Stability and Change in Personality Disorder Features

The Longitudinal Study of Personality Disorders

Mark F. Lenzenweger, PhD

Background: There exists no empirical literature documenting the long-term longitudinal stability of personality pathology comparable to that available for normal personality. A number of test-retest studies have usefully established the short-term reliability of Axis II measures. However, the test-retest design is methodologically inadequate for resolving issues related to the long-term stability of personality disorder (PD). This prospective longitudinal study evaluated the stability of PD features in multiwave perspective.

Methods: Subjects (N = 250) drawn from a nonclinical university population were examined for PD features at 3 different time points using the International Personality Disorders Examination (IPDE) and the Millon Clinical Multiaxial Inventory II (MCMI-II) during a study period of 4 years.

Results: Features of PD displayed considerable evidence of stability for individual differences and group means, at the dimensional level of analysis, on both the IPDE and the MCMI-II. Both measures revealed modest declines in PD features over time; however, the observed changes were associated with relatively small effect sizes.

Conclusion: Features of PD, viewed from a dimensional perspective, seem to be relatively stable in terms of individual differences and group means based on both clinical interview and self-administered PD assessments.

Arch Gen Psychiatry. 1999;56:1009-1015

RESULTS

AXIS I CHARACTERISTICS OF THE SAMPLE

The lifetime DSM-III-R Axis I diagnoses (Table 2) of the study subjects are for...
SUBJECTS AND METHODS

SUBJECTS

The 258 subjects in the LSPD were drawn from an initial sample that consisted of 2000 first-year undergraduate students at Cornell University, Ithaca, NY.21 Subjects were assigned to either a possible personality disorder (PPD) or no personality disorder (NPD) group as determined by the International Personality Disorder Examination (IPDE) Screen (IPDE-S) (completed screens = 1684, response rate = 84.2%). The PPD subjects had to meet the diagnostic threshold for at least 1 specific DSM-III-R Axis II PD, whereas NPD subjects did not meet the DSM-III-R-defined threshold for diagnosis and had fewer than 10 PD features across all disorders. Extensive details concerning subject selection procedure and sampling are given elsewhere.22 The 258 subjects consisted of 121 men (47%) and 137 women (53%); 134 (66 women) in the NPD group and 124 (71 women) in the PPD group. All subjects gave voluntary written informed consent and received an honorarium of $50 at each wave. Of the initial 258 subjects, 250 completed all 3 assessment waves. Of the 8 subjects who did not complete all waves, 5 were in the PPD and 3 were in the NPD group; the proportions did not differ (goodness of fit of $\chi^2 = 0.50, P > .47$). Of these 8 subjects, 6 transferred to other colleges and 2 died in motor vehicle crashes. This report concerns the 250 subjects completing all 3 assessment waves.

Compared with the US population, the sample over-represents Asian/Pacific Islanders and under-represents those of Hispanic, African American, and working-class origins (Table 1).

PD ASSESSMENTS

The LSPD has a prospective multiwave panel design8 in which subjects are evaluated at 3 points in time (ie, freshman, sophomore, and senior years in college). All interview assessments were conducted by experienced clinical psychologists (PhD level) or psychiatric social workers. The masters-level social workers had an average of 16 years of clinical experience. Subjects also completed a large battery of self-administered assessments of normal personality, temperament, and other psychological factors. Finally, as the LSPD is a naturalistic prospective study, subjects were free to seek psychological treatment of their own accord.

IPDE DSM-III-R Screen

This is a 250-item self-administered true-false PD screening inventory developed by Armanda W. Loranger, PhD. The diagnostic efficiency and psychometric properties of the IPDE-S in a 2-stage screen application were described in a prior report from the LSPD.21

International Personality Disorder Examination

The IPDE is the well-known semistructured interview designed for use by experienced clinicians for the assessment of both DSM and International Classification of Diseases, 10th Revision (ICD-10) PD features.13,21,22 The IPDE13 allows for both dimensional and categorical scoring of the DSM PDs. The DSM-III-R criteria were assessed in this study. Interviewers received training in IPDE administration and scoring by Loranger and were supervised throughout the project by the author. Supervision by the author was done blind to the subjects’ identity, putative PD status, and all prior assessment information. The interrater reliability for IPDE assessments was generally excellent at all 3 waves (intraclass correlations range: 0.84-0.92). Interviewers were blind to the screening results and putative PD status of the subjects and the same interviewer never saw the same subject more than once during the 3 study waves.

Structured Clinical Interview for DSM-III-R: Non-Patient Version

This is a well-known semistructured DSM-III-R Axis I clinical interview for use with nonpatients.23 It was administered prior to the IPDE.

Millon Clinical Multiaxial Inventory II

The Millon Clinical Multiaxial Inventory II (MCMI-II)26,27 is the well-known 175-item, true-false, self-administered inventory designed to coordinate with the DSM-III-R Axis II PDs. The MCMI-II possesses excellent psychometric properties26,27 and yields separate dimensional scores for each of the 11 PDs.

DATA ANALYSES

The primary analyses focused on the individual difference and mean level (or group) stability of PD features. These 2 forms of stability are conceptually independent and have different interpretations. Individual difference (ie, “rank order”17 or “normative”18) stability concerns the extent to which individuals maintain their relative position within a group ranking over time. Level, or mean level, stability concerns the extent to which group means on a variable (or disorder) of interest remain invariant over time. All data analyses were conducted at the level of dimensional scores on the IPDE and the MCMI-II.

Individual difference stability is assessed using the Pearson correlation coefficient ($r$) applied to actual symptom scores (ie, not ranks). The dimensional, cluster, and total scores for the PDs (IPDE and MCMI-II) were examined for rank order for wave 1 to wave 2, wave 2 to wave 3, and wave 1 to wave 3 comparisons. The differences between wave 1 to wave 2 vs wave 1 to wave 3 stabilities were tested to determine if the longer interval for the latter resulted in significantly lower stabilities; these correlated correlation coefficients were tested as per Meng et al.28

The stability of mean levels of the 11 Axis II disorders as well as the total number of PD features was evaluated for both measures using a repeated-measures analysis of variance (ANOVA) as implemented by MANOVA.29 The ANOVAs were all of a group × sex × time design, with 2 between-subjects factors (group, sex) and 1 within-subjects factor (time). All F tests involving the within-subjects factor (time) were corrected using the procedure of Greenhouse and Geisser.29 Two sets of 12 ANOVAs were conducted, one each for the IPDE and MCMI-II, for a total of 24 tests. A Bonferroni correction was used to maintain a nominal significance level of .05 for each effect repeatedly tested within each set of ANOVAs and only ANOVA results with $P < .004$ (ie, .05/12) are discussed as significant. Effect size estimates (partial $\eta^2$) were calculated for significant group, time, and group × time effects. The $\chi^2$ test was used to contrast proportions and the kappa ($\kappa$) coefficient was used to evaluate the stability of categorical PD diagnoses over time. All tests were 2-tailed.
Stability coefficients ($r$) to 0.70 (wave 1/wave 2) (Table 3). For the Axis "clusters," stability coefficient ranges were: cluster A (0.48-0.61), cluster B (0.60-0.78), and cluster C (0.52-0.67). Differences in the stabilities for the wave 1/wave 3 interval were tested to determine if lower stabilities for the wave 1/wave 3 interval were tested to determine if lower stabilities were observed for the longer time span. Wave 1/wave 2 stability coefficients were significantly higher for schizoid, antisocial, histrionic, and narcissistic PDs, clusters A and B, and total PD features (all $P < .05$). Overall, schizoid and antisocial PD features showed the highest level of individual differences stability on the IPDE.

For the overall mean MCMI-II PD dimensional score, stability coefficients range from 0.70 (wave 1/wave 3) to 0.77 (wave 1/wave 2) (Table 4). For clusters, stability coefficients ranges were: cluster A (0.64-0.73), cluster B (0.75-0.81), and cluster C (0.65-0.72). Differences in the MCMI-II stability coefficients were tested for the wave 1/wave 3 vs wave 1/wave 2 intervals. Wave 1/wave 2 stability coefficients were higher for paranoid, borderline, histrionic, narcissistic, and avoidant PDs; cluster A, B, and C, and total PD features (all $P < .05$). Overall, antisocial and passive-aggressive PD dimensions showed the highest level of individual differences stability on the MCMI-II.

### INDIVIDUAL DIFFERENCES STABILITY

definite and probable cases; a disorder was considered present if a subject met the criteria for the disorder at wave 1, wave 2, or wave 3 assessments. Eighty-one (62.8%) of the PPD subjects received an Axis I diagnosis compared with 32 (26.4%) NPD subjects ($\chi^2 = 33.30$, $P < .001$). Forty-one (31.8%) PPD subjects vs 21 (17.4%) NPD subjects reported a prior history of treatment by wave 3 ($\chi^2 = 6.97$, $P < .008$).
MEAN LEVEL STABILITY

The repeated-measures ANOVA results for IPDE are summarized with the F statistics, P values, and effect size estimates for the group, time, and group × time interactions (Table 5). The results for the sex factor will not be discussed owing to a relative absence of significant effects for the factor. Only effects associated with \( P < .004 \) are interpreted. For the IPDE, the PPD group consistently displayed greater PD symptomatology relative to the NPD group (main effect). The total PD features on the IPDE showed a main effect for time, indicating a decline over time. A significant group × time interaction superseded the main effects of group and time, indicating a more rapid decline in PD features among the PPD (vs NPD) subjects over time. The mean (± SD) total PD scores for the PPD group over time were: wave 1 = 16.72 (± 15.21), wave 2 = 8.91 (± 10.95), and wave 3 = 8.68 (± 10.58); and for the NPD group were: wave 1 = 4.90 (± 5.91), wave 2 = 3.24 (± 5.73), and wave 3 = 3.87 (± 5.52).

On the IPDE, paranoid, borderline, narcissistic, histrionic, dependent, obsessive-compulsive, and avoidant PD features followed a longitudinal pattern similar to that seen for the total scores (Table 5). Generally, most symptom change (decline) occurred from wave 1 to wave 2 assessments, whereas symptom levels remained relatively constant from wave 2 to wave 3 in both groups. An effect for time (as either a main effect or in interaction with group) was detected for 9 of the 11 PDs assessed by the IPDE. The mean effect sizes, for the significant findings, were as follows: group (.08), time (.08), and group × time interaction (.04).

On the MCMI-II overall average base rate–adjusted dimensional score, the PPD group consistently displayed higher levels of PD symptomatology (main effect) (Table 6). A group × time interaction indicated a more rapid decline in PD features among the PPD (vs NPD) subjects over time. The mean (± SD) for the average base rate–adjusted PD dimensional score on the MCMI-II were: for the PPD subjects, wave 1 = 53.42 (± 11.08), wave 2 = 48.49 (± 12.24), and wave 3 = 46.30 (± 12.03); for the NPD subjects: wave 1 = 37.82 (± 7.44), wave 2 = 35.87 (± 7.33), and wave 3 = 34.95 (± 8.44).

---

### Table 4. Stability Coefficients for DSM-III-R PD Features Assessed With the MCMI-II Across 3 Study Waves (N = 250)*

<table>
<thead>
<tr>
<th>Disorder</th>
<th>Wave 1/ Wave 2</th>
<th>Wave 2/ Wave 3</th>
<th>Wave 1/ Wave 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paranoid</td>
<td>0.75</td>
<td>0.70</td>
<td>0.63</td>
</tr>
<tr>
<td>Schizoid</td>
<td>0.68</td>
<td>0.61</td>
<td>0.61</td>
</tr>
<tr>
<td>Schizotypal</td>
<td>0.63</td>
<td>0.55</td>
<td>0.55</td>
</tr>
<tr>
<td>Cluster A</td>
<td>0.73</td>
<td>0.64</td>
<td>0.64</td>
</tr>
<tr>
<td>Antisocial</td>
<td>0.74</td>
<td>0.70</td>
<td>0.70</td>
</tr>
<tr>
<td>Borderline</td>
<td>0.73</td>
<td>0.72</td>
<td>0.64</td>
</tr>
<tr>
<td>Histrionic</td>
<td>0.76</td>
<td>0.73</td>
<td>0.69</td>
</tr>
<tr>
<td>Narcissistic</td>
<td>0.76</td>
<td>0.74</td>
<td>0.68</td>
</tr>
<tr>
<td>Cluster B</td>
<td>0.81</td>
<td>0.78</td>
<td>0.75</td>
</tr>
<tr>
<td>Avoidant</td>
<td>0.76</td>
<td>0.67</td>
<td>0.67</td>
</tr>
<tr>
<td>Passive-aggressive</td>
<td>0.78</td>
<td>0.76</td>
<td>0.73</td>
</tr>
<tr>
<td>Compulsive</td>
<td>0.66</td>
<td>0.69</td>
<td>0.60</td>
</tr>
<tr>
<td>Dependent</td>
<td>0.63</td>
<td>0.61</td>
<td>0.59</td>
</tr>
<tr>
<td>Cluster C</td>
<td>0.72</td>
<td>0.71</td>
<td>0.65</td>
</tr>
<tr>
<td>Total PD features</td>
<td>0.77</td>
<td>0.74</td>
<td>0.70</td>
</tr>
</tbody>
</table>

* MCMI-II indicates Millon Clinical Multiaxial Inventory II; PD, personality disorder. MCMI-II dimensions are base-rate transformed scores.
† Values are Pearson product-moment correlation coefficients based on 250 cases. All r values are significant at \( P < .001 \) (2-tailed).

### Table 5. Repeated-Measures ANOVA for PD Features for DSM-III-R Axis II Disorders Assessed by the IPDE*

<table>
<thead>
<tr>
<th>PD</th>
<th>Group</th>
<th>Time†</th>
<th>Group × Time†</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>P</td>
<td>ES‡</td>
<td>F</td>
</tr>
<tr>
<td>SZD</td>
<td>3.63</td>
<td>.06</td>
<td>.015</td>
</tr>
<tr>
<td>SIT</td>
<td>12.36</td>
<td>.001</td>
<td>.048</td>
</tr>
<tr>
<td>PAR</td>
<td>26.78</td>
<td>.001</td>
<td>.098</td>
</tr>
<tr>
<td>BDL</td>
<td>36.82</td>
<td>.001</td>
<td>.130</td>
</tr>
<tr>
<td>ANT</td>
<td>16.44</td>
<td>.001</td>
<td>.063</td>
</tr>
<tr>
<td>NAR</td>
<td>29.95</td>
<td>.001</td>
<td>.109</td>
</tr>
<tr>
<td>HST</td>
<td>23.13</td>
<td>.001</td>
<td>.086</td>
</tr>
<tr>
<td>DEP</td>
<td>19.25</td>
<td>.001</td>
<td>.073</td>
</tr>
<tr>
<td>CPS</td>
<td>34.03</td>
<td>.001</td>
<td>.122</td>
</tr>
<tr>
<td>AVD</td>
<td>9.15</td>
<td>.003</td>
<td>.036</td>
</tr>
<tr>
<td>PAG</td>
<td>20.29</td>
<td>.001</td>
<td>.076</td>
</tr>
<tr>
<td>Total PD</td>
<td>49.90</td>
<td>.001</td>
<td>.169</td>
</tr>
</tbody>
</table>

* ANOVA indicates analysis of variance; IPDE, International Personality Disorder Examination; PD, personality disorder; SZD, schizoid; SIT, schizotypal; PAR, paranoid; BDL, borderline; ANT, antisocial; NAR, narcissistic; HST, histrionic; DEP, dependent; CPS, obsessive-compulsive; AVD, avoidant; and PAG, passive-aggressive. Total PD indicates total personality disorder feature dimensional score on the IPDE. Number of cases equals 250 (8 cases with incomplete data omitted from the analysis). The data were analyzed at the level of dimensional scores on the IPDE. Results for the between-subjects main effect of sex as well as all 2-way and 3-way interactions involving the sex factor are omitted given the relative absence of significant results (see text). Degrees of freedom for F tests: group (1, 246), time (2, 492), time × group (2, 492). Degrees of freedom reflect the between-subjects sex factor.
† Within-subjects F tests evaluated for significance as per Greenhouse and Geisser.
‡ ES indicates effect size (partial \( \eta^2 \)), a statistic that describes the proportion of variability in PD features attributable to a factor (group, time, or group × time).
Table 6. Repeated-Measures ANOVA for PD Features for DSM-III-R Axis II Disorder Dimensions Assessed by the MCMI-II

<table>
<thead>
<tr>
<th>Disorder</th>
<th>Group</th>
<th>Time</th>
<th>Group × Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>P</td>
<td>ES‡</td>
</tr>
<tr>
<td>SZD</td>
<td>1.76</td>
<td>.19</td>
<td>.007</td>
</tr>
<tr>
<td>SZT</td>
<td>0.76</td>
<td>.91</td>
<td>.193</td>
</tr>
<tr>
<td>PAR</td>
<td>100.36</td>
<td>.001</td>
<td>.290</td>
</tr>
<tr>
<td>BDL</td>
<td>141.25</td>
<td>.001</td>
<td>.365</td>
</tr>
<tr>
<td>ANT</td>
<td>83.30</td>
<td>.001</td>
<td>.253</td>
</tr>
<tr>
<td>NAR</td>
<td>33.28</td>
<td>.001</td>
<td>.119</td>
</tr>
<tr>
<td>HST</td>
<td>10.49</td>
<td>.001</td>
<td>.041</td>
</tr>
<tr>
<td>DEP</td>
<td>0.04</td>
<td>.84</td>
<td>.000</td>
</tr>
<tr>
<td>CPS</td>
<td>2.57</td>
<td>.11</td>
<td>.010</td>
</tr>
<tr>
<td>AVD</td>
<td>49.92</td>
<td>.001</td>
<td>.169</td>
</tr>
<tr>
<td>PAG</td>
<td>133.12</td>
<td>.001</td>
<td>.351</td>
</tr>
<tr>
<td>Total PD</td>
<td>143.23</td>
<td>.001</td>
<td>.362</td>
</tr>
</tbody>
</table>

* ANOVA indicates analysis of variance; MCMI-II, Millon Clinical Multiaxial Inventory II; SZD, schizoid; SZT, schizotypal; PAR, paranoid; BDL, borderline; ANT, antisocial; NAR, narcissistic; HST, histrionic; DEP, dependent; CPS, obsessive-compulsive; AVD, avoidant; PAG, passive-aggressive; and PD, personality disorder.
† Time = overall average of the 11 MCMI-II personality disorder base rate–adjusted dimensions. Number of cases equals 250 (8 cases with incomplete data omitted from the analysis). The data were analyzed at the level of base rate–adjusted dimensional scores on the MCMI-II. Results for the between-subjects main effect of sex as well as all 2-way and 3-way interactions involving the sex factor are omitted given the relative absence of significant results (see text). Degrees of freedom for F tests: group (1, 246), time (2, 492), time × group (2, 492). Degrees of freedom reflect the between-subjects sex factor.
‡ Within-subjects F tests evaluated for significance as per Greenhouse and Geisser.
§ ES indicates effect size (partial h²), a statistic that describes the proportion of variability in PD features attributable to a factor (group, time, or group × time).

The results for the MCMI-II (Table 6) for specific PDs were somewhat less uniform than those observed for the IPDE. Borderline, avoidant, and passive-aggressive MCMI-II PD dimensions followed a longitudinal pattern similar to that seen for the overall average MCMI-II dimensional score (ie, PPD subjects displaying a more rapid decline over time) indicated by a significant time × group interaction. Subjects in the PPD group had higher scores on the schizotypal, paranoid, antisocial, narcissistic, and histrionic PD dimensions (main effects). Time main effects were observed for the schizotypal, paranoid, antisocial, and dependent PD dimensions, indicating a decline for both subject groups over time, with the largest changes typically occurring from wave 1 to wave 2. An effect for time (as a main effect or in interaction with group) was detected for 7 of the 11 PD dimensions assessed by the MCMI-II. The mean effect sizes, for the significant findings, were as follows: group = 0.22, time = 0.06, and group × time interaction = .04.

These analyses were judged to be the most sensitive for evaluating stability because they were based on a dimensional approach to the PDs. In clinical practice, however, categorical methods of classification remain prominent and therefore the stability of the classification of “any definite PD” in these data was examined. The proportion of subjects meeting the criteria for a definite PD of any type (including not otherwise specified) on the IPDE at wave 1 was 7.6%, wave 2 was 5.6%, and wave 3 was 2.8%. The k for wave 1/wave 2 = 0.45 (P<.001). The k coefficient could not be computed for other comparisons owing to the prevalences falling below 5%, and such categorical analysis, therefore, remains inconclusive. The MCMI-II does not generate categorical diagnoses. Furthermore, the presence of an Axis I disorder of any type at wave 1 did not significantly moderate PD feature changes over time for either the IPDE or MCMI-II.

**COMMENT**

These initial data from the LSPD suggest that PD features display relatively high levels of individual difference stability and appreciable mean level stability, with some change occurring over time. Categorical diagnoses could not be reliably evaluated for stability owing to a low prevalence, perhaps reflective of the conservative diagnostic thresholds embodied in the IPDE. To date, there are no other published reports or large-scale studies underway that address the full range of Axis II disorders from a prospective multiwave longitudinal perspective. The wave 1 to wave 2 findings of the LSPD are broadly consistent with the prior 2-occasion studies mentioned earlier; however, the absence of other multiwave data sets necessarily precludes comparison beyond wave 2 assessments.

The stability of individual differences is relatively high for total PD features assessed by both IPDE and MCMI-II. Differences in the rank order stability of the total PD feature scores across the 2 methods were expected given that the format of the MCMI-II remains constant over time, whereas clinical interviewers on the IPDE change. All 11 disorders revealed substantial stability of individual differences across the 4-year time span. The rank order stability of PD features is comparable to that found for major normal personality trait dimensions, where median stability coefficients range from 0.46 to 0.75. Finally, the rank order stability coefficients for these PD data would be higher were one to correct for...
unreliability of measurement, which serves to underesti- 
mate true stability. 

For mean level (ie, group) stability, the pattern of 
results reveals evidence of statistically significant changes 
in the mean levels of PD features, particularly within the 
PPD group, over time. However, despite being statisti- 
cally significant, the changes were relatively small as re- 
vealed through the effect size estimates. Overall, the data 
suggest a relatively high level of stability for PD group 
means over time for both measures (IPDE, MCMI-II). The 
greatest changes in the mean levels observed for both mea- 
sures tended to occur between the wave 1 and wave 2 
evaluations, a pattern of findings not unusual for pro- 
spective multiwave longitudinal studies.9 Change in mean 
levels from wave 2 to wave 3 was slight and longitudinal 
methodologists11 argue that this portion of the study pe- 
riod is the more accurate index of any real change. This is 
because changes from wave 1 to wave 2 might reflect 
some regression to the mean effect, even if relatively small, 
whereas change during the post–wave 2 assessments is 
more suggestive of genuine change free of regression to 
the mean effects.12

The overall theoretical implication of these mean 
level data suggests continuity consistent with the DSM de- 
nition of PD. Clinically, these data suggest also that change 
is possible in personality pathology, which may conceiv- ably be catalyzed through therapeutic efforts. One must 
remember that these are changes at the group level and 
some persons within the group may have changed rather 
notably. What, however, is causing the mean level de- 
clines over time, though they are small effects? The ob- 
served change is not likely to be explained entirely by a 
regression to the mean effect because (1) highly reliable 
measures were used, (2) subjects were assessed on 2 clinical 
instruments (IPDE, MCMI-II) that were different from 
the selection instrument (IPDE-S), (3) the variances within 
each of the 2 groups were broadly comparable over time, 
and (4) the groups were not at the ceiling or floor on ei- ther measure. Moreover, regression to the mean as an ef- 
fect per se loses much of its significance in a multiwave 
approach.2,12 The first-year college experience might have 
heightened PD symptomatology at wave 1, especially in 
those subjects who were already somewhat affected. How- 
ever, broadly comparable patterns of results (ie, small-
magnitude declines in group mean levels over time) have 
been observed for normal personality and such change 
is viewed as substantively trivial.5 Continued lifespan fol- 
low-up study of these individuals may help to shed fur- 
ther light on this issue of mean level stability and change.

Several caveats should be mentioned in consider- 
ing these LSPD data. Compared with the US population 
at large, the LSPD sample is more homogeneous in age, 
educational achievement, and social class, which may dif- 
ferentially affect the study results. Also, the LSPD was 
begun when the subjects were first-year university stu- 
dents and some of the most severely affected individu-
als with PD might have never successfully enrolled in col- 
lege and therefore would not be included in the study. 
However, one must recall the Axis I diagnostic data for 
the LSPD subjects (Table 2), diagnosed according to rig- 
orous clinical thresholds, before ascribing undue levels 
of mental health to these subjects. Multiwave longitudi- 
nal data from clinically based PD samples, mindful of their 
limitations for generalization, are needed for contrast with 
the LSPD results. Finally, one could argue that the sta- 
bility observed for PD features in this study might be 
artifactual (ie, subjects portray themselves as more con- 
sistent than they really are); however, a large methodo- 
logical literature in the personality assessment domain 
has rendered such an assertion untenable.1 The present 
data will be further disseminated using additional ap- 
proaches to the analysis of stability, including analysis 
of latent growth curves and factorial invariance, as well 
as structural equation–based approaches to the analysis 
of stability and change.8,10,11,21 William James33(p121) claimed 
that “by the age thirty, the character has set like plaster, 
and will never soften again,” and that indeed seems to 
be true of normal personality.1 However, clinical expe-
rience suggests that some PD features may diminish with 
age. Continued lifespan study of the LSPD subjects will 
offer the opportunity to determine if James’ view also holds 
for PD. Finally, this study is best viewed as heuristic and 
hypothesis-generating, undertaken within a context of 
exploration rather than a context of justification or con-
firmation.

Accepted for publication July 28, 1999.

This research was supported in part by grant MH45448 
from the National Institute of Mental Health, Washington, 
DC (Dr Lenzenweger).

Preliminary results of this study were presented at the 
Fifth Meeting of the International Congress on the Disor-
ders of Personality, International Society for the Study of 
Personality Disorders, University of British Columbia, Van-
couver, June 25, 1997; and at the 13th Annual Meeting of 
the Society for Research in Psychopathology, Harvard Uni-

I thank Margaret Dyer, MSW; Linda Foltz, PhD; 
Mickey Goldberg, MSW; Joan Lovejoy, MSW; Eileen Max-
well, MSW; Cynthia Neff, PhD; Judy Raabe, PhD; Carol Skin-
nert, MSW; and Ba Stopha, MSW for their able assistance 
in conducting clinical assessments. I thank Stephen Cor-
nelius, PhD, and Robert Rosenthal, PhD, for useful statis-
tical consultations and Jerome Kagan, PhD, for discussions 
regarding stability concepts. I also thank Lauren Korfine, 
PhD; Armand W. Loranger, PhD; Theodore Millon, PhD, 
and Paul T. Costa, PhD, for their comments on an earlier 
version of this article.

Reprints: Mark F. Lenzenweger, PhD, Department of 
Psychology, Harvard University, 33 Kirkland St, Cam-
bridge, MA 02138.

REFERENCES

1990.
Association; 1993.
3. American Psychiatric Association. Diagnostic and Statistical Manual of Mental 
1980.
1984.


