relationships: Factorial structure of the inventory of personality organization and incremental validity over neuroticism. *Journal of Personality Assessment*, *95*(1), 85–95. doi:10.1080/00223891.2012.713883

Wright, A. G. C., Calabrese, W. R., Rudick, M. M., Yam, W. H., Zelazny, K., Williams, T. F., … Simms, L. J. (2015). Stability of the DSM-5 section III pathological personality traits and their longitudinal associations with psychosocial functioning in personality disordered individuals. *Journal of Abnormal Psychology*, *124*(1), 199–207. doi:10.1037/abn0000018

Yaseen, Z. S., Zhang, X., Muran, J. C., Winston, A., & Galynker , I. I. (2016). Comparison of brain activity correlating with self-report versus narrative attachment measures during conscious appraisal of an attachment figure. *Frontiers in Human Neuroscience*, *10*, 90. doi:10.3389/fnhum.2016.00090

Yeomans, F. E., Clarkin, J. F., & Kernberg, O. F. (2015). *Transference-focused psychotherapy for borderline personality disorder*. Washington, DC: American Psychiatric Publishing.

Zanarini, M. C., Frankenburg, F. R., & Fitzmaurice, G. (2013). Defense mechanisms reported by patients with borderline personality disorder and axis II comparison subjects over 16 years of prospective follow-up: Description and prediction of recovery. *American Journal of Psychiatry*, *170*(1), 111–120. doi:10.1176/appi.ajp.2012.12020173

Zanarini, M. C., Frankenburg, F. R., Reich, D. B., & Fitzmaurice, G. (2012). Attainment and stability of sustained symptomatic remission and recovery among patients with borderline personality disorder and axis II comparison subjects: A 16-year prospective follow-up study. *American Journal of Psychiatry*, *169*(5), 476–483.

Zanarini, M. C., Weingeroff, J. L., & Frankenburg, F. R. (2009). Defense mechanisms associated with borderline personality disorder. *Journal of Personality Disorders*, *23*(2), 113–121. doi:10.1521/pedi.2009.23.2.113