HHS Public Access

Author manuscript

J Abnorm Psychol. Author manuscript; available in PMC 2019 April 02.

Published in final edited form as:

J Abnorm Psychol. 2009 August ; 118(3): 507–519. doi:10.1037/a0016478.

Ten-Year Stability and Latent Structure of the *DSM–IV* Schizotypal, Borderline, Avoidant, and Obsessive-Compulsive Personality Disorders

Charles A. Sanislow,
Department of Psychiatry, Yale University School of Medicine

Emily B. Ansell,
Department of Psychiatry, Yale University School of Medicine

Carlos M. Grilo,
Department of Psychiatry, Yale University School of Medicine

John C. Markowitz,
Department of Psychiatry, Columbia University and New York State Psychiatric Institute, New York, New York

Anthony Pinto,
Department of Psychiatry, Columbia University and New York State Psychiatric Institute, New York, New York

Shirley Yen,
Department of Psychiatry and Human Behavior, Brown University

Leslie C. Morey,
Department of Psychology, Texas A & M University

Todd D. Little,
Department of Psychology, University of Kansas and Schiefelbusch Institute for Life Span Studies, Lawrence, Kansas

Maria Daversa,
Department of Psychiatry, Harvard University Medical School and McLean Hospital, Belmont, Massachusetts

M. Tracie Shea,
Veterans Affairs Medical Center, Providence, Rhode Island, and Brown University Alpert Medical School

Correspondence concerning this article should be addressed to Charles A. Sanislow, who is now at the Department of Psychology, Wesleyan University, 207 High Street, Middletown, CT 06459. csanislow@wesleyan.edu.

Publisher's Disclaimer: This article was approved by the publication committee of the Collaborative Longitudinal Personality Disorders Study. Funding for the Collaborative Longitudinal Personality Disorders Study is provided by National Institutes of Health Grants MH050837 to Brown University Department of Psychiatry and Human Behavior, MH050839 to Columbia University and New York State Psychiatric Institute, MH050840 to Harvard Medical School and McLean Hospital, MH050838 to Texas A&M University, and MH050850 to Yale University School of Medicine; further support for this study was provided by National Institutes of Health Grants MH079078 to John C. Markowitz, MH080221 to Anthony Pinto, MH073708 to Charles A. Sanislow, and MH069904 to Shirley Yen.

Andrew E. Skodol,
Department of Psychiatry, University of Arizona College of Medicine and the Sunbelt Collaborative, Tucson, Arizona

Thomas H. McGlashan,
Department of Psychiatry, Yale University School of Medicine

John G. Gunderson, and
Department of Psychiatry, Harvard University Medical School and McLean Hospital, Belmont, Massachusetts

Mary C. Zanarini
Department of Psychiatry, Harvard University Medical School and McLean Hospital, Belmont, Massachusetts

Abstract

Evaluation of the validity of personality disorder (PD) diagnostic constructs is important for the impending revision of the *Diagnostic and Statistical Manual of Mental Disorders*. Prior factor analytic studies have tested these constructs in cross-sectional studies, and models have been replicated longitudinally, but no study has tested a constrained longitudinal model. The authors examined 4 PDs in the Collaborative Longitudinal Personality Disorders study (schizotypal, borderline, avoidant, and obsessive-compulsive) over 7 time points (baseline, 6 months, 1 year, 2 years, 4 years, 6 years, and 10 years). Data for 2-, 4-, 6- and 10-year assessments were obtained in semistructured interviews by raters blind to prior PD diagnoses at each assessment. The latent structure of the 4 constructs was differentiated during the initial time points but became less differentiated over time as the mean levels of the constructs dropped and stability increased. Obsessive-compulsive PD became more correlated with schizotypal and borderline PD than with avoidant PD. The higher correlation among the constructs in later years may reflect greater shared base of pathology for chronic personality disorders.

Keywords

schizotypal; borderline; avoidant; obsessive-compulsive; personality disorders

---

The *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*) defines personality disorders (PDs) as stable and enduring, reflecting a persistent pattern of maladaptive personality through-out the life course. Approaches used to evaluate PD construct validity include testing this stability assumption by examining time to remission for PD diagnosis, stability of criteria within the diagnosis, and factor structure of the PD diagnostic constructs or clusters. Prospective tests of stability by several research groups, including our own Collaborative Longitudinal Study of Personality Disorders (CLPS), have shown that PDs tend to remit at rates higher than the *DSM* definition implies (e.g., Grilo et al., 2004; Shea et al., 2002; see also Laptook, Klein, & Dougherty, 2006; Zanarini, Frankenburg, Hennen, & Silk, 2003; Zanarini et al., 2007). Individual variability of PDs across time has also emerged from nonclinical samples (Lenzenweger, Johnson, & Willett, 2004).

A second approach to testing the validity of *DSM* PDs is to examine their latent structure. Here, results have been mixed. Evidence from other studies supports the *DSM* constructs and suggests that subdiagnostic levels of PD pathology have prognostic value. Using latent class analyses, Clifton and Pilkonis (2007) identified a group of individuals with subclinical borderline PD (BPD) diagnoses who more closely resembled individuals meeting full diagnostic criteria in their social-interpersonal and occupational functioning than they did non-BPD participants. Dimensional scoring of *DSM* criteria has been shown to more accurately predict psychosocial functioning than do PD categories (Skodol, Oldham, et al., 2005). In some studies, *DSM* PD diagnoses appear more stable when examined dimensionally than when examined categorically (e.g., Morey et al., 2007). Further, tests of the stability of the relative order of PD criteria suggest that individuals remain consistent in rank order of criteria over time, even when they fluctuate in severity or number of PD features (Grilo et al., 2004). Together, these findings modestly support the validity of the PD constructs.

In several factor analytic studies with exploratory or confirmatory factor approaches, researchers have examined the *DSM* PD constructs. These studies have mainly addressed the three PD clusters: Cluster A is odd–eccentric (paranoid, schizoid, and schizotypal), Cluster B is dramatic–emotional–erratic (antisocial, borderline, histrionic, and narcissistic), and Cluster C is anxious–fearful (avoidant, dependent, and obsessive compulsive; American Psychiatric Association [APA], 1996). One notable inconsistency is whether loadings of constructs for the individual disorders empirically conform to the three clusters specified by the *DSM*. Although support for a three-factor solution was found in a non-clinical population, loadings did not correspond to the three *DSM* clusters (Moldin, Rice, Erlenmeyer-Kimling, & Squires-Wheeler, 1994). Using data obtained from clinician ratings of PDs in adolescents, Durrett and Westen (2005) found support for the individual PD diagnostic constructs but not for the three-cluster organization. Bell and Jackson (1992), attempting to fit a three-factor solution to an inpatient clinical sample, found that although fit statistics were less than optimal, the data best corresponded to the three *DSM* clusters. O'Connor and Dyce (1998) did a comparative analysis using self-report measures of personality as well as *DSM* diagnoses in a clinical sample carefully selected for a broad range of pathology. They concluded that five- and seven-factor models fit the data better than did the three-cluster *DSM* model. However, there was no incremental gain in model fit beyond four factors. In sum, there is little empirical consensus for the optimal number of higher order factors for personality pathology.

Several studies raised a second, more specific question, namely whether obsessive-compulsive PD (OCPD) stands apart from the three clusters (e.g., Hyler & Lyons, 1988; Kass, Skodol, Charles, Spitzer, & Williams, 1985; Livesley, Jackson, & Schroeder, 1992; Livesley, Jang, & Vernon, 1998; Nestadt et al., 1994; Tyrer & Alexander, 1979). Morey (1986), however, demonstrated that a three-cluster solution including OCPD could be forced with Procrustean procedures with the Kass et al. (1985) data. In an Italian patient sample, Fossati et al. (2000) found results supporting three factors, but only one (odd– eccentric) aligned with the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; *DSM–IV*; APA, 1996) clusters. Yang, Bagby, Costa, Ryder, and Herbst (2002) tested the *DSM–IV* cluster structure in a sample of Chinese patients, and their results did not support the three

clusters. Rodebaugh, Chambless, Renneberg, and Fydrich (2005) analyzed a combination of archival data sets, and their results revealed better support for a three-factor model than for one-factor model with confirmatory tests. Finally, O'Connor (2005) reported good fit for a three-factor model with data based on the five-factor model approach (Neuroticism, low Agreeableness, Extroversion/introversion) but found that adding a fourth factor capturing Conscientiousness and obsessive-compulsive features better accounted for the data. In sum, patient studies testing the *DSM* three-cluster model have produced mixed results.

The CLPS research group has also studied the covariance structures of PD constructs and found some support for the *DSM* constructs (Sanislow et al., 2002). Specifically, we tested a four-factor structural model corresponding to the *DSM–IV* diagnostic constructs schizotypal PD (STPD), BPD, avoidant PD (AVPD), and OCPD. Results showed good model fit at baseline (Year 0) and again at 2-year follow-up, based on diagnoses made by raters blind to baseline diagnoses (Sanislow et al., 2002). Although these findings supported the structure implied by the *DSM* diagnostic constructs, they did not uphold the relative weighting implicit in the ordering of symptom criteria within each diagnostic construct (APA, 1996). More recent work has also demonstrated inconsistencies in the criterion hierarchy for BPD (Karterud, Pedersen, Gude, & Falkum, 2004). However, variation in *DSM* criteria hierarchies across studies is not necessarily surprising, given that some criteria serve different functions for the constructs (e.g., predictive of the construct versus evidencing stability). Further, it seems reasonable to expect these functions, as well as the relation of the criteria to the constructs, to vary depending on the population sampled. For instance, suicidal behavior may predict a poorer outcome in a clinical sample than in a non–treatment-seeking sample. In a review of PD factor studies, Sheets and Craighead (2007) concluded that studies testing *DSM* structure with nonpatient community samples generally showed less support for the *DSM* than did those with patient populations.

Evaluating PD validity by testing stability at the level of the diagnostic construct can supersede these influences and fluctuations. Here, we examine the stability of four PD constructs longitudinally as well as their overlap with the other PD constructs. We extend prior work (Sanislow et al., 2002) in two ways. First, we test the stability of the four CLPS PDs (STPD, BPD, AVPD, and OCPD) over a longer, 10-year interval. Second, we examine the PD constructs across seven assessment points in a single longitudinal model. That is, rather than testing separate models at each time point (cf. Sanislow et al., 2002), we tested a single panel model using the entire 10-year CLPS sample to directly evaluate the stability of the constructs. PDs were modeled in a large treatment-seeking sample at seven time points: Year 0 (baseline at study entry), Month 6, and Year 1, then Year 2, Year 4, Year 6, and Year 10. Participants entered the study with a primary PD diagnosis of STPD, BPD, AVPD, or OCPD or with a diagnosis of major depressive disorder with no PD. Participants targeted for one of the four PDs were not excluded for the presence of comorbid PDs. To examine the joint characteristics of change across time, growth curves were estimated. The panel model allowed estimation of the stability of the individual constructs, whereas the growth curve model allowed estimation of the nature of the specific characteristics of change over time. Models were controlled for demographic characteristics of age, sex, and race. On the basis of the *DSM–IV* premise of stability and distinctiveness, we hypothesized that compared with diagnostic approaches based on criterion cutoffs, the *DSM* constructs STPD, BPD, and

AVPD would show stability within constructs and discriminant validity between constructs in the omnibus model. Given prior conflicting findings, we were uncertain whether the OCPD construct would demonstrate greater associations over time with the other Cluster C disorder, AVPD.

Method

Participants

Participants aged 18–45 years at study entry were evaluated as part of the CLPS. The CLPS is a prospective, repeated measures study that examined the course of PDs. For a more detailed description of the study design and aims, see Gunderson et al. (2000); for sample characteristics, see McGlashan et al. (2000). Primarily treatment-seeking participants at inpatient or outpatient facilities who were or had recently been in psychiatric treatment or psychotherapy were sampled for four representative PDs (borderline, schizotypal, avoidant, and obsessive-compulsive); a control group meeting criteria for major depressive disorder but no PD was also included. Media advertising and postings supplemented recruitment. Potential participants were prescreened to determine age eligibility and treatment status or history and to assist in excluding patients with active psychosis, acute substance intoxication or withdrawal, a history of schizophrenia-spectrum psychosis (i.e., schizophrenia, schizophreniform, or schizoaffective disorders), or organicity. The sample comprised 733 participants. The original cohort consisted of 668 participants followed through 10 years. The supplemental cohort of 65 additional minority participants, which was sampled to provide a more representative racial base, was not followed beyond Year 4 (Year 6 and Year 10 values were imputed as described below). The sample was 69% Caucasian, 15% African American, and 13% Hispanic, with the remainder from other ethnic backgrounds; 64% were women and 36% were men. All participants provided informed, written consent to study procedures prior to entry.

Three disorders were chosen to represent the *DSM–IV* Axis II Clusters A, B, and C (STPD, BPD, and AVPD, respectively). The fourth disorder, OCPD, was included because evidence suggested it might stand apart from the three clusters. The four targeted PD diagnoses and the treatment-seeking sample were drawn from varied settings that provided a spectrum of PD pathology, a distribution enhanced by the major depressive disorder contrast group. Presence of other PDs was not an exclusion criterion, and participants received 2.1 Axis II diagnoses on average, a rate comparable with other clinical studies (e.g., Blashfield, McElroy, Pfohl, & Blum, 1994; Oldham et al., 1995; Stuart et al., 1998). Further, participants' treatment-seeking status provided an ecologically valid study group.

Assessment

Extensively trained research interviewers with master's or doctoral degrees assessed all participants, and researchers were monitored for ongoing reliability. The Structured Clinical Interview for *DSM–IV* Axis I Disorders–Patient Version (SCID-I/P; First, Spitzer, Gibbon, & Williams, 1996) was used to assess Axis I disorders, and the Diagnostic Interview for *DSM–IV* Personality Disorders (DIPD-IV; Zanarini, Frankenburg, Sickel, & Yong, 1996) was used for Axis II disorders. The DIPD-IV is a semistructured diagnostic interview

containing several questions pertaining to each *DSM–IV* Axis II criterion. Each criterion is scored 0 for *absent*, 1 for *present but of uncertain clinical significance*, or 2 for *present and clinically significant*. In our sample, median kappa coefficients (Cohen, 1960) ranged from .69 to .97 for all Axis II disorders (Zanarini et al., 2000). The DIPD-IV was administered at baseline (Year 0). The Month 6 and Year 1 assessments for the four PDs had a modified version of the DIPD-IV, the DIPD-Follow Along Version (FAV; Zanarini & Shea, 1996), in which ratings are made on a scale with 0, 1, or 2 for each month during the time period being queried. Reliability on the DIPD-IV-FAV based on the rating of two overlapping time points (Month 6 was rated twice for 453 cases) resulted in kappa coefficients of .78 for STPD, .70 for BPD, .73 for AVPD, and .68 for OCPD (see Shea et al., 2002). The Month 6 and Year 1 assessments were followed by blind assessments with the DIPD-IV; interviewers had no knowledge of participants' PD diagnostic status from prior interviews at Years 2, 4, 6, and 10.

Analyses

Structural equation modeling (SEM; Hoyle & Smith, 1994) was used with LISREL 8.80 software (Jöreskog & Sörbom, 2008) with maximum likelihood estimation. Latent variables were computed for each of the four PDs—STPD, BPD, AVPD, and OCPD—to address our key question regarding the stability of the constructs these *DSM–IV* diagnoses represented. To represent the constructs, the indicators (i.e., individual PD criteria) were parceled following recommendations by Kishton and Widaman (1994; see also Little, Cunningham, Shahar, & Widaman, 2002). For the ratings obtained with the DIPD-IV-FAV, parcels were derived from the averaged values of each criterion over the 6 month assessment interval. Parceling items offers several advantages over use of individual criteria as indicators that are pertinent to our goals (Little et al., 2002). Aggregate sets of items produce indicators that are more likely to have continuous properties and to have a more normal distribution that better fulfills the maximum likelihood assumptions than do nonparceled criteria. For model estimation, parcels require fewer parameter estimates than do models that use the items individually. Finally, parcels are more reliable than items; hence, error variances of the parceled sets of items are smaller than are the items themselves.

To construct parcels for the present study, preliminary analyses of the items were carried out to determine the optimal balanced groupings of items. Three parcels were formed empirically for each of the four disorders, so that each grouping of averaged items evenly represented the common variance of the construct (i.e., all parcels exhibited an evenly distributed range of intercorrelations among the component criteria used to compose the parcels; see Little et al., 2002, for details of creating balanced parcels). The composition of the parcels, held constant across the assessment time points, is shown in Table 1.

Missing Data

Across the 10 years of the study, only 16.4% of the overall data was missing, meeting acceptable standards of less than 20%, to use modern imputation procedures (Schafer & Graham, 2002). For this study, we used the SAS procedure Proc MI (Version 9.12) to address missing data, specified the expectation maximization (EM) algorithm to establish prior estimates, and used the Markcov chain Monte Carlo procedure (MCMC) to impute

missing values. The imputation process included sex, race, and age as well as all diagnostic variables and a participant's membership in the original cohort or the supplemental minority sample (including all appropriate interaction terms). The imputation procedure was run 100 times to ensure maximal generalizability, given the presence of missing data (Enders, in press).

Evaluation of Model Fit

Three fit indices were used to evaluate model fit, each offering certain advantages: the root-mean-square–error of approximation (RMSEA; Steiger, 1990), the nonnormed fit index (Bentler, 1990), and the comparative fit index (CFI; Bentler, 1990). The RMSEA accounts for model parsimony when evaluating model fit. Values less than .08 indicate good fit (Browne & Cudeck, 1993). The nonnormed fit index and CFI both measure fit relative to the appropriate longitudinal and/or multiple-group null model (i.e., assuming no relationships or zero correlations among model indicators, no changes or differences in indicator means, and no changes or differences in indicator variances; Widaman & Thompson, 2003) with values above .90 generally considered a good fit and those over .95 considered an excellent fit.

Comparative Tests of Model Fit

To evaluate the comparative model fit, we used the maximum likelihood chi-square statistic to test for factorial similarities and differences across groups or across time in the form of nested-model comparisons. Because the chi-square difference test is overly sensitive to large sample sizes, we took appropriate, conservative measures. For instances in which the reliable structural components were being evaluated and the chi square difference test is appropriate, concerns of excessive power were addressed by adopting a more stringent p value. For omnibus chi square difference tests, we adopted a value of .005 (see Little & Slegers, 2005). For invariance tests, concerns arise over evaluating the invariance of the measurement parameters when a large number of parameter estimates are involved. Because these parameters reflect the fallible aspects of measurement (i.e., the loadings, residuals, intercepts), it is recommended in these instances that model invariance be evaluated with model-based information rather than an omnibus test (Cheung & Rensvold, 2002; Little, 1997). Therefore, when evaluating the tenability of the invariance constraints, we used two recommended criteria: (a) a change in CFI less than .01 and (b) the point estimate of the RMSEA falling within the confidence interval of the preceding model (see Cheung & Rensvold, 2002; Little, Bovaird, & Slegers, 2006).

Progression of Modeling

Our primary goal was to evaluate the factorial structure, construct stability, and intraindividual change patterns among the four PDs (STPD, BPD, AVPD, and OCPD) spanning 10 years of longitudinal data. The first set of analyses tested factorial invariance across the seven time points (baseline or Year 0, Month 6, Year 1, Year 2, Year 4, Year 6, Year 10), with the expected factorial structure specified at each assessment point. In addition, the four-factor model was tested across the seven waves of data separately for male and female participants, to determine whether different models for each sex were warranted. In the next step, factorial invariance for the sample as a whole was tested to discern stability characteristics of the four PDs. This first set of analyses examined correlations among the

disorders within each assessment as well as across the four time points. A second set of analyses specified growth curves to more clearly illustrate the relative intraindividual change patterns of each disorder.

Results

Part 1: Evaluation of Factorial Invariance

Table 2 shows model fit statistics for the progression of tests to examine factorial invariance. As indicated in Table 2, the progression of models showed excellent model fit (see Models 1–3). Moreover, inspection of the residuals and modification indices indicated that no further estimates would improve model fit. Specifically, the criteria for evaluating the steps of factorial invariance (i.e., a change in CFI less than .01 and the point estimate of the RMSEA falling within the confidence interval of the prior model) were well satisfied, indicating strong invariance across time and sex.

Next, we examined whether the correlations among the PD constructs were the same across men and women by testing whether the correlations among the PD constructs at each wave are the same across men and women (see Table 3). The chi-square difference test was significant, $\chi^2(378, N = 733) = 754.34$, $p < .0001$, indicating that there are sufficient differences among the correlations across the constructs for men and women to examine the longitudinal patterns separately by sex. In addition to the correlation differences, the omnibus test for any mean difference on any of these diagnostic constructs was significant, with a nested chi-square difference of $\chi^2(28, N = 733) = 109.74$, $p < .0001$ (see Table 3). Table 4 and Table 5 present the latent correlations, variances, and means for women and men, respectively.

Overall, the correlations among the constructs within each time point were quite discrete at baseline. For example, the highest correlation of .46 was between STPD and BPD in women, indicating that less than 20% of the variance was shared. These correlations do not change appreciably in women at 6 months or 1 year, but by Year 2, the OCPD diagnosis ratings begin to show modest correlations with the other diagnostic categories. In general, these correlations increase over time. The association between STPD and BPD also shows a steady increase for women, with correlations of around .45 during the first 2 years increasing to .73 (over 50% variance overlapping) by Year 10. The pattern for men was generally similar but with most correlations being somewhat smaller than for women. For example, the correlation between STPD and BPD is .35 at baseline (compared with .46 in women) and increases to .60 by Year 10. These longitudinal changes in the strength of the correlations were significant for both men and women (p .0001; see Table 3).

Because the correlations among the constructs were statistically different for women and men, we examined the longitudinal stability relationships separately for men and women. As evidenced in Tables 4 and 5, the cross-time stability correlations of each construct were reasonably high. These correlations are generally .7 or greater when the time span between is around 2 years (and the 4-year span between Year 6 and Year 10 shows similar levels of stability). When estimated as predictive (autoregressive) relationships over time, the indirect stability coefficients (i.e., when one or more time points separate the measurement

occasions) remain quite high for most time points (see Table 6). Note that the levels of stability when separated by the same amount of time are about the same, regardless of the specific time point of assessment. This pattern is consistent with a steady change process. Moreover, the indirect stability over the whole time span of the study (i.e., from baseline to Year 10) remains significant ($p < .0001$). In fact, all of the indirect effects are significant ($p < .0001$), indicating reliable stability in these indirect pathways over the 10 years span of the study.

Part 2: Evaluation of Change Relationships and Participant Characteristics

To examine the change relationships in mean levels, we fit simultaneous latent growth curve models to the four constructs. Fit statistics for the growth curves, which are shown in Table 2 (Model 4), were in the very good range. We then tested the similarities and the differences in the trajectories across men and women (see Table 2 for test results). We found that the proportions of change from time point to time point were functionally the same across men and women for all four constructs, $\chi^2(20, N = 733) = 34.72$, $p = .043$. When we tested whether the magnitude of these mean-level changes were similar or different for women and men, we found that the mean level at baseline of BPD and STPD were different and the mean change in STPD was different (see Table 3). The estimated trajectories for the growth curves are shown in Figure 1A (AVPD and OCPD) and Figure 1B (STPD and BPD, broken down by sex). Overall and as anticipated, the mean level of criteria dropped significantly over time (i.e., in the remitted direction) for all disorders (see Figure 1A and B, and Tables 4 and 5). This drop was most pronounced in the early years (e.g., Year 2 to Year 4) and then tended to level off. The STPD scores showed the least number of mean-level changes but did show sex differences in the change pattern. More pronounced drops in the means levels were found for BPD, AVPD, and OCPD, with the most pronounced drops seen between baseline and Year 2.

Discussion

Our work extends prior efforts to evaluate PD constructs by testing their latent structure longitudinally. In contrast to prior work, the stability of four CLPS *DSM–IV* PD constructs (AVPD, BPD, STPD, and OCPD) was tested in a single longitudinal model at seven measurement points over a 10 year period. Thus, the latent structure of the constructs was examined in the context of longitudinal stability, a key component of the PDs as they are currently defined. Our earlier work (Sanislow et al., 2002) lent some support to the constructs, though cross-sectional tests of the *DSM* structure have been mixed, depending on the samples and methods used (see Sheets & Craighead, 2007). Results from the present longitudinal test provide a very different picture than that seen with cross-sectional snapshots taken of latent structure in prior studies. Notably, the PD constructs become less distinct in this longitudinal context, and the PD constructs are more highly correlated at later time points, relative to the earlier observations. The distinctiveness of the four constructs at baseline, compared with the higher correlation among them 10 years later, suggests poor discriminant validity of enduring PDs. However, the results also support the proposition that a core aspect of personality pathology remains stable over time. It is simply not clear whether the *DSM*–PD constructs best represent a personality pathology that is both enduring

and distinct. Thus, if PDs are to retain the designation of “enduring patterns” (APA, 1996, p. 630) in the *Diagnostic and Statistical Manual of the Mental Disorders* (5th ed.; *DSM–V*), consideration of the increased correlation among the diagnoses is warranted.

Regarding stability, it is important to distinguish the significant statistical stability of the model from clinically meaningful stability. The general trajectories of the latent growth curve models showed a lessening of the constructs (i.e., in the direction of less pathology). The patterning of the indirect effects suggests that the majority of change in the constructs occurred early on, much of it during the first year. All four PD constructs exhibited increased stability in later years (Year 4 to Year 10). However, the mean levels of the constructs were much lower during these later time points and suggest a clinically significant reduction in pathology. Thus, only some aspect of each construct endures. However, due to the heterogeneity of the criteria as well as limitations imposed by the polythetic scoring system (i.e., different combinations of criteria can represent the diagnosis), it is not possible to tease this out with the present approach. We have suggested elsewhere, however, that some aspects of PDs may be more traitlike and enduring, whereas other aspects may be episodic in nature. For instance, affect-related criteria found in BPD are less likely to remit over time than are behavioral criteria (Zanarini et al., 2007) and are more frequently endorsed at later follow-up assessments (McGlashan et al., 2005; see also Sanislow & McGlashan, 1998).

The finding of lower mean levels on the constructs over the long term is interesting in light of the apparent disjunction between PD diagnoses, which appear to be less stable relative to their functional impairment (e.g., Skodol et al., 2002; Skodol, Pagano et al., 2005). Comparing BPD with an Axis II contrast group, Zanarini and colleagues (2005) noted that some improvement in psychosocial functioning was associated with BPD remission status, although vocational deficits were still pronounced (Zanarini, Frankenburg, Hennen, Reich, & Silk, 2005). The persistent low-grade stability shown by the constructs in the present study is consistent with patterns of functional impairment.

Our results may be considered in relation to comorbidity. It has been suggested that high rates of comorbidity reflect core traits shared by different PDs (see Lynam & Widiger, 2001). Perhaps it is those who experience the greatest range of disturbance across constructs who also suffer most enduringly. Such an explanation is consistent with the higher levels of comorbid pathology typical of more disturbed populations and with findings from other studies showing nonremitting BPD cases had greater comorbidity with other Axis II disorders than did those that did remit (Zanarini et al. 2004). Elsewhere, Tyrer and colleagues (Tyrer et al., 2007) have argued that the most severe cases of PD do not manifest as a single disorder, but rather, “personality disturbance extends, ripple-like, across all domains of personality” (p. s55). This assertion in context with our findings suggests that the most severe and chronic disturbance may include core, overlapping traits in which the manifestation of personality dysfunction perpetuates through non-prototypical diagnoses. We have also provided evidence supporting personality trait vulnerability with findings that show that a reduction in negative personality traits based on the five-factor model precedes a reduction in PD criteria (based on the *DSM* criteria; Warner et al., 2004). This suggests that

a stable core, not well captured by the *DSM* constructs, may have predictive value as a vulnerability factor.

These findings may also be viewed in light of the treatment seeking nature of the sample. Although our naturalistic designs precluded drawing conclusions about treatment effects, some reduction may have been the result of the varied (and uncontrolled) treatments that many of the participants received (for details, see Bender et al., 2001, 2007, 2006). There is also the possibility that some participants were misdiagnosed (i.e., over diagnosed) at study entry. If that were the case, Widiger (2005) has suggested that the disorders would then appear to lack discriminant validity. However, we would also expect to see greater differences in indirect effects between baseline diagnoses to Year 10 diagnoses and Month 6 diagnoses to Year 10 diagnoses. Instead, we identified a phenomenon of decreased overall level that is relatively consistent over time.

By the later years (Year 6 to Year 10), it was noteworthy that the OCPD construct, postulated in *DSM* to reside in Cluster C (anxious-fearful), was more correlated with the STPD–Cluster A– based construct (odd–eccentric) and the BPD–Cluster B–based construct (erratic–emotional–dramatic), although showing little overlapping variance with the AVPD–Cluster C–based construct. This finding echoes other reports noting higher co-occurrence of OCPD with Cluster A PDs than with Cluster C PDs (e.g., Blais, McCann, Benedict, & Norman, 1997; Rossi, Marinangeli, Butti, Kalyvoka, & Petruzzi, 2000) and raises interesting possibilities. It may be that there exists a persistent, maladaptive core of the OCPD construct that is less related to the anxious–fearful cluster(C) than to more severe clusters (A and B). Perhaps one component of OCPD is more associated with severe personality pathology, whereas other aspects reflect personality pathology in the anxious–fearful domain. This possibility is supported by recent factor analyses that identify two factors, perfectionism and rigidity, within the OCPD construct (Ansell, Pinto, Edelen, & Grilo, 2008). Clarifying this would help to explain prior inconsistent findings (e.g., Fossati et al., 2000; Hyler & Lyons, 1988; Kass et al., 1985; Livesley et al., 1992, 1998; Morey, 1986; Nestadt et al., 1994; Tyrer & Alexander, 1979; Yang et al., 2002) and might further identify a core personality trait prognostic for more enduring personality pathology. It is interesting to note that other studies have described a loss of the interpersonal control associated with OCPD to be related to explosive outbursts of anger (Villemarette-Pittman, Stanford, Greve, Houston, & Mathias, 2004) and a greater incidence of impulsive aggression relative to normal and noncompulsive PD controls (Stein et al., 1996).

Interesting sex differences were found between the BPD and STPD constructs, but not the AVPD and OCPD constructs. The BPD and STPD sex differences are best illustrated in the growth curves plotted separately for men and women for these two disorders (see Figure 1B). The mean level of the STPD construct was significantly higher in men than in women at baseline. This difference declined through the 10 years of the study, with the gap narrowing to a negligible difference by Year 10. For the BPD construct, the mean level was higher for women compared with men, and this difference persisted through the 10 years of the study. The BPD findings reflect prior-reported findings from our studies testing sex bias in PD diagnosis (e.g., Boggs et al., 2005; Johnson et al., 2003) and are consistent with

findings from other studies showing generally higher levels of BPD symptoms for women (e.g., Jane, Oltmanns, South, & Turkheimer, 2007).

This study has certain strengths and limitations. A cautionary note is that any revisions to the diagnostic system for personality pathology should take into account converging evidence from multiple sources and would ideally be informed by longitudinal studies of non–treatment-participating individuals with clinical levels of disturbance. As discussed above, many of the CLPS participants were receiving various forms of treatment, but because of the naturalistic design of the study, treatments were not controlled, and this precludes examination of treatment effects. Among the strengths of the present study is the large number of minority participants relative to other studies reported in the field. The focus on four PDs is a potential limitation. Participants were recruited with STPD, BPD, OCPD, and AVPD. Results may have been different if we had selected more broadly across all PDs. Even though those four disorders were targeted, participants typically met criteria for several PDs. Thus, this concern is moderated by the range of PD pathology evidenced in the CLPS sample, which was comparable with other clinical samples that used broader selection criteria (e.g., Blashfield et al., 1994; Oldham et al., 1995; Stuart et al., 1998). Nonetheless, results may not generalize to other clinical populations acquired with different selection criteria. Generalization to less disturbed symptomatic volunteers would also not be expected as different results in the latent structure of PDs have been found between clinical and community populations (Sheets & Craighead, 2007).

Our use of parceling is both a strength and limitation. A decided strength is that it increased the reliability of the estimations for the PD constructs (e.g., Little et al., 2002). By reducing measurement error, the diagnostic constructs can be better captured than they would be by simply summing the criteria. This suggests that dimensional approaches in which psychometric properties are carefully considered may better serve to capture PD constructs (see Cuthbert, 2005). However, the parceling approach does preclude an examination of the strength and ordering of the relationships of individual PD criteria to their presumed constructs. This limits our ability to draw conclusions about matters such as potential differential stability of certain criteria as noted above.

It is also not possible to completely characterize the apparent overlap that we might term *construct comorbidity* in the context of the present study. The overlap, of course, could be due to a variety of factors, including criterion overlap, related traits, or undifferentiated pathology in chronically disturbed individuals. These questions are of interest for future work. From the present study, it is clear that some central core of personality pathology evident from *DSM* constructs does endure; yet, the distinctiveness of the diagnostic categories does not. Thus, there does appear to be a problem with the *DSM* PDs in their present framework in that the most stable and enduring personality pathology does not retain the distinct qualities of the PD constructs. Clarifying the enduring qualities of personality pathology is an important consideration for the *DSM–V*.

References

American Psychiatric Association. (1996). Diagnostic and statistical manual of mental disorders (4th ed.). Washington, DC: Author.

Ansell EB, Pinto A, Edelen MO, & Grilo CM (2008). Obsessive-compulsive personality disorder in patients with binge eating disorder. Canadian Journal of Psychiatry, 53, 863–897. [PubMed: 19087485]

Bell RC, & Jackson HJ (1992). The structure of personality disorders in *DSM–III*. Acta Psychiatrica Scandanavica, 85, 279–287.

Bender DS, Dolan RT, Skodol AE, Sanislow CA, Dyck IR, McGlashan TH, et al. (2001). Treatment utilization by patients with personality disorders. American Journal of Psychiatry, 158, 295–302. [PubMed: 11156814]

Bender DS, Skodol AE, Dyck IR, Markowitz JC, Shea MT, Yen S, et al. (2007). Ethnicity and mental health treatment utilization by patients with personality disorders. Journal of Consulting and Clinical Psychology, 75, 992–999. [PubMed: 18085915]

Bender DS, Skodol AE, Pagano ME, Dyck IR, Grilo CM, Shea MT, et al. (2006). Prospective assessment of treatment use by patients with personality disorders. Psychiatric Services, 57, 254–257. [PubMed: 16452705]

Bentler PM (1990). Comparative fix indices in structural models. Psychological Bulletin, 107, 238–246. [PubMed: 2320703]

Blais MA, McCann JT, Benedict KB, & Norman DK (1997). Toward an empirical/theoretical grouping of the *DSM–III–R* personality disorders. Journal of Personality Disorders, 11, 191–198. [PubMed: 9203113]

Blashfield RK, McElroy RA, Pfohl B, & Blum N (1994). Comorbidity and the prototype model. Clinical Psychology: Science and Practice, 1, 96–99.

Boggs CD, Morey LC, Skodol AE, Sanislow CA, Grilo CM, McGlashan TH, et al. (2005). Differential impairment as an indicator of sex bias in *DSM–IV* criteria for four personality disorders. Psychological Assessment, 17, 492–496. [PubMed: 16393017]

Browne M, & Cudeck R (1993). Alternative ways of assessing model fit In Bollen KA & Long JS (Eds.), Testing structural equation models (pp. 136–162). Newbury Park, CA: Sage.

Cheung GW, & Rensvold RB (2002). Evaluating goodness-of-fit indices for testing measurement invariance. Structural Equation Modeling, 9, 233–255.

Clifton A, & Pilkonis PA (2007). Evidence for a single latent class of *Diagnostic and statistical manual of mental disorders* borderline personality pathology. Comprehensive Psychiatry, 48, 70–78. [PubMed: 17145285]

Cohen J (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20, 37–46.

Cuthbert BN (2005). Dimensional models of psychopathology: Research agenda and clinical utility. Journal of Abnormal Psychology, 114, 565–569. [PubMed: 16351380]

Durrett C, & Westen D (2005). The structure of axis II disorders in adolescents: A cluster and factor-analytic investigation of *DSM–IV* categories and criteria. Journal of Personality Disorders, 19, 440–461. [PubMed: 16178684]

Enders CK (in press). Applied missing data analysis. New York: Guilford Press.

First MB, Spitzer RL, Gibbon M, & Williams JBW (1996). Structured clinical interview for *DSM–IV* axis II personality disorders (SCID-II), Version 2.0. New York, Biometrics Research Department, New York State Psychiatric Institute.

Fossati A, Maffei C, Bagnato M, Battaglia M, Donati D, Donini M, et al. (2000). Patterns of covariation of *DSM–IV* personality disorders in a mixed psychiatric sample. Comprehensive Psychiatry, 41, 206–215. [PubMed: 10834630]

Grilo CM, Shea MT, Sanislow CA, Skodol AE, Gunderson JG, Stout RL, et al. (2004). Two-year stability and change in schizotypal, borderline, avoidant and obsessive-compulsive personality disorders. Journal of Consulting and Clinical Psychology, 72, 767–775. [PubMed: 15482035]

Gunderson JG, Shea MT, Skodol AE, McGlashan TH, Morey LC, Stout RL, et al. (2000). The collaborative longitudinal personality disorders study: Development, aims, design, and sample characteristics. Journal of Personality Disorders, 14, 300–315. [PubMed: 11213788]

Hoyle RH, & Smith GT (1994). Formulating clinical research questions as structural equation models: A conceptual overview. Journal of Consulting and Clinical Psychology, 62, 429–440. [PubMed: 8063970]

Hyler SE, & Lyons M (1988). Factor analysis of the *DSM–III* personality disorder clusters: A replication. Comprehensive Psychiatry, 29, 304–308. [PubMed: 3378417]

Jane JS, Oltmanns TF, South SC, & Turkheimer E (2007). Gender bias in diagnostic criteria for personality disorders: An item response theory analysis. Journal of Abnormal Psychology, 116, 166–175. [PubMed: 17324027]

Johnson DM, Shea MT, Yen S, Battle C, Zlotnick C, Sanislow CA, et al. (2003). Gender differences in borderline personality disorder: Findings from the collaborative longitudinal personality disorders study. Comprehensive Psychiatry, 44, 284–292. [PubMed: 12923706]

Jo¨reskog D, & So¨rbom D. (2008). LISREL 8: User's reference guide. Chicago: Scientific Software International.

Karterud JM, Pedersen G, Gude T, & Falkum E (2004). An investigation of the prototype validity of the borderline *DSM–IV* construct. Acta Psychiatrica Scandavica, 109, 289–298.

Kass F, Skodol AE, Charles E, Spitzer RL, & Williams JB (1985). Scaled ratings of *DSM–III* personality disorders. American Journal of Psychiatry, 142, 627–630. [PubMed: 3985201]

Kishton JM, & Widaman KF (1994). Unidimensional versus domain representative parceling of questionnaire items: An empirical example. Educational and Psychological Measurement, 54, 757–765.

Laptook RS, Klein DN, & Dougherty LR (2006). Ten-year stability of depressive personality disorder in depressed outpatients. American Journal of Psychiatry, 163, 865–871. [PubMed: 16648328]

Lenzenweger MF, Johnson MD, & Willett JB (2004). Individual growth curve analysis illuminates stability and change in personality disorder features. Archives of General Psychiatry, 61, 1015–1024. [PubMed: 15466675]

Little TD (1997). Mean and covariance structures (MACS) analyses of cross-cultural data: Practical and theoretical issues. Multivariate Behavioral Research, 32, 53–76. [PubMed: 26751106]

Little TD, Bovaird JA, & Slegers DW (2006). Methods for the analysis of change In Mroczek DK & Little TD (Eds.), Handbook of personality development (pp. 181–211). Mahwah, NJ: Erlbaum.

Little TD, Cunningham WA, Shahar G, & Widaman KF (2002). To parcel or not to parcel: Exploring the question, weighing the merits. Structural Equation Modeling, 9, 151–173.

Little TD, & Slegers DW (2005). Factor analysis: Multiple groups In Everitt B, Howell D (Eds.), & Rindskopf D (Section Ed.), Encyclopedia of statistics in behavioral science (Vol. 2, pp. 617–623). Chichester, England: Wiley.

Livesley WJ, Jackson D, & Schroeder ML (1992). Factorial structure of traits delineating PDs in clinical and general population samples. Journal of Abnormal Psychology, 101, 432–440. [PubMed: 1500600]

Livesley WL, Jang KL, & Vernon PA (1998). Phenotypic and genetic structure of traits delineating personality disorder. Archives of General Psychiatry, 55, 672–675.

Lynam DR, & Widiger TA (2001). Using the five-factor model to represent the *DSM–IV* personality disorders: An expert consensus approach. Journal of Abnormal Psychology, 110, 401–412. [PubMed: 11502083]

McGlashan TH, Grilo CM, Sanislow CA, Ralevski E, Morey LC, Gunderson JG, et al. (2005). Two-year prevalence and stability of individual criteria for schizotypal, borderline, avoidant, and obsessive-compulsive personality disorders. American Journal of Psychiatry, 162, 883–889. [PubMed: 15863789]

McGlashan TH, Grilo CM, Skodol AE, Gunderson JG, Shea MT, Morey LC, et al. (2000). The collaborative longitudinal personality disorders study: Baseline axis I/II and II/II diagnostic co-occurrence. Acta Psychiatrica Scandavica, 102, 256–264.

Moldin SO, Rice JP, Erlenmeyer-Kimling L, & Squires-Wheeler E (1994). Latent structure of *DSM–III–R* axis II psychopathology in a normal sample. Journal of Abnormal Psychology, 103, 259–266. [PubMed: 8040495]

Morey LC (1986). A comparison of three personality disorder approaches. Journal of Psychopathology and Behavioral Assessment, 8, 25–30.

Morey LC, Hopwood CJ, Gunderson JG, Zanarini MC, Skodol AE, Shea MT, et al. (2007). Comparison of diagnostic models for personality disorders. Psychological Medicine, 37, 983–994. [PubMed: 17121690]

Nestadt G, Eaton WW, Romanoski AJ, Garrison R Folstein MF,& McHugh PR (1994). Assessment of *DSM–III* personality structure in a general population survey. Comprehensive Psychiatry, 35, 54–63. [PubMed: 8149730]

O'Connor BP (2005). A search for consensus on the dimensional structure of personality disorders. Journal of Clinical Psychology, 61, 323–345. [PubMed: 15468325]

O'Connor BP, & Dyce JA (1998). A test of models of personality disorder configuration. Journal of Abnormal Psychology, 107, 3–16. [PubMed: 9505034]

Oldham JM, Skodol AE, Kellman HD, Hyler SE, Doidge N, Rosnick L, & Gallaher PE (1995). Comorbidity of axis I and axis II disorders. American Journal of Psychiatry, 152, 571–578. [PubMed: 7694906]

Rodebaugh TL, Chambless DL, Renneberg B, & Fydrich T (2005). The factor structure of the *DSM–III–R* personality disorders: An evaluation of the competing models. International Journal of Methods of Psychiatric Research, 14, 43–55.

Rossi A, Marinangeli MG, Butti G, Kalyvoka A, & Petruzzi C (2000). Pattern of comorbidity among anxious and odd personality disorders: The case of obsessive-compulsive personality disorder. CNS Spectrums, 5, 23–26.

Sanislow CA, & McGlashan TH (1998). Treatment outcome of personality disorders. Canadian Journal of Psychiatry, 43, 237–250. [PubMed: 9561312]

Sanislow CA, Morey LC, Grilo CM, Gunderson JG, Shea MT, Skodol AE, et al. (2002). Confirmatory factor analysis of *DSM–IV* borderline, schizotypal, avoidant and obsessive-compulsive personality disorders: Findings from the collaborative longitudinal personality disorders study. Acta Psychiatrica Scandavica, 105, 28–36.

Schafer JL, & Graham JW (2002). Missing data: Our view of state of the art. Psychological Methods, 7, 147–177. [PubMed: 12090408]

Shea MT, Stout RL, Gunderson JG, Morey LC, Grilo CM, McGlashan TH, et al. (2002). Short-term diagnostic stability of schizotypal, borderline, avoidant, and obsessive-compulsive personality disorders. American Journal of Psychiatry, 159, 2036–2041. [PubMed: 12450953]

Sheets E, & Craighead WE (2007). Toward an empirically based classification of personality pathology. Clinical Psychology: Science and Practice, 14, 77–93.

Skodol AE, Gunderson JG, McGlashan TH, Dyck IR, Bender DS, Grilo CM, et al. (2002). Functional impairment in schizotypal, borderline, avoidant, and obsessive-compulsive personality disorders. American Journal of Psychiatry, 159, 276–283. [PubMed: 11823271]

Skodol AE, Oldham JM, Bender DS, Dyck IR, Stout RL, Morey LC, et al. (2005). Dimensional representations of *DSM–IV* personality disorders: Relationships to functional impairment. American Journal of Psychiatry, 162, 1919–1925. [PubMed: 16199839]

Skodol AE, Pagano ME, Bender DS, Shea MT, Gunderson JG, Yen S, et al. (2005). Stability of functional impairment in patients with schizotypal, borderline, avoidant, or obsessive-compulsive personality disorder over two years. Psychological Medicine, 35, 443–451. [PubMed: 15841879]

Steiger JH (1990). Test for comparing elements of a correlation matrix. Psychological Bulletin, 87, 245–251.

Stein DJ, Trestman RL, Mitropoulou V, Cocarro EF, Hollander E, & Siever LJ (1996). Impulsivity and serotonergic function in compulsive personality disorder. Journal of Neuropsychiatry and Clinical Neurosciences, 8, 393–398. [PubMed: 9116474]

Stuart S, Pfohl B, Battaglia M, Bellodi L, Grove W, & Cadoret R (1998). The co-occurrence of *DSM–III–R* personality disorders. Journal of Personality Disorders, 12, 302–315. [PubMed: 9891285]

Tyrer P, & Alexander J (1979). Classification of personality disorder. British Journal of Psychiatry, 135, 163–167. [PubMed: 497619]

Tyrer P, Coombs N, Ibrahimi F, Mathilakath A, Bajaj P, Ranger M, et al. (2007). Critical developments in the assessment of personality disorder. British Journal of Psychiatry, 49, s51–s59. [PubMed: 17470943]

Villemarette-Pittman NR, Stanford MS, Greve KW, Houston RJ,& Mathias CW (2004). Obsessive-compulsive personality disorder and behavioral disinhibition. Journal of Psychology, 138, 5–22. [PubMed: 15098711]

Warner MB, Morey LC, Finch JF, Gunderson JG, Skodol AE, Sanislow CA, et al. (2004). The longitudinal relationship of personality traits and disorders. Journal of Abnormal Psychology, 113, 217–227. [PubMed: 15122942]

Widaman KF, & Thompson JS (2003). On specifying the null model for incremental fit indices in structural equation modeling. Psychological Methods, 8, 16–37. [PubMed: 12741671]

Widiger TA (2005). CIC, CLPS, and MSAD. Journal of Personality Disorders, 19, 586–593. [PubMed: 16274288]

Yang J, Bagby M, Costa PT, Ryder AG, & Herbst JH (2002). Assessing the *DSM–IV* structure of personality disorder with a sample of Chinese psychiatric patients. Journal of Personality Disorders, 16, 317–331. [PubMed: 12224125]

Zanarini MC, Frankenburg FR, Hennen J, Reich B, & Silk KR (2005). Psychosocial functioning of borderline patients and axis II comparison subjects followed prospectively for six years. Journal of Personality Disorders, 19, 19–29. [PubMed: 15899718]

Zanarini MC, Frankenburg FR, Hennen J, & Silk KR (2003). The longitudinal course of borderline psychopathology: 6-year prospective follow-up of the phenomenology of borderline personality disorder. American Journal of Psychiatry, 160, 274–283. [PubMed: 12562573]

Zanarini MC, Frankenburg FR, Reich DB, Silk KR, Hudson JL, & McSweeney LB (2007). The subsyndromal phenomenology of borderline personality disorder: A 10-year follow up study. American Journal of Psychiatry, 164, 929–935. [PubMed: 17541053]

Zanarini MC, Frankenburg FR, Sickel AE, & Yong L (1996). Diagnostic interview for DSM–IV personality disorders. Boston, MA: Laboratory for the Study of Adult Development, McLean Hospital and Department of Psychiatry, Harvard University.

Zanarini MC, Frankenburg FR, Vujanovic AA, Hennen J, Reich DB, & Silk KR (2004). Axis II comorbidity of borderline personality disorder: Description of 6-year course and prediction to time-to-remission. Acta Psychiatrica Scandanavica, 110, 416–420.

Zanarini MC, & Shea MT (1996). The Diagnostic Interview for DSM–IV Personality Disorders—Follow-Along Version (DIPD-IV-FAV). Belmont, MA: McLean Hospital.

Zanarini MC, Skodol AE, Bender DS, Dolan RT, Sanislow CA, Schaefer E, et al. (2000). The collaborative longitudinal personality disorders study: Reliability of axis I and II diagnoses. Journal of Personality Disorders, 14, 291–299. [PubMed: 11213787]

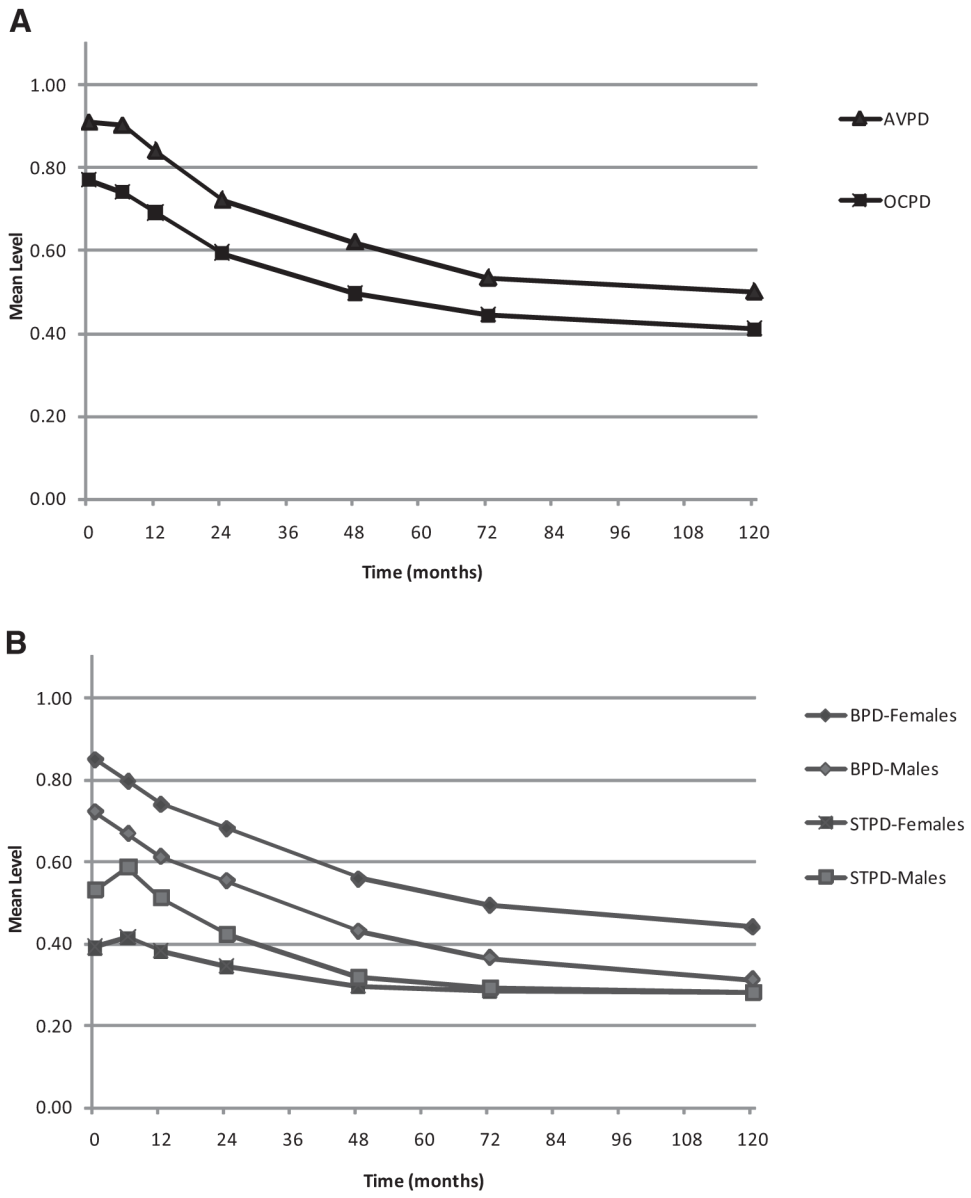

Figure 1.

A: Growth curve models of the four avoidant personality disorders (AVPD) and obsessive-compulsive personality disorders (OCPD) from baseline through Year 10. (Because of significant gender differences found for schizotypal personality disorder (STPD) and borderline personality disorder (BPD), growth curves are displayed separately for those disorders in Figure 1B.) Mean level reflects the range implied by the diagnosis (e.g., 0 = *not present*, 1 = *subclinical*, 2 = *clinical and significant*) for each construct. B: Growth curve models of the BPD and STPD broken down by sex through Year 10. Mean level reflects the range implied by the diagnosis (e.g., 0 = *not present*, 1 = *subclinical*, 2 = *clinical and significant*) for each construct.

Table 1

Parcels

Parcel	DIPD-IV criteria
STPD Parcel 1	Suspiciousness
STPD Parcel 1	Unusual perceptions
STPD Parcel 1	Odd behavior
STPD Parcel 2	Social anxiety
STPD Parcel 2	Odd thinking
STPD Parcel 2	Ideas of reference
STPD Parcel 3	Odd beliefs
STPD Parcel 3	Inappropriate affect
STPD Parcel 3	No close friends
BPD Parcel 1	Unstable relationships
BPD Parcel 1	Affective instability
BPD Parcel 1	Transient dissociation
BPD Parcel 2	Intense anger
BPD Parcel 2	Identity disturbance
BPD Parcel 2	Frequent suicidal behavior
BPD Parcel 3	Avoid abandonment
BPD Parcel 3	Impulsivity
BPD Parcel 3	Chronic emptiness
AVPD Parcel 1	Preoccupied with rejection
AVPD Parcel 1	Feels socially inept
AVPD Parcel 1	Avoids occupational activities
AVPD Parcel 2	Inhibited in interpersonal situations
AVPD Parcel 2	Reluctant to take risks
AVPD Parcel 3	Unwilling to get involved unless liked
AVPD Parcel 3	Shows restraint in relationships
OCPD Parcel 1	Reluctant to delegate
OCPD Parcel 1	Perfectionism
OCPD Parcel 2	Stubbornness
OCPD Parcel 2	Morality
OCPD Parcel 2	Workaholic
OCPD Parcel 3	Packrat
OCPD Parcel 3	Miserliness
OCPD Parcel 3	Detail-oriented

Note. Parcel scores are the average of the listed criteria. The same parcels were computed for each time point. DIPD-IV = Diagnostic Interview for *DSM–IV* = Personality Disorders; STPD = schizotypal personality disorder ciagnosis; BPD = borderline personality disorder diagnosis; AVPD = avoidant personality disorder diagnosis; OCPD = obsessive-compulsive personality disorder diagnosis; *DSM–IV = Diagnostic and Statistical Manual of Mental Disorders* (4th ed.).

Table 2

Model Fit Statistics

Model number	Model description	Chi square	*df*	RMSEA	Lower 90% confidence interval	Upper 90% confidence interval	Nonnormed fit index	Comparative fit index
1	Configural invariance (no constraints) over time or across sex	6,256.51	5,544	.019	.016	.021	.995	.996
2	Weak invariance (loadings invariant) across sex and time	6,422.41	5,648	.019	.017	.022	.995	.996
3	Strong invariance (loadings and intercepts) across sex and time	6,670.36	5,752	.021	.018	.023	.994	.995
4	Fit of the growth curve model for seven time points across men and women	1,535.55	760	.051	.047	.055	.968	.978

Note. RMSEA = root-mean-square error of approximation; *df* = degree of freedom.

Table 3

Results of the Chi-Square Difference Tests

Comparison	Chi-square difference	*df* difference	*p*
Test of latent correlations being equal across men & women in CFA model	754.34	378	<.0001
Test of latent means being equal across men & women in CFA model	109.74	28	<.0001
Test of within time correlations among four PDs being same over time: Women	150.891	36	<.0001
Test of within time correlations among four PDs being same over time: Men	102.413	36	<.0001
Test of basis weights being equal across males & females in LGC model	34.72	20	.043
Test of mean intercept difference in BPD (LGC model)	19.77	1	<.001
Test of mean slope difference in BPD (LGC model)	6.69	1	.010
Test of mean intercept difference in STPD (LGC model)	14.99	1	<.001
Test of mean slope difference in STPD (LGC model)	18.92	1	<.001
Test of mean intercept difference in AVPD (LGC model)	0.31	1	.578
Test of mean slope difference in AVPD (LGC model)	2.21	1	.137
Test of mean intercept difference in OCPD (LGC model)	0.01	1	.920
Test of mean slope difference in OCPD (LGC model)	1.13	1	.288

Note. CFA = confirmatory factor analysis; PD = personality disorder; BPD = borderline personality disorder; STPD = schizotypal personality disorder; LGC = latent growth curve; AVPD = avoidant personality disorder; OCPD = obsessive-compulsive personality disorder; *df* = degree of freedom.

Table 4.

Latent Correlations, Variances, and Means for Women

	Baseline				Month 6				Year 1				Year 2				Year 4				Year 6				Year 10			
Measure	**ST**	**BP**	**AV**	**OC**	**ST**	**BP**	**AV**	**OC**	**ST**	**BP**	**AV**	**OC**	**ST**	**BP**	**AV**	**OC**	**ST**	**BP**	**AV**	**OC**	**ST**	**BP**	**AV**	**OC**	**ST**	**BP**	**AV**	**OC**
Baseline																												
ST	—																											
BP	46	—																										
AV	37	37	—																									
OC	14	09	03	—																								
Month 6																												
ST	85	31	25	12	—																							
BP	40	85	36	04	42	—																						
AV	33	36	82	05	36	45	—																					
OC	05	01	08	86	10	09	14	—																				
Year 1																												
ST	81	30	25	12	91	37	34	11	—																			
BP	43	77	33	04	41	92	43	09	46	—																		
AV	32	32	79	06	32	39	90	15	37	46	—																	
OC	08	01	05	80	12	07	10	92	19	13	17	—																
Year 2																												
ST	73	23	26	12	83	31	36	08	85	36	41	15	—															
BP	38	70	34	11	34	78	43	12	40	82	44	19	45	—														
AV	30	28	70	06	28	40	79	12	31	40	82	16	44	49	—													
OC	11	05	03	70	19	10	09	74	22	11	13	78	29	33	22	—												
Year 4																												
ST	64	24	23	09	72	34	33	11	74	38	35	20	80	44	35	26	—											
BP	38	56	23	08	38	67	33	12	44	73	34	17	41	77	35	24	63	—										
AV	33	29	63	03	29	38	67	14	36	45	75	20	39	51	76	19	53	58	—									
OC	18	16	13	56	23	22	14	64	26	23	16	61	28	34	24	75	36	49	30	—								
Year 6																												
ST	61	29	29	14	67	36	37	17	75	41	41	23	77	46	38	26	85	59	49	40	—							

	Baseline				Month 6				Year 1				Year 2				Year 4				Year 6				Year 10			
Measure	**ST**	**BP**	**AV**	**OC**	**ST**	**BP**	**AV**	**OC**	**ST**	**BP**	**AV**	**OC**	**ST**	**BP**	**AV**	**OC**	**ST**	**BP**	**AV**	**OC**	**ST**	**BP**	**AV**	**OC**	**ST**	**BP**	**AV**	**OC**
BP	32	46	23	14	32	56	28	19	37	61	30	25	37	66	34	26	47	77	47	44	70	—						
AV	24	26	59	08	22	36	62	17	28	39	66	22	32	45	70	19	38	43	78	25	51	56	—					
OC	12	13	06	52	17	19	11	59	23	23	14	61	23	27	19	66	26	36	22	83	43	51	29	—				
Year 10																												
ST	55	32	31	15	64	39	35	13	72	45	45	19	69	48	37	30	76	61	51	42	84	64	44	39	—			
BP	36	42	24	13	31	51	27	18	36	54	30	22	33	58	30	27	45	65	41	40	53	72	34	36	73	—		
AV	22	26	47	07	23	33	53	12	29	40	62	21	30	44	64	23	31	42	71	30	46	54	80	30	57	47	—	
OC	13	18	09	42	20	26	13	50	30	31	16	50	24	30	21	57	28	35	23	70	38	46	29	81	50	45	38	—
Means	47	1.0	1.1	89	41	79	90	72	35	71	81	63	32	73	73	60	29	58	64	49	29	50	53	45	27	39	49	38
Variances	16	32	41	23	19	32	42	23	15	29	42	20	11	29	36	19	11	25	35	16	09	22	39	16	09	16	37	16

Note. $N = 467$. Estimates are based on 100 imputations of the missing data. Note also that the square of the correlations indicates the shared variance among constructs and that these values are estimates of the reliable variance (i.e., with measurement error removed). For example, a correlation of .71, indicates that 50% of the reliable variance overlaps between constructs. ST = schizotypal personality disorder; BP = borderline personality disorder; AV = avoidant personality disorder; OC = obsessive-compulsive personality disorder.

J Abnorm Psychol. Author manuscript; available in PMC 2019 April 02.

Table 5.

Latent Correlations, Variances, and Means for Men

	Baseline				Month 6				Year 1				Year 2				Year 4				Year 6				Year 10			
Measure	**ST**	**BP**	**AV**	**OC**	**ST**	**BP**	**AV**	**OC**	**ST**	**BP**	**AV**	**OC**	**ST**	**BP**	**AV**	**OC**	**ST**	**BP**	**AV**	**OC**	**ST**	**BP**	**AV**	**OC**	**ST**	**BP**	**AV**	**OC**
Baseline																												
ST	—																											
BP	.35	—																										
AV	.29	.32	—																									
OC	.19	.23	.09	—																								
Month 6																												
ST	.90	.36	.21	.04	—																							
BP	.26	.83	.20	.18	.41	—																						
AV	.29	.21	.86	.05	.30	.24	—																					
OC	.06	.06	.06	.87	.03	.17	.08	—																				
Year 1																												
ST	.84	.24	.19	.05	.91	.29	.26	.02	—																			
BP	.32	.75	.23	.19	.44	.83	.26	.12	.44	—																		
AV	.28	.16	.79	.06	.29	.25	.93	.09	.29	.34	—																	
OC	.15	.07	.13	.79	.15	.13	.09	.89	.15	.20	.17	—																
Year 2																												
ST	.75	.26	.24	.01	.82	.31	.35	.05	.83	.39	.40	.18	—															
BP	.30	.64	.22	.16	.34	.64	.24	.10	.34	.70	.29	.21	.44	—														
AV	.22	.24	.69	.07	.23	.28	.78	.09	.21	.29	.81	.14	.38	.43	—													
OC	.15	.12	.10	.54	.14	.21	.11	.66	.11	.19	.13	.73	.29	.43	.23	—												
Year 4																												
ST	.70	.36	.23	.05	.72	.38	.33	.07	.68	.44	.32	.14	.70	.47	.35	.18	—											
BP	.27	.56	.23	.16	.29	.57	.22	.15	.28	.66	.24	.16	.23	.67	.34	.21	.62	—										
AV	.27	.31	.60	.09	.31	.35	.67	.12	.28	.40	.71	.17	.36	.47	.75	.21	.59	.57	—									
OC	.19	.14	.14	.51	.17	.21	.18	.59	.19	.25	.14	.59	.22	.35	.24	.71	.36	.49	.34	—								
Year 6																												
ST	.57	.34	.32	.08	.63	.35	.40	.13	.61	.37	.37	.22	.63	.39	.34	.19	.81	.51	.49	.38	—							
BP	.17	.53	.23	.15	.22	.55	.22	.16	.22	.57	.21	.21	.20	.63	.28	.18	.41	.72	.44	.37	.65	—						
AV	.21	.25	.57	.05	.25	.29	.63	.08	.26	.31	.61	.17	.29	.36	.66	.09	.42	.39	.73	.22	.55	.56	—					
OC	.18	.17	.09	.37	.19	.15	.11	.52	.20	.20	.09	.55	.19	.29	.17	.54	.25	.31	.25	.75	.47	.45	.24	—				

	Baseline				Month 6				Year 1				Year 2				Year 4				Year 6				Year 10			
Measure	**ST**	**BP**	**AV**	**OC**	**ST**	**BP**	**AV**	**OC**	**ST**	**BP**	**AV**	**OC**	**ST**	**BP**	**AV**	**OC**	**ST**	**BP**	**AV**	**OC**	**ST**	**BP**	**AV**	**OC**	**ST**	**BP**	**AV**	**OC**
Year 10																												
ST	.54	.28	.27	.09	.58	.29	.30	.11	.55	.29	.37	.13	.53	.37	.32	.26	.66	.47	.40	.26	.72	.46	.35	.35	—			
BP	.16	.38	.17	.27	.11	.38	.12	.21	.09	.40	.13	.18	.07	.47	.16	.26	.29	.57	.25	.31	.31	.57	.14	.26	.60	—		
AV	.13	.19	.49	.04	.19	.26	.59	.04	.18	.30	.64	.10	.24	.30	.68	.17	.26	.31	.68	.20	.43	.41	.76	.20	.46	.29	—	
OC	.10	.09	.04	.33	.14	.16	.05	.47	.20	.18	.07	.46	.16	.25	.16	.52	.23	.29	.22	.64	.32	.33	.18	.82	.41	.37	.28	—
Means	0.62	0.85	1.1	0.91	0.58	0.62	0.87	0.71	0.47	0.54	0.79	0.65	0.46	0.55	0.74	0.61	0.33	0.42	0.65	0.53	0.31	0.41	0.55	0.48	0.29	0.35	0.54	0.43
Variances	0.24	0.25	0.36	0.23	0.25	0.19	0.36	0.22	0.23	0.19	0.34	0.20	0.18	0.18	0.35	0.20	0.08	0.16	0.35	0.16	0.08	0.15	0.36	0.18	0.07	0.14	0.35	0.15

Note. Note that the square of the correlations indicates the shared variance among constructs and that these values are estimates of the reliable variance (i.e., with measurement error removed). For example, a correlation of .71 indicates that 50% of the reliable variance overlaps between constructs. ST = schizotypal personality disorder; BP = borderline personality disorder; AV = avoidant personality disorder; OC = obsessive-compulsive personality disorder.

Table 6.

Direct and Indirect Effects Over Time for Women and Men

	Women				Men			
Measure	**STPD**	**BPD**	**AVPD**	**OCPD**	**STPD**	**BPD**	**AVPD**	**OCPD**
Baseline to Month 6	.85	.86	.84	.87	.91	.86	.87	.86
Baseline to Year 1	.79	.79	.77	.81	.84	.74	.81	.78
Baseline to Year 2	.69	.67	.65	.64	.72	.55	.67	.56
Baseline to Year 4	.56	.51	.52	.49	.54	.37	.52	.41
Baseline to Year 6	.49	.39	.41	.40	.46	.28	.40	.31
Baseline to Year 10	.41	.27	.33	.32	.35	.16	.32	.25
Month 6 to Year 1	.93	.92	.91	.93	.93	.86	.93	.90
Month 6 to Year 2	.81	.78	.77	.74	.79	.64	.78	.65
Month 6 to Year 4	.66	.59	.61	.57	.60	.43	.60	.48
Month 6 to Year 6	.58	.45	.48	.46	.51	.32	.46	.36
Month 6 to Year 10	.49	.31	.39	.37	.39	.19	.37	.29
Year 1 to Year 2	.88	.85	.85	.79	.86	.75	.83	.72
Year 1 to Year 4	.71	.64	.67	.61	.64	.50	.65	.53
Year 1 to Year 6	.63	.49	.53	.50	.54	.37	.50	.40
Year 1 to Year 10	.52	.34	.43	.40	.42	.22	.39	.33
Year 2 to Year 4	.81	.76	.79	.77	.75	.67	.78	.73
Year 2 to Year 6	.71	.58	.63	.63	.64	.50	.60	.55
Year 2 to Year 10	.60	.40	.51	.50	.49	.29	.47	.45
Year 4 to Year 6	.88	.77	.79	.82	.85	.75	.77	.75
Year 4 to Year 10	.74	.53	.64	.65	.65	.43	.61	.61
Year 6 to Year 10	.84	.69	.81	.80	.76	.58	.79	.82

Note. Standardized effects reflect the stability influences over time. Estimates are based on 100 imputations. These effects reflect the stability of the individual differences in diagnosis over the various time spans as either direct (when at adjacent measurement occasions) or indirect (when separated by 1 or more measurement occasions) effects. Because all of these effects are all significant at $p < .0001$, there is reliable stability in these pathways even over the 10-year span of the study. STPD = schizotypal personality disorder; BPD = borderline personality disorder; AVPD = avoidant personality disorder; OCPD = obsessive-compulsive personality disorder.