Published in final edited form as:

Assessment. 2023 September ; 30(6): 1764–1776. doi:10.1177/10731911221124340.

Levels of Personality Functioning Questionnaire 12–18 (LoPF-Q 12–18): Factor Structure, Validity, and Clinical Cut-Offs

Sophie Kerr[1], **Veronica McLaren**[1], **Kiana Cano**[1], **Salome Vanwoerden**[2], **Kirstin Goth**[3], **Carla Sharp**[1]

[1]University of Houston, TX, USA

[2]University of Pittsburgh, PA, USA

[3]Psychiatric University Clinics Basel, Switzerland

Abstract

The Levels of Personality Functioning Questionnaire 12–18 (LoPF-Q 12–18) is the only self-report measure informed by the Level of Personality Functioning (*Diagnostic and Statistical Manual of Mental Disorders* [5th ed.; *DSM-5*; American Psychiatric Association, 2013]) Alternative Model of Personality Disorders developed for adolescents. The present investigation includes two studies evaluating the English LoPF-Q 12–18. In Study 1, single-factor and bifactor structures (unidimensional severity criterion and four specific factors: identity, self-direction, empathy, intimacy) were evaluated in an ethnically diverse community sample (N = 453; age 10–18; 57% female). Study 2 used a community control (n = 298; age 10–18; 54.4% female) and clinical sample (n = 94; age 11–18; 58.5% female) to examine reliability, validity, and clinical utility. Study 1 results supported the bifactor model, with a robust general factor and little multidimensionality caused by the group factors, suggesting an essentially unidimensional structure. Study 2 revealed good internal consistency and construct validity and provided clinical cut-offs, supporting the use of the LoPF-Q 12–18 total score in research and clinical applications.

Keywords

personality functioning; adolescents; alternative model of personality disorder

The Alternative Model for Personality Disorders (AMPD; APA, 2013) was developed in response to concerns over the validity and utility of the categorical diagnostic system for personality disorders (PDs; Clark, 2007; Krueger et al., 2018; Morey et al., 2011). Although the AMPD was relegated to Section III (Emerging Models and Measures) of the *DSM-5* due to concerns that it was not ready for use in clinical practice, it gained traction as a more empirically sound diagnostic approach and informed the new ICD-11 criteria for PD

Corresponding Author: Carla Sharp, Department of Psychology, University of Houston, 3695, Cullen Blvd, Room 126, Houston, TX 77204, USA., csharp2@uh.edu.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Supplemental Material

Supplemental material for this article is available online.

(Sharp & Miller, 2022). The AMPD includes two primary criteria. Criterion A requires moderate or greater impairment in Level of Personality Function (LPF): self (identity, self-direction) and interpersonal (empathy, intimacy). Criterion B requires that individuals show at least one pathological personality trait across five domains, which describe the behavioral manifestations of personality pathology.

Criterion A meets shortcomings of the categorical system by describing PD severity on a single continuum (dimension) of healthy to unhealthy self-interpersonal functioning, with the latter extreme identifying the core feature that differentiates PDs from other psychopathology (Morey et al., 2011). Specifically, the *DSM-5* states that

> to use the Level of Personality Functioning Scale (LPFS), the clinician selects the level that most closely captures the individual's current *overall* level of impairment in personality functioning. The rating is necessary for the diagnosis of a personality disorder (moderate or greater impairment) and can be used to specify the severity of impairment present for an individual with any personality disorder at a given point in time. The LPFS may also be used as a *global indicator of personality functioning* without specification of a personality disorder diagnosis, or in the event that personality impairment is subthreshold for a disorder diagnosis. (p. 772)

As such, the *DSM* suggests the LPFS to be a unidimensional severity criterion consistent with the original intention of the LPF as described in the early publications leading up to the publication of the AMPD (e.g., Morey et al., 2011). Using the *DSM-5* Level of Personality Functioning Scale (LPFS), clinicians can rate Criterion A functioning in each element on a 5-point scale from *little to no impairment* to *extreme impairment*, with the overall goal of deriving a single severity score consistent with the idea of a single dimension of personality functioning. The LPFS has shown incremental predictive utility for psychosocial functioning over the categorical *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; *DSM-IV*; American Psychiatric Association, 1994) disorders, good inter-rater reliability, and external validity (Busmann et al., 2019; Few et al., 2013; Morey et al., 2013). To standardize assessment, researchers have developed several semi-structured interviews and self-report measures of LPF.

Although comparatively more empirical work has focused on Criterion B, there is a growing literature on Criterion A and its measurement. Recent research has evaluated whether the latent structure of LPF measures aligns with the *DSM-5* conceptualization of a unidimensional severity criterion. Support for unidimensionality of the LPFS—Self-Report (LPFS-SR) has been demonstrated in community samples (Hopwood et al., 2018; Morey, 2017). However, in another study, the LPFS-SR failed to demonstrate a unidimensional or a four-factor structure, and instead revealed a different but related four-factor structure (Sleep et al., 2019). In addition, the LPFS—Brief Form (LPFS-BF, Hutsebaut et al., 2016) and its revised version (LPFS-BF 2.0; Weekers et al., 2019) have shown a two-factor structure (self, interpersonal) in a clinical (Hutsebaut et al., 2016; Weekers et al., 2019) and combined clinical and incarcerated sample (Bach et al., 2016). However, these studies evidenced high correlations between domains, suggesting a unidimensional construct. This has led to arguments against defining two domains or four LPF components (Sleep et al., 2019). An alternative is to conceive of LPF as best captured by a second-order or bifactor structure

(Sharp & Wall, 2021). In support of this idea, the *DSM-5* Levels of Personality Functioning Questionnaire—Short Form (Siefert et al., 2020) and the Self and Interpersonal Functioning Scale (Gamache et al., 2019) both demonstrated second-order factor structures consisting of the four LPF elements and an overarching personality dysfunction factor in combined community and clinical samples. Bliton et al. (2021) was the first to test bifactor models of LPF using comprehensive indices, an approach that posits both a general factor explaining covariance across all items and orthogonal specific factors explaining excess shared variance among item clusters. They found mixed support for single-factor, two-factor, four-factor, and bifactor models across the LPFS-SR, LPFS-SRA, LPFS-BF. However, the bifactor model parameters indicated the majority of variance was accounted for by the general factor and little to no variance accounted for by the specific factors, supporting a unidimensional structure.

Almost all AMPD research has focused on adults despite strong evidence for the validity and early intervention of PD in adolescence (see Chanen et al., 2016; Sharp & Fonagy, 2015 for reviews). Early intervention is not possible without early detection, which requires empirically validated assessment tools that map onto the most current psychiatric nosology. The AMPD offers particular advantages for assessment of personality function in adolescence over traditional *DSM-5* Section II approaches because its dimensional nature allows for assessment of personality pathology at earlier stages of disorder (Sharp, Kerr, et al., 2021) through the identification of at-risk adolescents who do not yet meet PD criteria. Moreover, consistent with developmental models of personality (McAdams, 2015) and personality dysfunction (e.g., Sharp & Wall, 2017, 2021), the assessment of self- and interpersonal function is critical during adolescence when these functions first come on line in adult form.

The Levels of Personality Functioning Questionnaire (LoPF-Q 12–18; Goth et al., 2018b) is the only measure of LPF specifically developed for use with adolescents. It is a self-report measure intended for youth ages 12 to 18 that yields a total score of PD severity and four optional scale scores corresponding to the LPF dimensions (identity, self-direction, empathy, and intimacy). The LoPF-Q 12–18 was initially developed in German by a Swiss research group. Test construction was based on the general descriptions of Criterion A in the AMPD (*DSM-5*; APA 2013) and beta draft of ICD-11 and was informed by an in-depth content analysis of related models and measures and their reported clinical validity (for details see Goth et al., 2018a). The final test version was constructed in a balanced sample according to gender and age, combining 351 adolescents from a general population and a clinical sample of 241 patients from Switzerland, Germany, and Austria. It was validated in a large sample of 823 students and inpatient and outpatient youth (age 11–19; Goth et al., 2018a). The LoPF-Q 12–18 demonstrated good to excellent internal consistency, construct validity with significant, large differences in scores between students and patients diagnosed with PD (effect size $d = 2.1$) and high correlations with BPFSC-11 scores, and clinical utility in identifying PD patients in receiver operating characteristic (ROC) analyses. The LoPF-Q 12–18 has since been translated into English (Sharp & Vanwoerden, 2018), Spanish (Kassin & Hackradt, 2019); Lithuanian (Barkauskienė & Skabeikytė, 2020), and Turkish (Cosgun et al., 2021), but has only been used in two other published studies. In a student ($n = 282$) and clinical ($n = 52$) sample of Turkish adolescents, Cosgun et

al. (2021) demonstrated good scale reliability, convergent and discriminant validity, and clinical utility in predicting PD diagnosis. In a hybrid community-patient sample of German adolescents, Gander and colleagues (2020) demonstrated construct validity with significant mean differences emerging between community adolescents and patients without PD, and those with PD on the total and scale scores.

Although these studies provide a good basis for the LoPF-Q 12–18, several gaps remain. As in adult studies of LPF, there are mixed findings regarding the factor structure of the LoPF-Q 12–18 and further investigation is needed. Although Cosgun et al. (2021) demonstrated a unidimensional solution, Goth et al. (2018a) found that exploratory factor analyses (EFAs) at the item level supported a unidimensional structure while EFAs at the subscale level supported a four-factor structure, leading the authors to suggest future studies examine a bifactor structure. Second, no studies have evaluated the psychometric properties of the LoPF Q 12–18 in an English-speaking population or in an ethnically diverse sample of youth, rendering its use in English-speaking populations uncertain. Finally, further work is needed to determine clinical cut-offs to facilitate use by clinicians as previous suggestions range from 163 to 180 (Cosgun et al., 2021; Goth et al., 2018a). Although dimensional scores are clinically useful and should be employed, clinicians also find benchmarks useful to interpret severity in psychopathology.

To address these gaps, the present investigation consists of two studies. In Study 1, we examined the factor structure of the LoPF-Q 12–18 in an ethnically diverse community sample of youth ages 10 to 18. Based on previous evidence for both unidimensional and multidimensional aspects of LPF (see Sharp, Vanwoerden, et al., 2021 for a more detailed review), we hypothesized that the LoPF-Q 12–18 would demonstrate a bifactor structure, with a general personality dysfunction factor and four specific factors corresponding to each scale (identity, self-direction, empathy, intimacy). Fitting a bifactor structure allows us to test the ability to interpret the factor structure as unidimensional or multidimensional and examine whether specific scale factors yield reliable and meaningful information after accounting for the general personality dysfunction factor score. Given arguments that LPF is an entirely unidimensional construct and the domains offer no unique descriptive value (Sleep et al., 2019), as well as the relative simplicity of a unidimensional solution for future research and clinical work with the LoPF-Q 12–18, we also tested the fit of a one-factor solution without any specific factors.

Study 2 used a community control and clinical samples of youth to examine reliability, demographic correlates, validity, and clinical utility of the LoPF-Q 12–18. The scoring of the LoPF-Q 12–18 (i.e., whether to further examine the total score, scale scores, or both) was based on the findings of Study 1. Based on previous studies of the LoPF-Q 12–18 and the conceptualization of LPF as a unidimensional severity criterion, we hypothesized that the LoPF-Q 12–18 would show adequate internal consistency. Second, we examined relationships between the LoPF-Q 12–18 and demographic variables (i.e., gender, age, race, and ethnicity) to provide preliminary information on how the measure may perform in different samples and potential relationships between LPF and demographic variables in youth. Regarding gender, Cosgun et al. (2021) found that boys scored significantly higher, though only slightly, than girls on the Turkish LoPF-Q 12–18. Findings using Section II

measures of personality pathology have been mixed, with females showing higher rates of personality pathology in clinical samples but no significant gender differences in a large community sample (Ha et al., 2014; Zanarini et al., 2011), leading to no a priori hypotheses. Regarding age, there were no significant differences in Turkish LoPF-Q 12–18 between younger and older adolescent age groups (Cosgun et al., 2021). However, a study evaluating mean differences in a latent variable of maladaptive identity (a core aspect of LPF) found normative increases in maladaptive identity from age 12 to17 followed by a decrease back to levels observed at age 12 (Sharp, Vanwoerden, et al., 2021). This is consistent with longitudinal studies using Section II measures suggesting normative increases in personality dysfunction through early to mid-adolescence followed by declines into adulthood, though groups with more severe levels of personality pathology do not show age-related declines (Cohen et al., 2005; Wright, Zalewski, et al., 2016). Therefore, we hypothesized that older youth would show lower levels of personality dysfunction in the community control sample, but perhaps not in the clinical sample. To our knowledge, no studies have examined differences in LPF based on race or ethnicity and, while worthy of a full, separate investigation, these exploratory analyses will provide preliminary evidence using a diverse sample of youth.

Construct validity was examined in multiple ways in Study 2. First, we tested associations with measures of personality dysfunction based on (a) the traditional, *DSM-5* Section II perspective of BPD, (b) an AMPD Criterion A identity perspective, and (c) an AMPD Criterion B trait perspective. We hypothesized that greater personality dysfunction on the LoPF-Q 12–18 would be associated with greater personality dysfunction on these measures. As an additional index of construct validity, we examined differences in LoPF-Q 12–18 scores between the community control and clinical samples and hypothesized that there would be higher scores (i.e., greater impairment) in the clinical sample. To provide information on clinical utility and to clarify cut-off scores, we examined the ability of the LoPF-Q 12–18 to (a) predict sample type (community control vs. clinical) and (b) predict youth above a clinical cut-off of Section II-defined borderline personality disorder (BPD) across both samples, given that clinical cut-offs are available for BPD measures of youth personality. A further reason for using the BPD measure to derive clinical cut-offs is the evidence that Section II BPD closely reflects general PD severity (Clark et al., 2018; Sharp et al., 2015; Wright, Hopwood, et al., 2016). We hypothesized that the LoPF-Q 12–18 total score would show accuracy in discriminating between samples of youth above and below the clinical cut-off of BPD symptoms and that derived cut-off scores would help clarify discrepant findings of previous studies.

Study 1

Method

Participants and Procedures.—Study 1 used a sample of 453 participants recruited from schools and non-profit youth programs in a major metropolitan area in the Southwestern United States. Inclusion criteria for the study required participants to be aged 10 to 18, students at a general education school, and proficient in English to complete surveys. The sample consisted of 258 (57%) females, 191 (42.2%) males, and 5 (.9%)

gender diverse adolescents. Participants were an average age of 13.1 (*SD* = 2.0), with 15.5% in 5th grade, 17.9% in 6th grade, 21.4% in 7th grade, 13.2% in 8th grade, 14.3% in 9th grade, 8.8% in 10th grade, 2.9% in 11th grade, and 5.5% in 12th grade. The racial and ethnic breakdown was: 59.6% Hispanic or Latin American, 11.9% Black or African American, 10.4% Caucasian or White, 5.5% Asian, 1.3% Native American, and 11.3% multiracial or other. Details of data collection procedures are provided in the Supplementary Materials. We report on the adequacy of our sample size and explain the purpose of each measure in the study.

Measures

Levels of Personality Functioning Questionnaire.—The LoPF-Q 12–18 (Goth et al., 2018b) is a 97-item self-report measure of impairment in personality functioning as described by Criterion A of the *DSM-5* Section III AMPD and ICD-11 that was specifically developed for use with adolescents age 12 to 18. Items are rated on a 5-point scale ranging from 0 (*no*) to 4 (*yes*). The measure yields a total score and four scale scores corresponding to each element of Criterion A: identity (e.g., “I am confused about what kind of person I really am”), self-direction (e.g., “I have difficulties reaching goals I set for myself”), empathy (e.g., “I often don’t understand other people’s reactions to my behavior”), and intimacy (e.g., “I prefer others to not get too close to me”). The LoPF-Q 12–18 was initially developed in German by a Swiss research group. There are now several translations that are available in electronic format on the project website (academic-tests.com) and can be requested for free for research purposes. Details on the English translation process are provided in the Supplementary Materials. The validation study of the German version of the LoPF-Q 12–18 (Goth et al., 2018a) showed good to excellent internal consistency, strong construct validity, and clinical utility (see introduction). In the current sample, internal consistency was excellent for the total score (α = .96) and good to excellent for the scale scores (identity: 23 items, α = .83, self-direction: 25 items, α = .95, empathy: 26 items, α = .87, intimacy: 23 items, α = .86).

Data Analytic Strategy.—We conducted confirmatory factor analyses (CFAs) to evaluate bifactor and single factor structures of the LoPF-Q 12–18 in Mplus (Version 8; Muthén & Muthén, 2017). The bifactor structure included a general factor that consists of all items as well as four specific factors that consist of items specific to identity, self-direction, empathy, and intimacy. The general factor was specified to be orthogonal to the specific factors and specific factors were specified to be orthogonal to each other. Weighted least squares means and variance adjusted (WLSMV) estimation was used due to the categorical nature of the data. Models were identified by constraining the first-factor loading to one and freely estimating each factor variance. Model fit for both CFAs was determined by using the following indices: the root-mean-square error of approximation (RMSEA), with values of less than .08 indicating reasonable fit and values above .10 suggesting poor fit; the comparative fit index (CFI) and Tucker–Lewis Index (TLI), with values between 0.95 and 1.00 indicating excellent fit and values between .90 and .95 indicating acceptable fit; and the standardized root-mean-square residual (SRMR), with values <.08 indicating acceptable fit (Browne & Cudeck, 1993; Hu & Bentler, 1999). Chi-square tests are reported; however, the

chi-square test is sensitive to sample size and is therefore not given as much weight as the fit indices listed above (Fan et al., 1999).

For the bifactor model, in addition to conventional goodness-of-fit cut-off scores, we calculated statistical indices including Omega reliability coefficients, factor determinacy (FD), construct reliability (H), explained common variance (ECV), percent of uncontaminated correlations (PUC), and relative parameter bias. Details on calculation and interpretation are in the Supplementary Materials. This approach reduces the sensitivity to sample size by relying on goodness-of-fit statistics and follows best practices for bifactor analysis (Rodriguez et al., 2016).

Results

The bifactor model demonstrated acceptable fit on all indices except TLI and CFI which were .01 below the .90 cutoff, x^2 (4,456) = 8,440.94, $p < .001$; RMSEA = .04; CFI = .89; TLI = .89; SRMR = .07. For the single factor CFA, CFI, TLI, and SRMR were all outside acceptable ranges, x^2 (4,559) = 13,098.49, $p < .001$; RMSEA = .06; CFI = .77; TLI = .76; SRMR = .09.

Factor loadings of each item in the bifactor model are displayed in Table S1 Supplementary Materials. The general factor was most heavily represented by self-direction scale. All but one of the 25 self-direction items loaded strongly (≥ .54) onto the general factor. The remaining item loaded moderately onto the general factor (.39) and strongly, but negatively, onto the specific factor (−.48). Of the 23 identity items, 13 items exhibited strong factor loadings (≥.52) on the general factor, nine items exhibited strong loadings (≥ .51) on the specific factor of identity, and the remaining item loaded moderately onto the general factor (.35). Twelve of the 26 empathy scale items exhibited strong general factor loadings (≥.41), while 10 exhibited strong specific factor loadings (≥.43). Six items did not load strongly onto either factor. Of the 23 intimacy items, 11 loaded strongly (≥.48) onto the general factor, 8 loaded strongly (≥.41) onto the specific factor, and 4 items did not load strongly onto either factor.

Regarding model-based reliability indices, omega was high for the general factor (.97) and omega subscale was high for all four subscales (> .89). OmegaH was .90, indicating that 90% of the variance of unit-weighted total scores can be attributed to the individual differences on the general factor. Comparing omega and omegaH shows that only 7% (.97–.90) of the reliable variance in total scores can be attributed to the multidimensionality caused by the group factors and 3% is estimated to be due to random error. Therefore, the raw total scores can be interpreted as essentially unidimensional (Rodriguez et al., 2016). OmegaHS was low for self-direction (.00) and moderate for all other subscales (.34–.44), showing that subscale reliability dwindles substantially after partitioning out the variance for the general factor, especially for self-direction. Therefore, the reliability judged by omega subscale was largely attributable to individual differences on the general factor.

All scales were above the .90 cut-off for FD (all >.93; Gorsuch, 1983) except for self-direction (.83). In addition, all scales were above the .70 cut-off for H (all > .84; Hancock & Mueller, 2001) except for self-direction (.56). These results suggest that all

scales except for self-direction represent well-defined latent variables and add justification for their use in future measurement models using a structural equation modeling (SEM) framework (Rodriguez et al., 2016). However, these results again suggest caution in using the self-direction scale. ECV of the general factor was .67 and PUC was .76. Therefore, the general ECV falls slightly below the cut-offs (.70 ECV and .70 PUC) provided by Rodriguez et al. (2016) to determine when common variance can be regarded as essentially unidimensional. However, Reise and colleagues (2013) explain that when a model shows PUC values <.80, general ECV values >.60, and omegaH values greater than .70, the multidimensionality is not severe enough to disqualify the interpretation of the instrument as primarily unidimensional. Average relative parameter bias was 26%, which is considered an unacceptable degree of relative bias (Muthén et al., 1987), suggesting caution if collapsing the measure into a unidimensional space in an SEM context (Rodriguez et al., 2016).

Although the single-factor CFA did not show adequate fit, the bifactor model demonstrated a robust general factor with only 7% of reliable variance in total scores attributed to multidimensionality caused by the group factors. Therefore, findings support an essentially unidimensional structure. Combined with the complexity that would be required to use a bifactor scoring approach in future research and clinical practice, we use only the total score in Study 2.

Study 2

Method

Participants and Procedures.—The community control sample was drawn from the sample described in Study 1. From those 453 participants, we excluded 155 who scored above the clinical cut-off on the Brief Problem Monitor (BPM; Total Problems *T*-score > 65), resulting in a final sample size of 298. Demographic characteristics are presented in Table 1. Procedures were identical to Study 1.

The clinical sample consisted of 94 participants drawn from two outpatient clinics. Participants and their primary caregivers provided assent and consent, respectively. Inclusion criteria required that youth be age 11 to 18, be receiving or actively seeking psychiatric treatment, and speak fluent English. The only exclusion criterion was the presence of a psychotic disorder. Demographic characteristics are described in Table 1. Although the BPM was used to exclude participants in the community control sample, this screening measure was not completed by the clinical sample participants and therefore cannot be used to describe the nature of the clinical sample. However, 82 of the 94 clinical sample participants completed a more comprehensive measure of adolescent psychopathology, the Personality Assessment Inventory—Adolescent version. Of these 82 participants, 57.3% of the sample had an elevated (*T*-score > 60) depression score and 25.6% had clinically significant (*T*-score > 70) depression scores. Anxiety scores were elevated for 61.7% and clinically significant for 33.3% of the sample. BPD scale scores were elevated for 50% and clinically significant for 17.1% of the sample. Aggression scores were elevated for 22.2% and clinically significant for 6.2% of the sample.

Power analysis using G*Power was conducted to determine the sample size for independent samples t-tests (power = .80, α = .05, allocation = 3:1) and found that sample sizes of 34 and 100 per group were sufficient to detect a moderate effect size (d = .05).

Measures

LoPF-Q 12–18.—The LoPF-Q 12–18 was described in Study 1. Based on the findings of Study 1, we used only the total score.

The Assessment of Identity Development in Adolescents.—The Assessment of Identity Development in Adolescents (AIDA; Goth & Schmeck, 2018) is a 58-item self-report questionnaire for adolescents assessing identity development in terms of impairments in personality functioning. Items are rated on a 5-point Likert-type scale and correspond to two scales: discontinuity and incoherence. The total score, representing identity diffusion, was used for the current study to evaluate construct validity. In addition, given that the 23 LoPF-Q 12–18 identity items were drawn from the AIDA, we also calculated the total diffusion score excluding these overlapping items excluded ("AIDA—no LoPF"). The AIDA has demonstrated reliability and validity in adolescents from a wide range of cultural backgrounds (e.g., Ercegovic et al., 2018; Kassin et al., 2013; Ragelienė & Justickis, 2016; Tardivo et al., 2016) and in an inpatient sample of American adolescents (Lind et al., 2019). In the current study, Cronbach's alpha indicated high internal consistency in both the healthy control (AIDA: α = .94, AIDA-no LoPF: α = .91) and clinical sample (α = .95, AIDA-no LoPF: α = .95).

Borderline Personality Features Scale for Children-11.—The Borderline Personality Features Scale for Children-11 (BPFS-C-11; Sharp et al., 2014) is an 11-item self-report instrument that assesses borderline personality features among children and adolescents. The BPFS-C-11 is a reduced item version of the 24-item BPFS-C, which was developed by Crick et al. (2005). Items are rated on a 4-point Likert-type scale to yield a total score and four subscales (affective instability, identity problems, negative relationships, and self-harm). The BPFS-C-11 has shown good construct validity (Sharp et al., 2014) and longitudinal invariance over the period of adolescence (Vanwoerden et al., 2019). In addition to using the total score as a dimensional measure of borderline features, we also used a cutoff score to dichotomize the clinical sample into BPD and non-BPD groups for ROC analyses. We used a cutoff score of 34, which has demonstrated high diagnostic accuracy for BPD (Sharp et al., 2014). Cronbach's alpha for the total score was .79 in the healthy control sample and .86 in the clinical sample.

Personality Inventory for the* DSM-5 *Brief Form.—The Personality Inventory for the *DSM*-5 Brief Form (PID-5-BF; APA, 2013), a 25-item self-report measure, was used to further evaluate construct validity by assessing maladaptive personality traits as defined by Criterion B of the AMPD. This measure yields a total score and five scale scores with equal numbers of items that correspond to five trait dimensions (negative affect, detachment, antagonism, disinhibition, and psychoticism). We used the average score of all items and average item scores for each scale in the current study. Previous studies in adolescents have supported the use of the total score as an indicator of overall profile elevation (Fossati et

al., 2017). In addition, the PID-5-BF has demonstrated adequate psychometric properties in adolescents including replicable factor structure, reliability, and construct validity (Anderson et al., 2018; Bach et al., 2016; Fossati et al., 2015). In the community control sample, internal consistency was good (α = .89) for the total scale and questionable to acceptable for the scales (negative affect: α = .65, detachment: α = .62, antagonism: α = .73, disinhibition: α = .77, psychoticism: α = .77). In the clinical sample, internal consistency was excellent (α = .91) for the total scale and questionable to acceptable for the scales (negative affect: α = .82, detachment: α = .68, antagonism: α = .69, disinhibition: α = .78, psychoticism: α = .82).

Data Analytic Strategy.—For Study 2, all analyses were conducted using *SPSS (Version 26)* (IBM, 2019). We first examined the internal consistency of the LoPF-Q 12–18 in each sample using Cronbach's alpha. To examine relations between demographic characteristics, we ran Pearson's correlations between LoPF-Q 12–18 scores and age, t-tests comparing females and males, and one-way ANOVA's with follow-up Tukey's HSD tests comparing racial groups. To examine construct validity, we examined associations with external variables related to personality functioning, following American Educational Research Association et al. (2014) standards for determining construct validity. Specifically, we ran Pearson's correlations with measures of AMPD Criterion A identity (AIDA), *DSM-5* Section II BPD (BPFSC-11), and AMPD Criterion B pathological personality traits (PID-5-BF total and subscale average scores). As another index of construct validity, we conducted a t-test to examine the difference between the clinical and control groups on the LoPF-Q 12–18 total score. We also examined group differences on construct validity measures and demographic variables (age, race, gender). In the case that age, race, or gender varied between groups and were significantly associated with LoPF-Q 12–18 scores, we also ran ANCOVA's to examine group differences on the LoPF-Q 12–18 total score while controlling for relevant demographic variables. Finally, we examined clinical utility and determined clinical cut-off scores for the LoPF-Q 12–18 total score using ROC analyses. Details on ROC analysis and interpretation are in the Supplementary Materials. Using participants from both the control and clinical samples, we ROC analyses predicting: (a) the control versus clinical sample and (b) youth above the BPFSC-11 clinical cut-off.

Results

Internal Consistency and Correlations Between Scales.—Cronbach's alpha indicated excellent internal consistency of the total LoPF-Q 12–18 scale in both the control (α = .95) and clinical samples (α = .96).

Relations With Demographic Variables.—In the control sample, older age was significantly associated with lower scores on the LoPF-Q 12–18 (Table 2). In contrast, age showed a non-significant positive correlation with the LoPF-Q 12–18 in the clinical sample. This pattern was also shown with other indices of personality dysfunction: Older age was significantly associated with lower scores on the AIDA, BPFSC-11, the total PID-5-BF scale, and three PID-5-BF scales in the control sample but was significantly associated with higher scores on the BPFSC-11, PID-5-BF total score, and four PID-5-BF scales in the clinical sample. *T*-tests revealed that females had significantly higher total scores than males

in the community control sample, $M = 143.09$, $SD = 47.09$ versus $M = 131.33$, $SD = 45.54$, $t(292) = 2.16$, $p < .05$; $d = .25$, but no significant gender difference in the clinical sample. In the control sample, a one-way ANOVA revealed significant differences in LOPF-Q 12–18 total scores between racial and ethnic groups, $F(5,292) = 4.77$, $p < .001$, $\eta^2 = .08$. Follow-up Tukey's LSD comparisons showed that total scores for white youth ($M = 111.09$, $SD = 47.63$) were significantly lower than total scores of Hispanic ($M = 145.80$ $SD = 46.01$, $p < .001$) and multiracial youth ($M = 154.12$, $SD = 50.26$, $p < .001$). African American youth also had significantly lower scores ($M = 122.80$, $SD = 42.37$) than Hispanic ($M = 145.80$, $SD = 46.01$, $p < .01$) and multiracial youth ($M = 154.12$, $SD = 50.26$, $p < .01$). Comparisons were not computed for Asian or Native American youth given low sample sizes. In the clinical sample, a one-way ANOVA did not reveal significant differences between racial groups.

Construct Validity.—Pearson's correlations between LoPF-Q 12–18 scores and external measures of personality function are displayed in Table 2. The AIDA identity diffusion score demonstrated large correlations with the LoPF-Q 12–18 in the control and clinical samples. When excluding items included on both the AIDA and LoPF-Q 12–18 ("AIDA—no LoPF"), correlations remained high in both samples. Correlations with the BPFS-C-11, PID-5-BF overall average score, and all PID-5-BF subscales were similarly strong across both samples.

Differences between the control and clinical samples on demographic and clinical variables are displayed in Table 1. A *t*-test revealed that the clinical sample scored higher on the LoPF-Q 12–18. Regarding demographic covariates, the clinical sample was significantly older and contained significantly more females, and there were significant group differences in race and ethnicity. The clinical sample also had higher scores on AIDA identity diffusion, the BPFSC-11, the PID-5-BF overall score, and all PID-5-BF subscales except for antagonism. Due to significant associations between the LOPF-Q 12–18, age, gender, and race, and significant group differences in these variables, we conducted a two-way ANCOVA comparing LoPF-Q 12–18 scores with sample and race entered as fixed factors and age and gender entered as covariates and found that significant group differences remained, $F(1,378) = 54.03$, $p < .001$, $\eta p^2 = .13$.

Clinical Utility.—Two ROC curves were drawn to examine the clinical utility of the LoPF-Q 12–18 (see Figure S1 in Supplementary Materials). The LoPF-Q 12–18 demonstrated moderate accuracy [Area under the curve (AUC) = .83] in discriminating between the control and clinical samples (optimal cut point = 177.5, sensitivity = .745, specificity = .748). The LoPF-Q 12–18 was moderately accurate (AUC = .82) in discriminating between youth above and below the clinical cut-off of the BPFSC-11 across both samples, with an optimal cut-off of 176.5 (sensitivity = .760, specificity = .757). The group above the BPFS-C-11 cut-off ($n = 288$) had a mean LoPF-Q 12–18 score of 139.11 ($SD = 53.43$), compared with 209.03 ($SD = 55.27$) in the group above the cut-off ($n = 104$).

Discussion

In Study 1, goodness of fit statistics generally supported a bifactor structure of the LoPF-Q 12–18 with a general factor and four specific factors corresponding to each scale. In

contrast, fit statistics were largely inadequate for the single factor CFA. However, we note that some indices (CFI, TLI) are likely to yield biased, poor fit in CFAs with this number of indicators (97 items; Shi et al., 2019). For the bifactor model, the general factor was most represented by items on the self-direction scale, followed by identity, empathy, and then intimacy. This suggests that the self-domain scales are most closely related to global personality functioning, consistent with the idea that self-functioning is the driver or nexus of Criterion A (Sharp, 2020; Sharp & Wall, 2021). Specifically, compared with the interpersonal component, the self-component has been shown to be a stronger predictor of psychosocial impairment (Buer Christensen et al., 2020), severity on the Brief Symptom Inventory (Hutsebaut et al., 2017), general distress on the Symptom Checklist-90-Revised (Bach & Hutsebaut, 2018), and change in treatment (Weekers et al., 2019).

Given the tendency for bifactor models to overfit data, we calculated several psychometrically informative bifactor-derived statistics alongside model fit statistics (Rodriguez et al., 2016). Model-based reliability indices revealed that after partitioning out the variance for the general factor, reliability for the scale scores dwindled substantially, especially for the self-direction score, suggesting that specific factors may not represent unique constructs beyond general personality dysfunction. In addition, only 7% of reliable variance in total scores could be attributed to multidimensionality caused by the group factors, suggesting that the measure can be regarded as essentially unidimensional. However, indices also suggested that three of the four subscales (identity, empathy, and intimacy) yielded some reliable and meaningful information after accounting for the general factor, such that a bifactor model may more appropriately account for the additional variability explained by these specific factors. With these complexities in mind, the variance that could be attributed to multidimensionality is too small to justify calculating scale scores to capture distinct constructs. This is further supported by excellent internal consistency of the total score. These results add to only two existing studies on the factor structure of the LoPF-Q 12–18, which found both unidimensional and four-factor structures (Cosgun et al., 2021; Goth et al., 2018a). Although adult studies of LPF factor structure have also been mixed, our findings are consistent with Bliton et al.'s (2021) bifactor models of LPF, which also found essentially unidimensional solutions.

The inadequate model fit obtained from the single factor CFA suggests a need to further scrutinize and revise the LoPF-Q 12–18. It is possible that items contain local dependence, and the factor structure could be improved through Item Response Theory analyses. This approach would also improve clinical utility by shortening the relatively lengthy (97 item) measure. It is also possible that the single-factor CFA was under-powered given the high number of items despite the large sample size ($N = 453$). Regardless, in its current form, we suggest using the total raw score in future research and clinical practice with the LoPF-Q 12–18 to capture a unidimensional severity criterion of personality dysfunction. Doing so is consistent with the description of the LPF as a unidimensional severity criterion in the *DSM-5* AMPD (2013, p. 772), and the theory that guided its development (Morey et al., 2011), as well as its recommendation to use one score to indicate severity in personality dysfunction. Indeed, the *DSM*-5 articulates Criterion A thus:

> Personality disorders are optimally characterized by a generalized personality severity continuum with additional specification of stylistic elements, derived from personality disorder symptom constellations and personality traits. At the same time, the core of personality psychopathology is impairment in ideas and feelings regarding self and interpersonal relationships; this notion is consistent with multiple theories of personality disorder and their research bases. The components of the Level of Personality Functioning Scale-identity, self-direction, empathy, and intimacy are particularly central in describing a personality functioning continuum. (p. 772)

The use of the total score rather than scoring of specific factors is further underscored by the additional effort and complexity that would be required to use a bifactor scoring approach. Of course, more research is needed to replicate these findings and to further examine whether there might be clinical utility in calculating subscale scores, particularly in clinical samples.

Based on the findings of Study 1, we chose to only examine the total score in Study 2 to further examine psychometric properties of the measure in both control and clinical samples. Consistent with our hypothesis and with a unidimensional conceptualization of LPF, internal consistency was excellent in both samples. Regarding demographic variables, analyses revealed interesting relationships between the LoPF-Q 12–18 scores and age. Consistent with our hypotheses, older community control participants had significantly lower LoPF-Q 12–18 scores while older clinical sample participants had higher (though not significant) scores. This pattern was generally mirrored by the other measures of personality impairment: a negative relationship between age and personality impairment in the community control sample and a nonsignificant or positive relationship between age and personality dysfunction in the clinical sample. Control sample results are consistent with longitudinal studies showing normative declines in personality dysfunction in late adolescence and into adulthood (Cohen et al., 2005; Wright, Zalewski, et al., 2016), suggesting youth are successfully consolidating self-other processes. Therefore, higher scores on the LoPF-Q 12–18 in younger healthy children or adolescents may represent not-yet-developed self-other processes, rather than maladaptive self-other function. In contrast, findings in the clinical sample suggest a persistence or exacerbation of early emerging personality pathology, consistent with longitudinal trajectories of youth with clinically significant personality problems (Cohen et al., 2005; Wright, Zalewski, et al., 2016). It is important to note that the community control sample was significantly younger than the clinical sample, and different patterns may be observed if age ranges were otherwise restricted. In addition, while the LoPF-Q 12–18 was developed for age 12–18, it performed well beyond this range in the German study (Goth et al., 2018a; age 11–20), Turkish study (Cosgun et al., 2021; age 11–18) and current study (community control: 10–18, clinical: 11–18). Future studies could use the LoPF-Q 12–18 to investigate age-related differences in personality function using measurement invariance or longitudinal methods.

We also examined the differences in LoPF-Q 12–18 scores based on gender and race or ethnicity. Females had significantly higher total scores in the community control sample but no significant group differences were found in the clinical sample. Although the validation

study of the German LoPF-Q 12–18 (Goth et al., 2018a) did not find significant gender differences, this adds to the mixed literature on gender differences in adolescent BPD, which, in contrast to our findings, has shown higher rates in females in clinical samples but no significant gender differences in a large community sample (Ha et al., 2014; Zanarini et al., 2011). Differences in LoPF-Q 12–18 scores between racial and ethnic groups were present in the community control, but not clinical, sample. In the community control sample, both White and African American groups of youth showed significantly lower levels of LoPF-Q 12–18 personality dysfunction compared with Hispanic/Latin American and multiracial youth. We note that a large proportion of participants were recruited from predominantly Hispanic/Latin American schools.

Results further supported the construct validity of the LoPF 12–18. In both samples, the LoPF-Q 12–18 scores were correlated with identity diffusion as measured by the AIDA (a Criterion A, identity-oriented perspective of personality functioning), BPD features as measured by the BPFS-C-11 (a traditional, *DSM-5* Section II perspective of PD), and the PID-5-BF overall and subscale scores (a Criterion B pathological personality trait perspective). These findings are consistent with the validation study of an interview measure of LPF in adolescents (STiP-5.1, Weekers et al., 2020), which found that LPF dysfunction was associated with number of *DSM-IV* BPD criteria and PID-5-BF total scores, detachment, disinhibition, and psychoticism, though we found stronger and more consistent correlations with PID-5-BF subscales. Construct validity was also supported by significant differences between the control and clinical samples on the total and scale scores, even when controlling for relevant demographic variables.

Finally, ROC analyses demonstrated clinical utility of the total score, with moderate accuracy in determining (a) the control and clinical samples (cut-off = 177.5) and (b) the clinical cut-off of the BPFSC-11 across both samples (cut-off = 176.5). This adds further clarification to the cutoff scores found by the German validation study (180 to differentiate adolescents with a PD from clinical sample without a PD, 163 to differentiate PD from community sample; Goth et al., 2018a) and the Turkish validation study (176 to differentiate adolescents with a PD from school sample; Cosgun et al., 2021). Taken together, we suggest a preliminary cut-off score of 177 for use in future research or in clinical practice to identify adolescents who may need additional assessment and treatment focused specifically on personality functioning.

There were several limitations in the current study. First, given the relatively high number of items on the LoPF-Q 12–18, Study 1 CFAs may have been under-powered despite the large sample size, potentially contributing to the occasional below-threshold fit index for the bifactor model and the inadequate fit indices for the single factor model. In addition, due to lower sample size in the clinical sample (n = 94), we were only able to examine the factor structure of the LoPF-Q 12–18 in the community sample. However, we did not exclude adolescents with psychopathology from this sample and 34.2% of the participants scored above the clinical cut-off on the BPM, indicating that we captured a range of pathology. Moreover, the LPF is intended to capture the full spectrum of personality dysfunction, including healthy function, and therefore should be validated in community samples. However, it is possible that there are differences in the LPF construct depending

on the clinical nature of the sample. For instance, we found different relations between personality functioning and age depending on sample. On a related note, our demographic comparisons are limited by sample size in the clinical sample and provide only preliminary information. The examination of measurement invariance of the LoPF-Q 12–18 regarding age, gender, race, and ethnicity is worthy of a full, separate investigation. Researchers may also consider using latent growth curve models or other longitudinal methods to examine changes in LPF with age using the LoPF-Q 12–18. However, we believe that it is important to present our preliminary findings nonetheless given that this is the first examination of LPF in a diverse sample of youth

Another limitation is that we were not able to examine relationships with other measures of LPF. However, we did find significant associations with PID-5-BF pathological personality traits, although this measure was not developed for youth. Researchers should consider further examining construct validity using interview measures of LPF such as the STiP-5, which has shown preliminary evidence of reliability and validity in adolescents (Weekers et al., 2020), and using parent-report measures. This is also important considering that all measures in the current study were self-report, which inflates correlations due to shared method variance. Future studies should also examine whether the LoPF-Q 12–18 predicts other important outcomes such as other PDs, psychosocial functioning, suicidality, or treatment response. Another important future direction is to develop and evaluate a short version of LoPF-Q 12–18 to improve efficiency.

Despite these limitations, our study makes important contributions to the literature as the first psychometric evaluation of the LoPF-Q 12–18 in an English-speaking and ethnically diverse sample of youth, one of only four studies (Cosgun et al., 2021; Goth et al., 2018a; Weekers et al., 2020) to examine LPF in adolescents, and the first to evaluate a bifactor structure of LPF in adolescents. Although the hypothesized bifactor model exhibited adequate fit, the small amount of variance that could be attributed to multidimensionality does not justify the calculation and use of subscale scores as unique constructs. Results suggested an essentially unidimensional structure, aligning with the *DSM-5* (2013) operationalization of Criterion A as a unidimensional severity criterion, and consistent with theory behind the development of the LPF (Morey et al., 2011), and supporting the use of the total score in research and clinical practice. Moreover, Study 2 analyses revealed adequate internal consistency, construct validity, clinical utility across community and clinical samples. Although further work is needed to refine and further evaluate the LoPF-Q 12–18 and determine the clinical utility of the subscales, our findings support the use of the total score as a valid and reliable assessment tool in future research and to aid in the early detection and intervention with youth with personality problems.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: F31 MH113291 (PI: Vanwoerden)

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (Eds.). (2014). Standards for educational and psychological testing. American Educational Research Association.

American Psychiatric Association. (1994). Diagnostic and statistical manual of mental disorders (4th ed.).

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). American Psychiatric Publishing.

Anderson JL, Sellbom M, & Salekin RT (2018). Utility of the Personality Inventory for DSM-5 Brief Form (PID-5-BF) in the measurement of maladaptive personality and psychopathology. Assessment, 25(5), 596–607. 10.1177/1073191116676889 [PubMed: 27827808]

Bach B, & Hutsebaut J.(2018). Level of Personality Functioning Scale–Brief Form 2.0: Utility in capturing personality problems in psychiatric outpatients and incarcerated addicts. Journal of Personality Assessment, 100(6), 660–670. 10.1080/00223891.2018.1428984 [PubMed: 29494782]

Bach B, Maples-Keller JL, Bo S, & Simonsen E.(2016). The alternative DSM–5 personality disorder traits criterion: A comparative examination of three self-report forms in a Danish population. Personality Disorders: Theory, Research, and Treatment, 7(2), 124–135. 10.1037/per0000162

Barkauskienė R, & Skabeikytė G.(2020). Culture-adapted version Lithuanian of the self-report questionnaire LoPF-Q 12–18 (Levels of Personality Functioning Questionnaire; authors Goth & Schmeck) - Short manual. Academic-Tests.

Bliton CF, Roche MJ, Pincus AL, & Dueber D.(2021). Examining the structure and validity of self-report measures of DSM-5 Alternative Model for Personality Disorders criterion A. Journal of Personality Disorders, 36, Article 531. 10.1521/pedi_2021_35_531

Browne MW, & Cudeck R.(1993). Alternative ways of assessing model fit. In Bollen KA, & Long JS (Eds.), Testing struclural equalion riiodels (pp. 136–162). Sage.

Buer Christensen T, Eikenaes I, Hummelen B, Pedersen G, Nysæter T-E, Bender DS, Skodol AE, & Selvik SG (2020). Level of personality functioning as a predictor of psychosocial functioning: Concurrent validity of criterion A. Personality Disorders: Theory, Research, and Treatment, 11(2), 79–90. 10.1037/per0000352

Busmann M, Wrege J, Meyer AH, Ritzler F, Schmidlin M, Lang UE, Gaab J, Walter M, & Euler S.(2019). Alternative Model of Personality Disorders (DSM-5) predicts dropout in inpatient psychotherapy for patients with personality disorder. Frontiers in Psychology, 10, Article 952. 10.3389/fpsyg.2019.00952

Chanen AM, Berk M, & Thompson K.(2016). Integrating early intervention for borderline personality disorder and mood disorders. Harv Rev Psychiatry, 24(5), 330–341. 10.1097/HRP.0000000000000105 [PubMed: 27144298]

Clark LA (2007). Assessment and diagnosis of personality disorder: Perennial issues and an emerging reconceptualization. Annual Review of Psychology, 58, 227–257. 10.1146/annurev.psych.57.102904.190200

Clark LA, Nuzum H, & Ro E.(2018). Manifestations of personality impairment severity: Comorbidity, course/prognosis, psychosocial dysfunction, and “borderline” personality features. Current Opinion in Psychology, 21, 117–121. 10.1016/j.copsyc.2017.12.004 [PubMed: 29291458]

Cohen P, Crawford TN, Johnson JG, & Kasen S.(2005). The children in the community study of developmental course of personality disorder. Journal of Personality Disorders, 19(5), 466–486. 10.1521/pedi.2005.19.5.466 [PubMed: 16274277]

Cosgun S, Goth K, & Cakiroglu S.(2021). Levels of Personality Functioning Questionnaire (LoPF-Q) 12–18 Turkish Version: Reliability, validity, factor structure and relationship with comorbid psychopathology in a Turkish adolescent sample. Journal of Psychopathology and Behavioral Assessment, 43, 620–631. 10.1007/s10862-021-09867-2

Crick NR, Murray–Close D, & Woods K.(2005). Borderline personality features in childhood: A short-term longitudinal study. Development and Psychopathology, 17(4), 1051–1070. 10.1017/S0954579405050492 [PubMed: 16613430]

Ercegovic N, Paradzik L, Boricevic Marsanic V, & Marcinko D.(2018). Nonsuicidal self-injury and identity development in adolescents. Socijalna Psihijatrija, 46(4), 457–470.

Fan X, Thompson B, & Wang L.(1999). Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. Structural Equation Modeling: A Multidisciplinary Journal, 6(1), 56–83. 10.1080/10705519909540119

Few LR, Miller JD, Rothbaum AO, Meller S, Maples J, Terry DP, Collins B, & MacKillop J. (2013). Examination of the Section III DSM-5 diagnostic system for personality disorders in an outpatient clinical sample. Journal of Abnormal Psychology, 122(4), 1057–1069. 10.1037/a0034878 [PubMed: 24364607]

Fossati A, Krueger RF, Markon KE, Borroni S, Maffei C, & Somma A.(2015). The DSM-5 Alternative Model of Personality Disorders from the perspective of adult attachment: A study in community-dwelling adults. The Journal of Nervous and Mental Disease, 203(4), 252–258. 10.1097/NMD.0000000000000274 [PubMed: 25756706]

Fossati A, Somma A, Borroni S, Markon KE, & Krueger RF (2017). The Personality Inventory for DSM-5 Brief Form: Evidence for reliability and construct validity in a sample of community-dwelling Italian adolescents. Assessment, 24(5), 615–631. 10.1177/1073191115621793 [PubMed: 26676917]

Gamache D, Savard C, Leclerc P, & Côté A.(2019). Introducing a short self-report for the assessment of DSM–5 level of personality functioning for personality disorders: The Self and Interpersonal Functioning Scale. Personality Disorders: Theory, Research, and Treatment, 10(5), 438–447. 10.1037/per0000335

Gander M, Buchheim A, Bock A, Steppan M, Sevecke K, & Goth K.(2020). Unresolved attachment mediates the relationship between childhood trauma and impaired personality functioning in adolescence. Journal of Personality Disorders, 34(Supplement B), 84–103. [PubMed: 31990614]

Gorsuch RL (1983). Factor analysis (2nd ed.). Lawrence Erlbaum Associates.

Goth K, Birkhölzer M, & Schmeck K.(2018a). Assessment of Personality Functioning in Adolescents with the LoPF–Q 12–18 Self-Report Questionnaire. Journal of Personality Assessment, 100(6), 680–690. 10.1080/00223891.2018.1489258 [PubMed: 30907712]

Goth K, Birkhölzer M, & Schmeck K.(2018b). LoPF-Q 12–18 (Levels of Personality Functioning Questionnaire) German version: A self-report questionnaire for measuring personality functioning in adolescence—Short description. Academic-Tests.

Goth K, & Schmeck K.(2018). AIDA (Assessment of Identity Development in Adolescence) German Version: A self-report questionnaire for measuring identity development in adolescence - Short manual. Academic-Tests.

Ha C, Balderas JC, Zanarini MC, Oldham J, & Sharp C.(2014). Psychiatric comorbidity in hospitalized adolescents with borderline personality disorder. Journal of Clinical Psychiatry, 75(5), e457–e464. 10.4088/JCP.13m08696 [PubMed: 24922498]

Hancock GR, & Mueller RO (2001). Rethinking construct reliability within latent variable systems. In Cudeck R, du Toit S, & Sörbom D.(Eds.), Structural equation modeling: Present and future—A Festschrift in honor of Karl Jöreskog (pp. 195–216). Scientific Software International.

Hopwood CJ, Good EW, & Morey LC (2018). Validity of the DSM–5 levels of personality functioning scale–self report. Journal of Personality Assessment, 100(6), 650–659. [PubMed: 29424568]

Hu L, & Bentler PM (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal, 6(1), 1–55. 10.1080/10705519909540118

Hutsebaut J, Feenstra DJ, & Kamphuis JH (2016). Development and Preliminary Psychometric Evaluation of a Brief Self-Report Questionnaire for the Assessment of the DSM–5 level of Personality Functioning Scale: The LPFS Brief Form (LPFS-BF). Personality Disorders: Theory, Research, and Treatment, 7(2), 192–197. 10.1037/per0000159

Hutsebaut J, Kamphuis JH, Feenstra DJ, Weekers LC, & De Saeger H.(2017). Assessing DSM–5-oriented level of personality functioning: Development and psychometric evaluation of the Semi-Structured Interview for Personality Functioning DSM–5 (STiP-5.1). Personality Disorders: Theory, Research, and Treatment, 8(1), 94–101. 10.1037/per0000197

Kassin M, De Castro F, Arango I, & Goth K.(2013). Psychometric properties of a culture-adapted Spanish version of AIDA (Assessment of Identity Development in Adolescence) in Mexico. Child and Adolescent Psychiatry and Mental Health, 7(1), Article 25. 10.1186/1753-2000-7-25

Kassin M, & Hackradt J.(2019). Culture-adapted version Spanish Mexico of the self-report questionnaire LoPF-Q 12–18 (Levels of Personality Functioning Questionnaire; authors Goth & Schmeck) - Short manual. Academic-Tests.

Krueger RF, Kotov R, Watson D, Forbes MK, Eaton NR, Ruggero CJ, Simms LJ, Widiger TA, Achenbach TM, Bach B, Bagby RM, Bornovalova MA, Carpenter WT, Chmielewski M, Cicero DC, Clark LA, Conway C, DeClercq B, DeYoung CG, & Zimmermann J.(2018). Progress in achieving quantitative classification of psychopathology. World Psychiatry, 17(3), 282–293. 10.1002/wps.20566 [PubMed: 30229571]

Lind M, Vanwoerden S, Penner F, & Sharp C.(2019). Inpatient adolescents with borderline personality disorder features: Identity diffusion and narrative incoherence. Personality Disorders: Theory, Research, and Treatment, 10(4), 389–393. 10.1037/per0000338

McAdams DP (2015). The art and science of personality development. Guilford Press.

Morey LC (2017). Development and initial evaluation of a self-report form of the DSM–5 Level of Personality Functioning Scale. Psychological Assessment, 29(10), 1302. [PubMed: 28240933]

Morey LC, Bender DS, & Skodol AE (2013). Validating the proposed diagnostic and statistical manual of mental disorders, 5th Edition, severity indicator for personality disorder. Journal of Nervous & Mental Disease, 201(9), 729–735. 10.1097/NMD.0b013e3182a20ea8

Morey LC, Berghuis H, Bender DS, Verheul R, Krueger RF, & Skodol AE (2011). Toward a model for assessing level of personality functioning in DSM–5, part II: Empirical articulation of a core dimension of personality pathology. Journal of Personality Assessment, 93(4), 347–353. 10.1080/00223891.2011.577853. [PubMed: 22804673]

Muthén BO, Kaplan D, & Hollis M.(1987). On structural equation modeling with data that are not missing completely at random. Psychometrika, 52, 431–462. 10.1007/BF0229436.

Muthén LK, & Muthén BO (2017). Mplus: Statistical Analysis with Latent Variables: User's Guide (Version 8). Los Angeles, CA: Authors.

Ragelienė T, & Justickis V.(2016). Interrelations of adolescent's identity development, differentiation of self and parenting style. Psichologija, 53, 24–43.

Reise SP, Scheines R, Widaman KF, & Haviland MG (2013). Multidimensionality and structural coefficient bias in structural equation modeling: A bifactor perspective. Educational and Psychological Measurement, 73, 5–26. 10.1177/001316441244983.

Rodriguez A, Reise SP, & Haviland MG (2016). Applying bifactor statistical indices in the evaluation of psychological measures. Journal of Personality Assessment, 98(3), 223–237. 10.1080/00223891.2015.1089249 [PubMed: 26514921]

Sharp C.(2020). Adolescent personality pathology and the Alternative Model for Personality Disorders: Self development as nexus. Psychopathology, 53(3), 198–204. [PubMed: 32464626]

Sharp C, & Fonagy P.(2015). Practitioner Review: Borderline personality disorder in adolescence–recent conceptualization, intervention, and implications for clinical practice. Journal of Child Psychology and Psychiatry, 56(12), 1266–1288. [PubMed: 26251037]

Sharp C, Kerr S, & Chanen A.(2021). Early identification and prevention of personality pathology: An AMPD informed model of clinical staging. In Skodol A, & Oldham J.(Eds.), The American Psychiatric Association textbook of personality disorders (pp. 285–337). American Psychiatric Association.

Sharp C, & Miller J.(2022). Ten-year retrospective on the DSM–5 Alternative Model of Personality Disorder: Seeing the forest for the trees. Personality Disorders: Theory, Research and Treatment, 13(4), 301–304.

Sharp C, Steinberg L, Temple J, & Newlin E.(2014). An 11-item measure to assess borderline traits in adolescents: Refinement of the BPFSC using IRT. Personality Disorders: Theory, Research, and Treatment, 5(1), 70–78. 10.1037/per0000057

Sharp C, & Vanwoerden S.(2018). Culture-adapted version English USA of the self-report questionnaire LoPF-Q 12–18 (Levels of Personality Functioning Questionnaire; authors Goth & Schmeck)—Short manual. Academic-tests.

Sharp C, Vanwoerden S, Schmeck K, Birkhölzer M, & Goth K.(2021). An evaluation of age-group latent mean differences in maladaptive identity in adolescence. Frontiers in Psychiatry, 12, Article 730415. 10.3389/fpsyt.2021.730415

Sharp C, & Wall K.(2017). Personality pathology grows up: Adolescence as a sensitive period. Current Opinion in Psychology, 21, 111–116. 10.1016/j.copsyc.2017.11.010 [PubMed: 29227834]

Sharp C, & Wall K.(2021). DSM-5 level of personality functioning: Refocusing personality disorder on what it means to be human. Annual Review of Clinical Psychology, 17(1), 313–317. 10.1146/annurev-clinpsy-081219-105402

Sharp C, Wright AGC, Fowler JC, Frueh BC, Allen JG, Oldham J, & Clark LA (2015). The structure of personality pathology: Both general ("g") and specific ("s") factors? Journal of Abnormal Psychology, 124(2), 387–398. 10.1037/abn0000033 [PubMed: 25730515]

Shi D, Lee T, & Maydeu-Olivares A.(2019). Understanding the model size effect on SEM fit indices. Educational and Psychological Measurement, 79(2), 310–334. 10.1177/0013164418783530 [PubMed: 30911195]

Siefert CJ, Sexton J, Meehan K, Nelson S, Haggerty G, Dauphin B, & Huprich S.(2020). Development of a short form for the DSM–5 levels of Personality Functioning Questionnaire. Journal of Personality Assessment, 102(4), 516–526. 10.1080/00223891.2019.1594842. [PubMed: 31107606]

Sleep CE, Lynam DR, Widiger TA, Crowe ML, & Miller JD (2019). An evaluation of DSM–5 Section III personality disorder Criterion A (impairment) in accounting for psychopathology. Psychological Assessment, 31(10), 1181–1191. 10.1037/pas0000620 [PubMed: 31580131]

Tardivo LSPC, Miura P, & Alhanat M.(2016). Family violence against children and adolescents in Brazil and relations with identity development. Acta Psychopathologica, 2(4), Article 35. 10.4172/2469-6676.100061

Vanwoerden S, Garey L, Ferguson T, Temple JR, & Sharp C.(2019). Borderline Personality Features Scale for Children-11: Measurement invariance over time and across gender in a community sample of adolescents. Psychological Assessment, 31(1), 114–119. [PubMed: 30080065]

Weekers LC, Hutsebaut J, & Kamphuis JH (2019). The Level of Personality Functioning Scale-Brief Form 2.0— Update of a brief instrument for assessing level of personality functioning: The Level of Personality Functioning Scale— Brief Form 2.0. Personality and Mental Health, 13(1), 3–14. 10.1002/pmh.1434 [PubMed: 30230242]

Weekers LC, Verhoeff SCE, Kamphuis JH, & Hutsebaut J.(2020). Assessing Criterion A in adolescents using the Semistructured Interview for Personality Functioning DSM–5. Personality Disorders: Theory, Research, and Treatment, 12, 312–319. 10.1037/per0000454

Wright AGC, Hopwood CJ, Skodol AE, & Morey LC (2016). Longitudinal validation of general and specific structural features of personality pathology. Journal of Abnormal Psychology, 125(8), 1120–1134. 10.1037/abn0000165 [PubMed: 27819472]

Wright AGC, Zalewski M, Hallquist MN, Hipwell AE, & Stepp SD (2016). Developmental trajectories of borderline personality disorder symptoms and psychosocial functioning in adolescence. Journal of Personality Disorders, 30(3), 351–372. 10.1521/pedi_2015_29_200 [PubMed: 26067158]

Zanarini MC, Horwood J, Wolke D, Waylen A, Fitzmaurice G, & Grant BF (2011). Prevalence of DSM-IV borderline personality disorder in two community samples: 6,330 English 11-year-olds and 34,653 American adults. Journal of Personality Disorders, 25(5), 607–619. 10.1521/pedi.2011.25.5.607 [PubMed: 22023298]

Table 1.

Study 2 Sample Characteristics and Group Comparisons.

	Healthy (N = 298)		Clinical (N = 94)		
	n* or *M	***SD* or %**	***n* or *M***	***SD* or %**	**Group differences**
Age	13.08	2.09	14.39	1.74	$t(390) = -5.51^{***}$
Gender (female)	162	54.4%	55	58.5%	$\chi^2 = 8.23^{*}$
Race\ethnicity					$\chi^2 = 32.05^{**}$
Asian	17	5.7%	3	3.2%	
Black	40	13.4%	15	16.0%	
Caucasian	32	10.7%	27	28.7%	
Hispanic	171	57.4%	39	41.5%	
Native American	4	1.3%	5	5.3%	
Multi-racial/other	34	11.4%	5	5.3%	
LoPF total	139.14	47.92	216.37	65.52	$t(125.91) = -10.57^{***}$
AIDA diffusion	93.32	33.95	135.29	42.90	$t(390) = -9.78^{***}$
BPFSC-11 total	27.10	7.61	30.31	9.34	$t(132.33) = -3.02^{**}$
PID-5-BF average	.85	.50	1.14	.59	$t(144.79) = -4.21^{**}$
Negative affect	1.03	.76	1.54	.90	$t(145.21) = -4.87^{***}$
Detachment	.82	.69	1.02	.70	$t(343) = -2.53^{*}$
Antagonism	.53	.55	.54	.58	$t(337) = -.24$
Disinhibition	.80	.65	1.08	.75	$t(255) = -3.69^{**}$
Psychoticism	1.05	.70	1.51	.89	$t(138.58) = -4.48^{***}$

Note. AIDA = Assessment of Identity Development in Adolescents; BPFSC-11 = Borderline Personality Features Scale for Children-11; PID-5-BF = Personality Inventory for the *DSM-5* Brief Form.

$^{*}p < .05$.

$^{**}p < .01$.

$^{***}p < .001$.

Assessment. Author manuscript; available in PMC 2023 September 01.

Table 2.

Study 2 Pearson's Correlations Among Continuous Variables (Control Sample on Top, Clinical Sample on Bottom).

		1	2	3	4	5	6	7	8	9	10	11
1	LoPF-Q 12–18	1	.84**	.81**	.66**	.71**	.45**	.62**	.47**	.55**	.62**	−.35**
2	AIDA diffusion	.95**	1	.97**	.68**	.57**	.41**	.44**	.33**	.46**	.54**	−.48**
3	AIDA diff. (no LoPF)	.91**	.98**	1	.71**	.59**	.45**	.43**	.33**	.47**	.56**	−.40**
4	BPFSC-11 Total	.62**	.63**	.62**	1	.67**	.57**	.46**	.35**	.52**	.60**	−.13*
5	PID-5-BF Average	.73**	.69**	.67**	.70**	1	.76**	.75**	.62**	.80**	.83**	−.20**
6	Negative affect	.67**	.67**	.67**	.75**	.84**	1	.40**	.27**	.49**	.56**	−.03
7	Detachment	.53**	.50**	.48**	.54**	.74**	.60**	1	.47**	.47**	.56**	−.18**
8	Antagonism	.38**	.33**	.30**	.30**	.56**	.22**	.29**	1	.38**	.34**	−.10
9	Disinhibition	.54**	.51**	.50**	.58**	.84**	.65**	.48**	.47**	1	.62**	−.26**
10	Psychoticism	.64**	.58**	.56**	.48**	.83**	.62**	.48**	.34**	.62**	1	−.20**
11	Age	.18	.12	.09	.28**	.27**	.28*	.22*	.21*	.19*	.14	1

Note. AIDA = Assessment of Identity Development in Adolescents; BPFSC-11 = Borderline Personality Features Scale for Children-11; PID-5-BF = Personality Inventory for the *DSM-5* Brief Form.

* $p < .05$.

** $p < .01$.