Development of a Short Form for the *DSM–5* Levels of Personality Functioning Questionnaire

Caleb J. Siefert, Jim Sexton, Kevin Meehan, Sharon Nelson, Greg Haggerty, Barry Dauphin & Steven Huprich

Published online: 20 May 2019.

Submit your article to this journal ⍉

View related articles ⍉

View Crossmark data ⍉

Routledge
Taylor & Francis Group

Check for updates

Development of a Short Form for the *DSM–5* Levels of Personality Functioning Questionnaire

Caleb J. Siefert[1], Jim Sexton[2], Kevin Meehan[3], Sharon Nelson[4], Greg Haggerty[5], Barry Dauphin[6], and Steven Huprich[6]

[1]Department of Psychology, University of Michigan–Dearborn; [2]Department of Professional Psychology, George Washington University; [3]Department of Psychology, Long Island University; [4]Department of Psychology, Eastern Michigan University; [5]Department of Psychiatry, Nassau University Medical Center; [6]Department of Psychology, University of Detroit–Mercy

ABSTRACT

The *Diagnostic and Statistical Manual of Mental Disorders* (5th ed. [*DSM–5*]; American Psychiatric Association, 2013) introduced the Alternative *DSM–5* Model for Personality Disorders (AMPD). Criterion A of the AMPD conceptualizes level of personality functioning (LOPF) in terms of self- and interpersonal functioning. This article describes the development of a short form for the *DSM–5* Level of Personality Functioning Questionnaire (DLOPFQ). A sample of 1,279 participants was drawn from community, clinical, and college settings. All participants completed the DLOPFQ full form. The sample was split into a derivation sample ($n = 640$) and a validation sample ($n = 639$). Exploratory factor analysis of the derivation sample data was used to select short-form items. Using the validation sample, confirmatory factor analyses (CFAs) were used to assess fit for proposed item-to-subscale assignments. Short-form subscales had good internal consistency estimates, correlated strongly with full-form subscales, correlated with one another, and were associated with relevant constructs. CFA supported a second-order factor model (i.e., four factors loading onto a higher order LOPF factor). Overall, the DLOPFQ Short Form provides a brief assessment of the constructs measured by the full form. Limitations of the study are reviewed, speculations for improving the measure are discussed, and suggestions for future directions are provided.

Huprich and colleagues (2018) presented a new self-report measure for assessing level of personality functioning (LOPF), the *DSM–5* Levels of Personality Functioning Questionnaire (DLOPFQ). The DLOPFQ is based on the Alternative Model for Personality Disorders (AMPD) described in Section III of the *Diagnostic and Statistical Manual of Mental Disorders* (5th ed. [*DSM–5*]; American Psychiatric Association, 2013). It provides an overall LOPF score, scores for self- and interpersonal functioning, and element subscales assessing problems with identity, self-direction, empathy, and intimacy. Initial studies supported the psychometric adequacy, construct validity, and test–retest reliability of the DLOPFQ (Huprich et al., in press; Sexton et al., 2019). The DLOPFQ contains 132 items, and this length might pose problems in some settings. This article describes the development of a short form for the DLOPFQ. The goal was to develop a short form with the same structure as the full form (i.e., four element subscales that contribute to an overall index of LOPF) and that showed strong agreement with full-form scores.

The Alternative *DSM–5* Model for Personality Disorders

Prior to the publication of *DSM–5*, the Personality and Personality Disorder Work Group (PPDWG) revised the manual's approach to personality and personality disorders. They developed and proposed the AMPD, which is now included in Section III of the *DSM–5* (Skodol et al., 2011). Criterion A of the AMPD conceptualizes LOPF in terms of self- and interpersonal functioning (Morey et al., 2011). Each functional domain is broken into specific elements; self-functioning includes the identity and self-direction elements and interpersonal functioning contains the empathy and intimacy elements. Although somewhat unique in how they are characterized, these elements are considered reciprocally influential and interdependent (e.g., one's capacity for intimacy is likely affected by one's ability to engage in empathy). Criterion B of the AMPD includes five independent maladaptive traits: negative affectivity, detachment, antagonism, disinhibition, and psychoticism. Each trait can also be further broken into facets.

Clinicians are advised to assess LOPF and maladaptive traits separately, although they are expected to be related (e.g., high levels of maladaptive traits are expected among those with more problematic LOPF; Hopwood, Good, & Morey, 2018). There are practical reasons for this recommendation. For example, different constellations of maladaptive traits can produce similar levels of personality dysfunction, but might create this dysfunction in different ways. A client with notable issues with disinhibition and a

client with a pathologically negative self-image might both experience similar levels of interpersonal functioning deficits, but for different reasons. Likewise, similar trait profiles could result in different areas of dysfunction. One client with negative affectivity might experience dysfunction primarily in his sense of self (e.g., identity problems), whereas another routinely withdraws from close relationships (e.g., intimacy problems). To ascertain if differentiating the assessment of problematic traits from LOPF is useful, further research is necessary (Few et al., 2013). Strong measures are needed for such research.

At present, multiple measures for assessing the maladaptive personality traits exist (Bagby & Widiger, 2018). There is also growing consensus that the Personality Inventory for the *DSM–5* (PID–5; Krueger, Derringer, Markon, Watson, & Skodol, 2012) robustly assesses the maladaptive traits. PID–5 profiles have been linked to specific personality disorders and functional problems (Keeley, Flanagan, & McCluskey, 2014). Fewer options exist to assess LOPF. This could be due to the fact that the LOPF framework is "younger" (Krueger, Hopwood, Wright, & Markon, 2014) relative to the maladaptive traits, which are variations of the traditional Big Five traits (Trull & Widiger, 2013). Although fewer in number, rating scales and self-report measures for assessing LOPF have shown promise.

Assessing level of personality functioning

Given the clinical research tradition on which Criterion A is based, it is of little surprise that many early LOPF assessment tools are clinician rated (see Bender, Morey, & Skodol, 2011; Zimmerman et al., 2015). The Level of Personality Functioning Scale (LPFS; Morey, 2017) was one of the first clinician-rated approaches to assess LOPF using the AMPD model. Other clinician-rated LOPF assessments include the Clinical Assessment of Level of Personality Functioning scale (Thylstrup et al., 2016), the Semi-Structured Interview for Personality Functioning (Hutsebaut, Kamphuis, Feenstra, Weekers, & De Saeger, 2017), the SCID–5–AMPD Level of Personality Functioning Scale (Module 1; Christensen et al., 2018), and methods for rating LOPF based on client interviews are also available (see Zimmerman et al., 2015). Studies support the psychometric adequacy, validity, and clinical utility of these rating methods (Bach, Markon, Simonsen, & Krueger, 2015; Zimmerman et al., 2015). Given the value of multimethod personality assessment (Bornstein & Hopwood, 2014; Meyer et al., 2001), self-report LOPF measures to complement clinical rating systems have been developed.

Early self-report LOPF measures drew from clinician-rated counterparts. Morey (2017) developed the 80-item Levels of Personality Functioning Scale–Self-Report (LPFS–SR) by adapting the items of the clinician-rated LPFS. The LPFS–SR produces scores for each of the four elements. These scales show good internal consistency, have strong associations with one another, and converge with measures of personality functioning (Hopwood et al., 2018). Huprich et al. (2018) recently developed an alternative LOPF self-report measure. This measure, the DLOPFQ, also produces an overall LOPF score and four subscales that

correspond to the LOPF elements. Evidence for the internal consistency and test–retest reliability have been demonstrated for the DLOPFQ subscales (Huprich et al., 2018; Sexton et al., 2019). Studies suggest that DLOPFQ subscales incrementally increase the prediction of variance in relevant personality functioning variables (e.g., interpersonal functioning; attachment status) beyond PID–5 scales in clinical and community samples (Huprich et al., 2018; Sexton et al., 2019).

The length of the LPFS–SR and the DLOPFQ might limit their utility in some settings. Fortunately, brief LOPF measures have also been developed. Prior to the publication of the LPFS–SR, Hutsebaut, Feenstra, and Kamphuis (2016) created the Levels of Personality Functioning–Brief Form (LPFS–BF). Items for this 12-item, self-report questionnaire were based on the clinician-rated LPFS (Bender, Morey, & Skodol, 2011). Roche, Jacobson, and Pincus (2016) also developed a 12-item LOPF scale based on the LPFS, the Alternative Model for Personality Disorders –Criterion A Scale (AMPD–CAS). Both the LPFS–BF and AMPD–CAS can be scored to produce a global score and scores for self- and interpersonal functioning. The AMPD–CAS also provides scores for the four elements (i.e., identity, self-direction, empathy, and intimacy). Internal consistency estimates for both the LPFS–BF and AMPD–CAS have generally been within acceptable ranges and there is some support for each measure's validity. LOPF–BF scores successfully discriminate respondents with and without personality disorders and correlate with indicators of personality pathology (Hopwood et al., 2018). AMPD–CAS scores are linked to variations in day-to-day functioning over time (Roche et al., 2016) and to patterns of interpersonal deficits and sensitivities (Dowgwillo, Roche, & Pincus, 2018). The latter study also supported the utility of considering LOPF score at multiple levels. For example, the global AMPD–CAS score was associated with overall interpersonal problems, whereas specific element scales yielded mildly differentiated associations with specific interpersonal deficits and sensitivities (e.g., problems with empathy were associated with greater interpersonal vindictiveness and aversion to dependence). Findings for both suggest that brief measures can be useful in assessing LOPF.

Currently, the DLOPFQ has no short form. The full form requires respondents to complete 132 items, which might not be practical for screening in health settings, for studies containing several measures, or in clinical settings where less burdensome alternatives are desired. This article describes the development of a DLOPFQ short form (DLOPFQ–SF). Our goal was to develop a shorter measure that was not context specific. Like the AMPD–CAS, we developed four element subscales, corresponding to the respective Criterion A elements, that all contribute to an overall LOPF index. Procedures for developing the DLOPFQ–SF are detailed next.

Method

Participants

Community sample

A total of 555 adults were recruited through Amazon's Mechanical Turk (MTurk). Of these, 500 participants were

retained and 55 were removed due to failure to complete measures or poor performance on data quality checks. All community participants resided within the United States and were over 18 years of age (M age = 24.27 years [$SD = 10.78$]). Regarding gender, 55.1% were men and 44.9% were women. Regarding ethnicity, 76.7% identified as White or Caucasian; 7.6% identified as Black or African American; 7.4% identified as Latino, Latina, or Hispanic; 6.3% identified as Asian or Asian American; and 2.0% identified as other. All participants in the community sample were paid $5 for participation.

Clinical sample

Clinical participants were recruited from an outpatient psychiatric clinic and an internal medicine clinic at a large midwestern hospital. Patients with a known organic cause for mental health issues (e.g., stroke, head trauma) and patients with a psychotic-spectrum disorder (e.g., schizophrenia, schizoaffective disorder) were not eligible for participation. A total of 184 patients agreed to participate in a study of personality and health. Of these, 135 participants completed the measures (26.5% men, 73.5% women; $n = 83$ from outpatient psychiatry, $n = 52$ from internal medicine). Ages ranged from 18 to 79, with a mean of 47.77 ($SD = 13.37$). Regarding ethnicity, 76.5% identified as White or Caucasian; 12.5% identified as Black or African American, 2.9% identified as Latino, Latina, or Hispanic; 1.5% identified as Native American; and 6.6% identified as other. Participants were paid $25 to compensate them for their participation.

College sample 1

A sample of 339 college students (22.4% men, 77.6% women) from a large eastern university agreed to participate in a study of personality and functioning. All participants were over the age of 18 (M age = 20.08 [$SD = 3.98$]). Regarding ethnicity, 15.2% identified as White or Caucasian; 24.2% identified as Black or African American; 11.5% identified as Latino, Latina, or Hispanic; 27.9% identified as Asian or Asian American; 9.1% identified as Middle Eastern or Arab American; and 12.1% identified as other. Participants were given course credit for participation.

College sample 2

A total of 305 college students (42.3% men; 57.7% women) at a midwestern university agreed to participate in a study of personality and well-being. Ages ranged from 18 to 50, with an average age of 20.12 ($SD = 3.99$). Regarding ethnic identification, 54.4% identified as White or Caucasian; 8.2% identified as Black or African American; 6.2% identified as Latino, Latina, or Hispanic; 8.2% identified as Asian or Asian American; 20.7% identified as Middle Eastern or Arab American; and 2.3% identified as other. Course credit was given for participation.

Derivation and validation samples

We pooled all participants into a single file ($N = 1,279$) and randomly split the group into a derivation sample ($n = 640$; 40% from the community sample, 10% from the clinical sample, 25% from College Sample 1, and 25% from College Sample 2) and a validity sample ($n = 639$; 38% from the community sample, 11% from the clinical sample, 28% from College Sample 1, and 23% from College Sample 2). Samples did not differ significantly in terms of age, gender makeup, ethnic identification composition, or scores on the DLOPFQ.

Measures

DSM–5 Levels of Personality Functioning Questionnaire

The DLOPFQ contains 66 items that respondents rate on a Likert scale ranging from 1 (*strongly disagree*) to 6 (*strongly agree*). Respondents rate statements based on functioning in work and school settings and again based on functioning in social and personal contexts (a total of 132 items). Items are worded identically for both contexts. Scores can be generated for an overall LOPF (sum of all items) and four subscales (identity problems, self-direction problems, empathy problems, and intimacy problems) can be generated for both contexts. Further, an overall score can be calculated, as can scores for self- and interpersonal problems. Higher scores indicate greater dysfunction. Prior studies report acceptable internal consistency (i.e., coefficient alphas ranging from .77–.94; Huprich et al., 2018) and test–retest reliability over a 1-month period (Sexton et al., 2019).

The Relationship Questionnaire

The Relationship Questionnaire (RQ; Bartholomew & Horowitz, 1991) is an adult attachment measure. Respondents are given four paragraphs, each of which describes an attachment style (i.e., secure, fearful-avoidant, preoccupied, dismissive-avoidant), and they rate each paragraph on a 1 (*not at all like me*) to 7 (*very much like me*) scale for how well it describes their relationships. RQ ratings have shown good test–retest reliability and construct validity in prior studies (Scharfe & Bartholomew, 1994).

The Swartz Outcome Scale–10

The Swartz Outcome Scale–10 (SOS–10; Blais et al., 1999) is composed of 10 items tapping subjective well-being (e.g., I am generally satisfied with my psychological health; I have confidence in my ability to sustain important relationships). Respondents rate items on a 7-point Likert scale ranging from 0 (*never*) to 6 (*all the time or nearly all the time*). Items are summed to generate a total score. Higher scores indicate more positive perceptions of one's quality of life. Studies support the unidimensional nature and internal consistency of the SOS–10 (Blais et al., 1999), and scores have been linked to life outcomes in college samples (e.g., Haggerty, Blake, Naraine, Siefert, & Blais, 2010).

The Inventory of Interpersonal Problems–Short Circumplex

The Inventory of Interpersonal Problems–Short Circumplex (IIP–SC; Horowitz, Alden, Wiggins, & Pincus, 2000) contains 32 items and assesses relational problems. Items are grouped into things the respondent finds "hard to do" (e.g., Be assertive with another person) and things the respondent "does too much" (e.g., I open up to people too much). Items are rated on a scale ranging from 0 (not at all) to 4 (extremely). The IIP–SC produces an overall score, scores for the eight subscales of the Interpersonal Circumplex (IPC), and scores for the interpersonal warmth and dominance dimensions of the IPC. The IIP–SC's psychometric properties and scale structure have been found to be strong in college samples (Hopwood, Pincus, DeMoor, & Koonce, 2008).

In addition to calculating a total score (i.e., sum of all items), we used formulas by Ruiz et al. (2004, p. 218) to generate scores for interpersonal warmth and dominance. Interpersonal dominance is calculated as follows: $.25 \times [\text{Domineering}_{PA} - \text{Nonassertive}_{HI} + .71 \times (\text{Intrusive}_{NO} + \text{Vindictive}_{BC} - \text{Socially Inhibited}_{FG} - \text{Overly Accommodating}_{JK})]$. Interpersonal warmth is calculated as follows: $.25 \times [\text{Self-Sacrificing}_{LM} - \text{Cold}_{DE} + .71 \times (\text{Intrusive}_{NO} - \text{Vindictive}_{BC} - \text{Socially Inhibited}_{FG} + \text{Overly Accommodating}_{JK})]$. We also used formulas from Ruiz and colleagues to calculate IPC quadrant scores for Friendly-Dominance ($\text{Intrusive}_{NO} + [.707 \times \text{Domineering}_{PA}] + [.707 \times \text{Self-Sacrificing}_{LM}]$), Friendly-Submissive ($\text{Overly Accommodating}_{JK} + [.707 \times \text{Self-Sacrificing}_{LM}] + [.707 \times \text{Nonassertive}_{HI}]$), Hostile-Submissive ($\text{Socially Inhibited}_{FG} + [.707 \times \text{Nonassertive}_{HI}] + [.707 \times \text{Cold}_{DE}]$), and Hostile-Dominant ($\text{Vindictive}_{BC} + [.707 \times \text{Domineering}_{PA}] + [.707 \times \text{Cold}_{DE}]$).

Procedure

Data collection

Data collection methods differed by sample. Community sample participants learned of the study through MTurk. If interested, they clicked a link for a Qualtrics survey, reviewed a consent form, and, if consent was given, completed a series of self-report questionnaires (including the DLOPFQ). Participants in the clinical sample were initially approached by health care providers. Those interested met with a study team member, reviewed a consent form, and were given access to an online survey. For both college samples, participants learned of the study through a university subject pool system and completed all survey measures online after giving consent.

Short-form item selection

Items for the DLOPFQ–SF were selected using a mixed approach involving exploratory factor analysis (EFA), classical test theory, and expert input. For item selection, all analyses were conducted using the derivation sample only. As part of efforts to eliminate the need for specific contexts, we examined correlation coefficients for respective item

pairs across contexts (e.g., the association between Item 1 for the work or school context and Item 1 for the social or personal context). Six items with small cross-context correlations were removed, leaving a pool of 60 items.

We next examined item skew and found very little evidence of skew. For the items rated in the work or school context, absolute values for item skew ranged from 0.01 to 1.08, with a mean of 0.36 ($SD = 0.28$). In the social or personal context, absolute values for item skew ranged from 0.01 to 1.11, with a mean of 0.37 ($SD = 0.29$). Thus, we employed maximum likelihood (ML) extraction with direct oblimin rotation (Schmitt, Sass, Chappelle, & Thompson, 2018). While removing items, we sought a simple structure within the pattern matrix.[1] We defined a simple structure as follows: (a) each factor having at least three items, (b) all item loadings being interpretable, and (c) a dearth of problematic cross-loading. Based on Tabachnick and Fidell (2007), loadings of < .40 were considered uninterpretable, those of .40 were considered adequate, those of .50 to .70 were considered good, and those greater than .70 were considered excellent. Cross-loadings were considered problematic when an item's largest loading was within .20 of its second largest loading. Although we preferred items with smaller cross-loadings in the pattern matrix, we anticipated cross-loadings within the structural matrix because full-form subscales are correlated (Huprich et al., 2018). This overlap is consistent with the AMPD, which views element domains as interpenetrating and nonindependent. By seeking a simple structure within the pattern matrix, we sought to limit item redundancy across subscales, while retaining items that would collectively provide an index of LOPF.

A series of EFAs were conducted using the 60 work or school items and the 60 social or personal items. Each series had two phases. The first phase sought to identify core items. To do this, we entered our 60 items and removed items one at a time. Items with no interpretable loadings were first removed. When no such items remained, we removed items with problematic cross-loadings in the pattern matrix. We treated the remaining items as core items. The second phase examined if additional items would now produce interpretable loadings. To explore this, we readded previously removed items one by one. When a newly added item showed a primary loading with one of the factors and no problematic cross-loadings, the item was retained. If the later addition of an item resulted in a newly retained item becoming uninterpretable or generating a problematic cross-loading, we addressed this by rerunning the model with the newly added item included and the newly retained item removed. If the newly added item generated a larger loading on a factor compared to the newly retained item (i.e., the

---

[1] Because LOPF is conceptualized as a higher order factor with interpenetrating, nonindependent element domains, one could argue that we should have sought items with high cross-loadings. Although we sought a simple structure in the pattern matrix, we anticipated higher cross-loadings for selected items within the structural matrix, as DLOPFQ full-form subscales are correlated. This is precisely what occurred. Thus, this approach balanced our desire to minimize item redundancy across conceptually distinct subscales, while producing subscales that are nonetheless correlated and contribute to a global index of LOPF.

item just removed), we retained the new item and continued on. If the newly added item became uninterpretable, generated a problematic cross-loading, or generated a primary loading that was smaller than the removed item, we removed the newly added item, readded the original item, and continued. This process was repeated until all items were given multiple chances to enter the models.

After this process was completed for both item contexts, we stopped our use of EFA and compared models. Subscales ranged from seven possible items (self-direction problems) to 10 possible items (identity problems). To reduce the number of items down to six per subscale we reviewed item content for redundancy. For example, although Items 15 and 16 both had strong loadings, we removed Item 16 (When people ask what my future plans are, I have a difficult time answering and often feel that it doesn't really matter) due to redundancy with Item 15 (Unlike other people, I don't have a sense of what my future plans are). Based on this approach, 24 items were selected for the DLOPFQ–SF; six items for the identity problems subscale (6, 7, 9, 10, 15, and 25), six items for the self-direction subscale (4, 17, 21, 27, 28, and 30), six items for the empathy problems subscale (8, 22, 32, 41, 43, and 61), and six for the intimacy problems subscale (12, 48, 49, 51, 52, and 54). Selected items are shown in the Appendix. As anticipated, these items generally showed a simple structure in the pattern matrix, but cross-loaded in the structural matrix. Ultimately, factors were significantly correlated.

Final subscale analyses

Given that the items selected for the DLOPFQ–SF were highly correlated across item contexts (average correlation of .67 [$SD = .06$]), we averaged each item's ratings across context. For example, the score for the work or school Item 6 was averaged with the score for the social or personal Item 6 (context-specific EFAs are available from the first author on request). All items not selected for inclusion in the DLOPFQ–SF were removed. We conducted parallel analysis (PA) using O'Connor's (2000) SPSS syntax. We used permutations of the derivation sample to generate 1,000 samples. Actual eigenvalues were compared with the 99th percentile estimates from the permutation samples. We explored models with the number of factors suggested by PA and models with two additional and two less factors (Osborne, 2014; Wright, 2017). We then examined item-to-scale assignments with confirmatory factor analysis (CFA) based on the data from the validation sample.

CFA models were generated using ML estimation with the Analysis of Moment Structures (AMOS 22) software. To evaluate fit, we considered the root mean square error of approximation (RMSEA) and two incremental fit indexes: the nonnormed fit index (NNFI) and the comparative fit index (CFI). We adopted fit criteria common in psychological research (see Byrne, 2010; Jackson, Gillaspy, & Purc-Stephenson, 2009; Kline, 2005). Specifically, CFI and NNFI values $\geq$ .90 indicate adequate fit and values > .95 indicate good fit. RMSEA values of < .10 indicate adequate fit and those of < .06 indicate good fit (Schmitt et al., 2018). When

respecifying nested models, we considered a respecified model an improvement if there was a statistically significant reduction in the model chi-square (i.e., $\Delta\chi^2$) and CFI increased by at least .01 (Kline, 2005). For nonnested respecified models, we considered only the CFI increase.

Model comparisons

CFA was used to examine fit for four potential models: a one-factor model (all items assigned to a single factor), a two-factor model (identity problems and self-direction problem items assigned to a factor; empathy problems and self-direction problems assigned to a factor), a four-factor model (all items assigned to factors for theorized subscales), and a second-order model (all items assigned to factors for theorized subscales that were then assigned to a second-order factor). Factors were allowed to correlate when models contained multiple factors (except in the second-order factor model). To evaluate model fit, we used the same criteria noted in the preceding paragraph.

Construct validity

Although not a primary focus, we did explore construct validity for the DLOPFQ–SF subscales using the two college samples. Both samples contained the IIP–SC, the SOS–10, and the RQ. These subscales were selected because problems in personality functioning and risk for personality disorders have been linked to patterns of interpersonal problems (Hopwood, Zimmermann, Pincus, & Krueger, 2015; Wilson, Stroud, & Durbin, 2017), adult attachment status (Gunderson & Lyons-Ruth, 2008; Levy, Johnson, Clouthier, Scala, & Temes, 2015), and well-being (Cloninger, 2006; Cramer, Torgersen, & Kringlen, 2006). Construct validity was explored with zero-order correlations and stepwise hierarchical regressions.

Results

Parallel analysis and exploratory factor analyses: Derivation sample

The PA suggested six factors, so we explored models with four, five, six, seven, and eight factors. The four-factor model accounted for 52.9% of the variance. Communalities (after extraction) and factor loadings from the pattern matrix are shown in Table 1. Item 30 was the only item with a problematic cross-loading, loading at .58 on the fourth factor (self-direction problems) and showing a problematic cross-loading of .45 on the third factor (empathy problems). Although Items 4 and 27 had small cross-loadings, these were not problematic (i.e., they were < .40 and not within .20 of their primary loadings).

The five-factor model explained 55.1% of the variance. Communalities (after extraction) ranged from .39 (Item 61) to .79 (Item 9). Similar to the prior model, all of the identity problems items loaded onto a factor (loadings ranged from .43–.90), all items on the empathy problems subscale loaded onto a factor (loadings ranged from .50–.67), and five of the six self-direction items clearly loaded onto the fourth factor

Table 1. Exploratory factor analysis item communalities and loadings in the pattern matrix.

Item	h^2	Identity	Intimacy	Empathy	Self-direction
6	.60	.67			
7	.68	.73			
9	.79	.93			
10	.76	.89			
15	.42	.46			
25	.52	.55			
4	.44	*.21*			.45
17	.62				.64
21	.57				.74
27	.68			*.30*	.68
28	.58				.75
30	.69			**.46**	.59
8	.39			.50	
22	.48			.52	
32	.50			.65	
41	.48			.67	
43	.54			.66	
61	.39			.53	
12	.30		.51		
48	.34		.58		
49	.47		.57		
51	.52		.59		
52	.47		.67		
54	.48		.70		

Note. Identity = identity problems factor; Self-direction = self-direction problems factor; Empathy = empathy problems factor; Intimacy = intimacy problems factor. Communalities reported postextraction. Factor loadings of < .20 are not shown. Unintended cross-loadings are shown in italics; those within .20 of the largest loading are shown in bold.

(loadings ranged from .45–.76). Item 30 loaded at .58 on the fourth factor and a problematic cross-loading of .46 on the third factor. The intimacy problem items split into two factors, with items assessing discomfort with open sharing in relationships (i.e., 49, 51, and 54) loading on Factor 2 and two items focusing on being more task- or achievement-oriented than relationally oriented loading on Factor 5. Item 52 cross-loaded on both factors.

The six-factor model explained 57.4% of the variance. Communalities (after extraction) ranged from .43 (Item 12) to .99 (Item 21). The model was very similar to the five-factor model (i.e., all identity problem items loaded onto a factor; all empathy problems items loaded onto a factor; the intimacy items split into two factors). However, in this model, the self-direction items split. Further, only two items (21 and 27) had primary loadings. All other items had problematic cross-loadings; as such, we did not further consider this model. Similar problems occurred with the seven- and eight-factor models (i.e., multiple factors contained fewer than three items and some items generated problematic cross-loadings).

In our EFAs, only the four-factor model had factors with more than three items per factor (each with an interpretable loading, and each factor having at least one item with a strong loading). Although this model contained an item with a problematic cross-loading, it contained the fewest problematic cross-loadings of the models considered. Thus, we further examined this model with CFA.

Confirmatory factor analyses: Validation sample

Identity problems subscale
We assigned Items 6, 7, 9, 10, 15, and 25 to a single factor. Respective loadings for Items 6, 7, 9, 10, 15, and 25 were

.75, .82, .90, .89, .57, and .72, respectively. The model χ^2 ($df = 9$) was 37.41 ($p < .001$), NNFI was .98, CFI was .99, and RMSEA was .070. The lower 90% confidence interval (CI) for RMSEA was .048 and the upper 90% CI was .094.

Self-direction problems subscale
Items 4, 17, 21, 27, 28, and 30 were assigned to a single factor and had respective loadings of .57, .77, .74, .82, .75, and .79. The model χ^2 ($df = 9$) was 127.00 ($p < .001$), NNFI was .90, and CFI was .94. RMSEA, however, was .141. Modification indexes suggested allowing error terms for Items 27 and 30 to correlate (both items measure tendencies to follow others and avoid leadership). The nested respecified model produced a significant reduction in the model χ^2 ($p < .01$) and CFI increased by .04. With the exception of Item 4, all items had loadings of $> .70$ onto the factor. Item 4 had a loading of .57. Error terms for Items 27 and 30 correlated at .45. The model χ^2 ($df = 8$) was 33.59 ($p < .001$), NNFI was .98, CFI was .99, and RMSEA was .071. The lower 90% CI for RMSEA was .047 and the upper 90% CI was .096.

Empathy problems subscale
Items 8, 22, 32, 41, 43, and 61 were assigned to a single factor and had respective loadings of .59, .63, .73, .69, .70, and .57. The model χ^2 ($df = 9$) was 61.07 ($p < .001$), NNFI was .92, CFI was .95, and RMSEA was .095, with a lower 90% confidence interval of .073 and an upper 90% confidence interval of .119. Quality of fit was acceptable for RMSEA and NNFI, and good for CFI. Modification indexes suggested allowing error terms for Items 22 to 43 to correlate. However, we did not feel the item content was sufficiently redundant to warrant allowing the error terms to correlate.

Intimacy problems subscale
Items 12, 48, 49, 51, 52, and 54 were assigned to a single factor. They produced respective loadings of .35, .47, .69, .76, .51, and .69. The model χ^2 ($df = 9$) was 85.78 ($p < .001$), NNFI was .85, CFI was .91, and RMSEA was .116. As both NNFI and RMSEA failed to meet our criteria for at least adequate fit, we examined modification indexes that suggested allowing the error terms for Items 48 and 52 to correlate. Both items assess comfort with working and depending on others, and we allowed the error terms to correlate. The nested respecified model produced a significant reduction in the model χ^2 ($p < .01$) and increased CFI by .05. The loading for Item 12 dropped to .33, whereas all others remained highly similar to initial loadings. Error terms for Item 48 and Item 52 correlated at .23. The model χ^2 ($df = 8$) was 56.94 ($p < .001$), NNFI was .89, CFI was .94, and RMSEA was .098, with a lower 90% CI of .075 and an upper 90% CI of .123. Thus, we dropped Item 12 due to a low loading and reran the model with only five items. CFI increased by .05. Loadings ranged from .41 to .80. The error terms for Items 48 and 52 correlated at .24. The model χ^2 ($df = 4$) was 7.22 ($p = .13$), NNFI was .99, CFI was .99, and RMSEA was .036 (lower 90% CI of .001, upper 90% CI of .076).

Table 2. Internal consistency estimates for and intercorrelations among *DSM–5* Levels of Personality Functioning Questionnaire–Short Form (DLOPFQ–SF) scales.

	Subscale intercorrelations				Alpha		Full-form correlations	
	1	2	3	4	DS	VS	DS	VS
1. Identity problems	—	.68	.57	.52	.90	.90	.90	.90
2. Self-direction problems	.70	—	.53	.37	.87	.88	.87	.87
3. Empathy problems	.56	.55	—	.44	.81	.76	.69	.71
4. Intimacy problemsa	.51	.37	.46	—	.80	.76	.86	.86
5. Self-problems					.93	.92	.91	.96
6. Interpersonal problems				.85	.83		.96	.89
7. DLOPFQ–SF total					.93	.93	.97	.97

Note. DS = derivation sample; VS = validation sample. Values above the diagonal reflect intercorrelations for the derivation sample and those below the diagonal reflect values calculated from the validity sample. All intercorrelations were significant at *p* < .05. Self-problems = mean value of all identity problems and self-direction problem items (12 items); Interpersonal problems = mean value of all empathy problems and intimacy problems items (11 items); DLOPFQ–SF total = mean of all 23 items.
aThe five-item intimacy problems scale was used for all analyses.

Table 3. Fit statistics for the factor models.

Model	χ^2	df	CFI	NNFI	RMSEA	LCI	HCI
One-factor	2081.71	228	.75	.72	.113	.108	.117
Two-factor	1581.14	227	.82	.80	.097	.092	.101
Three-factor	1268.95	225	.86	.84	.085	.081	.090
Four-factor	843.78	222	.92	.91	.066	.062	.071
Second-order	872.36	224	.91	.90	.067	.063	.072

Note. CFI = comparative fit index; NNFI = nonnormed fit index; RMSEA = root mean square error of approximation; LCI = lower 90% confidence limit; HCI = higher 90% confidence limit.

Internal consistency, agreement with full-form scores, and subscale intercorrelations

Table 2 displays internal consistency estimates for the DLOPFQ–SF derivation and validation samples. We used the six-item version of the identity problems, self-direction problems, and empathy problems subscales, as well as the five-item intimacy problems subscale. All were within the acceptable ranges and all subscales were strongly correlated with one another. Agreement with full-form subscales was also generally strong (although agreement with the full-form empathy subscale was only moderate). Alpha values were also strong for the self-problems (all items from the identity problems and self-direction problems), interpersonal problems (all items from the empathy problems and intimacy problems subscales), and overall LOPF (all 23 items), as were correlations with the DLOPQ full form.

Model comparisons

Using the validation sample and DLOPFQ–SF subscales, we examined various models: a single-factor model, a two-factor model, a three-factor model, a four-factor model, and a second-order factor model. The respecified model (with one correlated error term) for the self-direction problems subscale and the respecified, five-item intimacy problems subscale were used.

Results are presented in Table 3. Fit was poor for the model with all items assigned to a single factor; no fit statistics achieved our criteria for at least adequate fit and several items had poor loadings (i.e., < .40). In the two-factor model, items for the intimacy problems and empathy problems subscales were assigned to a factor, and the items for the identity problems and self-direction problems subscales were assigned to a separate factor. RMSEA for this model was acceptable, but CFI and NNFI were below our criteria (see Table 3) and Items 48 and 52 showed poor loadings (i.e., < .40). As such, we explored a three-factor model in which items from the empathy problems subscale were assigned to a factor, those from the intimacy problems subscales were assigned to a factor, and the items from the identity problems and self-direction problems subscales were assigned to a factor. The three-factor model was the first in which all item loadings were > .40 and RMSEA met our criteria for adequate fit, but CFI and NNFI fell below expectations (see Table 3). In the four-factor model, each item was assigned to its intended factor. All items had loadings of > .40 and CFI, NNFI, and RMSEA all achieved our criteria for adequate fit. Factors were highly correlated. The identity problems factor correlated with the self-direction problems factor (*r* = .75), the intimacy problems factor (*r* = .60), and the empathy problems factor (*r* = .62). The self-direction factor correlated with the intimacy problems factor (*r* = .47) and the empathy problems factor (*r* = .62). The intimacy problems factor correlated with the empathy problems factor at .59.

The second-order factor contained a model in which the four factors were assigned to a higher order factor (LOPF). As shown in Table 2, the second-order factor model generated adequate fit scores for CFI, NNFI, and RMSEA, with values nearly identical to the four-factor model. Again, items all loaded onto respective factors at > .40. Loadings onto the higher order factor were .89 for the identity problems factor, .82 for the self-direction factor, .74 for the empathy problems factor, and .66 for the intimacy problems factor.

Construct validity

To explore construct validity, we combined the two college samples and calculated scores for the six-item versions of the identity problems, self-direction problems, and empathy problems subscales, and the five-item version of the intimacy problems subscale. Zero-order correlations are reported in Table 4. Results from the hierarchical regressions predicting interpersonal dominance, interpersonal warmth, and well-being can be found in Table 5, and those for the

Table 4. Zero-order correlations among *DSM–5* Levels of Personality Functioning Questionnaire–Short Form (DLOPFQ–SF) scales and dependent variables.

	IIP total	WRM	DOM	Fr Dom	Fr Sub	Ho Dom	Ho Sub	SOS–10	Sec	Fear	Preoc	Dism
Identity problems	.48	.13	−.14	.49	.48	.28	.40	−.56	−.30	.38	.34	*ns*
Self-direction problems	.45	.11	−.15	.45	.48	.26	.38	−.42	−.19	.25	.28	*ns*
Empathy problems	.42	*ns*	*ns*	.41	.36	.35	.29	−.24	−.15	.14	.23	*ns*
Intimacy problems	.35	*ns*	−.21	.22	.35	.24	.34	−.30	−.38	.39	.12	.21
Self-problems	.51	.13	−.15	.51	.51	.29	.42	−.54	−.27	.35	.34	*ns*
Interpersonal problems	.45	*ns*	−.11	.37	.43	.35	.38	−.33	−.32	.32	.21	.16
DLOPFQ–SF total	.53	.08	−.15	.50	.52	.35	.45	−.50	−.33	.37	.31	*ns*

Note. IIP total = Inventory of Interpersonal Problems–Short Circumplex (IIP–SC) total score; WRM = IIP–SC interpersonal warmth dimension; DOM = IIP–SC interpersonal dominance dimension; Fr Dom = IIP–SC friendly-dominant quadrant score; Fr Sub = IIP–SC friendly-submissive quadrant score; Ho Dom = hostile-dominant quadrant score; Ho Sub = hostile-submissive quadrant score; SOS–10 = Swartz Outcome Scale–10 total score; Self-problems = DLOPFQ–SF self-problems scale (identity problems + self-direction problems); Interpersonal problems = DLOPFQ–SF interpersonal problems scale (empathy problems + intimacy problems); DLOPFQ–SF total = mean of all 23 items. *ns* = nonsignificant, all other correlations are significant at $p < .05$.

hierarchical regressions predicting the four attachment prototype dimensions are reported in Table 6.

Discussion

Using both EFA and CFA, a short form for the DLOPFQ was developed. Score agreement was adequate to strong at all levels (i.e., overall LOPF, scores for self- and interpersonal functioning, and subscale scores). Psychometrically, short-form subscales were internally consistent and unidimensional. In CFA models, items showed good-to-strong loadings on theorized subscales and model fit was consistently adequate or better for each subscale. Overall, findings suggest that the DLOPFQ–SF is assessing constructs in a manner consistent with the full form, and is structurally sound.

The authors of the DLOPFQ full form proposed a four-scale structure, but assumed each would be indicators of overall LOPF (i.e., subscales of LOPF; Huprich et al., 2018). Although this early study found subscales to be positively correlated (at similar magnitudes to other scales, such as the AMPD–CAS [Dowgwillo et al., 2018]), a second-order factor model (i.e., overall LOPF) was never tested, because participant-to-item ratios were not sufficient. Because the short form contains fewer items, we examined this issue directly. Consistent with the AMPD's assertion that functional elements are interconnected and reciprocally influential, meaning that functioning in one element domain likely affects functioning in others (Bender et al., 2011), fit for the second-order factor model was acceptable. Items grouped into conceptually distinct element subscales, all of which loaded onto a second-order LOPF factor. This pattern of findings supports the use of the DLOPFQ–SF's overall LOPF score as an index of overall functioning.

The construct validity findings provide initial support for the DLOPFQ–SF as a measure of LOPF. Similar to prior research (e.g., Dowgwillo et al., 2018; Hutsebaut et al., 2016), scores for self-problems, interpersonal problems, and overall LOPF were associated with indicators of personality dysfunction. The overall LOPF score was associated with lower life satisfaction, total interpersonal problems, and less romantic attachment security. Although there was some variation in magnitudes, the four element subscales were

Table 5. Hierarchical regressions predicting interpersonal variables and subjective well-being.

Dependent	Predictor	β	t	R	R²	F	df
IIP total	Identity problems	.24	4.75**				
	Empathy problems	.17	4.05**				
	Self-direction problems	.15	3.11**				
	Intimacy problems	.10	2.54*				
	Final model			.54	.29	64.18	4, 638
Warmth	Identity problems	.31	6.48**				
	Intimacy problems	.20	4.09**				
	Final model			.18	.03	10.33**	2, 640
Dominance	Intimacy problems	−.24	5.62**				
	Empathy problems	.24	5.17**				
	Self-direction problems	−.19	5.06**				
	Final model			.30	.08	20.43**	3, 639
SOS–10	Identity problems	−.56	16.89**				
	Final model			.56	.31	285.18**	1, 641

Note. IIP = Inventory of Interpersonal Problems; Warmth = IIP–32 warmth dimension; Dominance = IIP–32 dominance dimension; SOS–10 = the Swartz Outcome Scale–10 total score.

Table 6. Hierarchical regressions predicting fit with the four attachment prototypes.

Dependent	Predictor	β	t	R	R²	F	df
Secure	Intimacy problems	−.31	7.44**				
	Identity problems	−.15	3.67**				
	Final model			.41	.17	62.71**	2, 633
Fearful	Intimacy problems	.30	7.16**				
	Identity problems	.31	7.07**				
	Empathy problems	−.14	3.41**				
	Final model			.46	.21	57.23**	3, 631
Preoccupied	Identity problems	.53	7.98**				
	Final model			.34	.11	81.12**	1, 633
Dismissive	Intimacy problems	.29	6.93**				
	Self-direction problems	−.21	4.96**		.56	.31	285.18** 1, 641
	Final model			.28	.08	27.05**	2, 631

Note. Self-Direction Probs. = Self-Direction Problems Scale..

also associated with indicators in similar directions. With a few exceptions, however, associations between the overall LOPF score were generally larger than corresponding associations with specific subscales. These findings, when combined with the findings for the second-order factor model, suggest that the DLOPFQ–SF's overall LOPF score might have utility as a screening measure for Criterion A personality difficulties.

Although not a primary aim of this article, regressions did suggest that some subscales might be more uniquely predictive of some specific indicators than others. For example, although highly correlated, the identity problems

and self-direction problems subscales were linked to different patterns of attachment. The former was predictive of less security and greater fearful-avoidance, whereas an absence of the latter was predictive of dismissive avoidance. Similarly, although empathy problems and intimacy problems were positively correlated with one another, intimacy problems increased risk for interpersonal submissiveness and empathy problems increased risk for dominance. Thus, although an overall measure of LOPF is likely to be most useful to clinicians, present findings and prior work suggest that some elements' subscales might show more specific associations with certain types of dysfunction (Dowgwillo et al., 2018) and day-to-day patterns of difficulty (Roche et al., 2016).

Although the short form achieved many of our aims, some challenges emerged during development. First, although consistent with the AMPD model in term of structure, the DLOPFQ–SF subscales might not tap all aspects of dysfunction. Like other self-report LOPF measures, DLOPFQ–SF subscales construe problems within element domains in a fairly narrow and unidirectional manner (e.g., the absence of an identity, a lack of empathy). It could be that being too self-directed, having too fixed an identity, seeking too much intimacy with others, or having so much empathy that one lacks boundaries, could all indicate problems in functioning. Currently, DLOPFQ and DLOPFQ–SF items are unidirectional (with one pole suggesting problems and one suggesting an absence of problems). Further, our item pool was limited to the original DLOPFQ. As in the original version, some aspects of the *DSM–5* AMPD model are lacking. For example, the identity subscale contains very few items tapping self-esteem or emotional regulation. Reducing the number of items to create a short form might have exacerbated this existing issue. Overall, this might limit the range of personality dysfunction assessed by the DLOPFQ–SF. Thus, although organized around the four element domains of the AMPD, the DLOPFQ–SF might not capture the full range of these element domains as they are described in the *DSM–5* (American Psychiatric Association, 2013, pp. 775–778) and elsewhere (e.g., Bender et al., 2011). Although sensitive to some problems within each domain, the limited range of item content might result in a lack of sensitivity to other problems. Thus, whereas scores for the DLOPFQ–SF are strongly correlated with those of the full form, the range of construct coverage might be more limited.

A second challenge involved a handful of items shifting from their original DLOPFQ full-form subscales. Specifically, Items 8, 12, 15, 22, 25, and 61 all ended up on a short-form subscale that differed from their subscale of origin on the full-form version. To date, we are unaware of any EFA or CFA study of the DLOPFQ. Thus, such findings might suggest opportunities to improve full-form subscales by examining item loadings. This assertion is further bolstered by our experience of working with the empathy problems subscales.

While developing the DLOPFQ–SF, an interesting pattern emerged within the full-form empathy problems subscale.

The initial series of EFAs suggested two possible item sets within the empathy problems subscale. One set assessed low investment in others and poor listening (these items comprise our empathy problems subscale) and a separate item set involved oversensitivity to others' emotions and excessive investment in others' perspective (none of these items are included in this version of the DLOPFQ–SF). Although both item sets assess empathy challenges, these two item sets did not cluster together and were possibly orthogonal in nature. Perhaps the full form assesses two types of empathy disturbance that are somewhat unique. Ultimately this is an issue to be resolved by studying the full form. For our purposes, we selected the items clearly worded as problems, that generated a subscale with positive associations to the other three scales, and that had high item loadings. This observation might explain why the short-form empathy problems subscale has a lower level of agreement with its full-form counterpart relative to the other three subscales.

This study, like all studies, has limitations. We used a mixed sample, containing respondents from an outpatient psychiatric clinic, a primary care clinic, community respondents, and college students. Although mixed samples are optimal for factor analytic studies (Osborne, 2014), we would have ideally had more participants from clinical settings. A relative dearth of clinical participants, for example, might have elevated correlations among our subscales and could limit generalizability of findings. With regard to analytic approach, we evaluated our subscales based primarily on their dimensional structure. Given the heterogeneous nature of the constructs within each element subscale, this approach might not be ideal for assessing the quality of the DLOPFQ subscales. It is possible that alternative approaches, such as item–response theory, would have yielded different item sets for the short-form subscales. This is an issue worthy of further investigation. Our investigation of validity was limited to our nonclinical college sample, as data beyond the DLOPFQ full form were not available to us in the other samples. The results presented here provide only a tentative exploration of the validity of these subscales and our analyses were limited to cross-sectional correlations among self-report subscales. We are in the process of collecting additional data that include self-report, clinician ratings, and life data (e.g., history of inpatient hospitalization, number of depressive episodes) to further examine the validity of DLOPFQ–SF subscales. Additionally, future DLOPFQ–SF studies would likely benefit from including additional self-report LOPF measures (e.g., the LPFS–SR) to better determine how such measures overlap and diverge. Such research is necessary to further determine the DLOPFQ–SF's merit. Finally, this study did not include a measure of the maladaptive traits (i.e., Criterion B). In the future, studies using the DLOPFQ–SF would optimally include a brief measure of the maladaptive traits to determine how these relate to one another and if there is support for assessing both sets of constructs.

In conclusion, the DLOPFQ–SF assesses LOPF in a manner that corresponds with the DLOPFQ full form (albeit with fewer items). The DLOPFQ–SF offers clinicians and

researchers a low-burden, self-report measure that can be used to assess LOPF at multiple levels. Given that some aspects of personality functioning might be less amenable to self-report (at least for some individuals), the DLOPFQ–SF should only be used as a complement to clinician-rating systems, which remain the gold standard for LOPF assessment. In research settings, additional work with the DLOPFQ–SF is needed. Future efforts should seek to confirm present findings, determine how the DLOPFQ–SF relates to specific aspects of interpersonal functioning, consider the utility of the measure for differentiating those with and without personality disorders, examine relationships between the DLOPFQ–SF and maladaptive personality traits, and determine if the measure's structure is invariant across various groups (e.g., genders; clinical and nonclinical groups). In addition, research that examines convergence and divergence between the DLOPFQ–SF and existing self-report LOPF measures (e.g., LPFS–SR, AMPD–CAS) would be of benefit.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Bach, B., Markon, K., Simonsen, E., & Krueger, R. F. (2015). Clinical utility of the *DSM-5* alternative model of personality disorders: Six cases from practice. *Journal of Psychiatric Practice*, *21*(1), 3–25.

Bagby, R. M., & Widiger, T. A. (2018). Five factor model personality disorder scales: An introduction to a special section on assessment of maladaptive variants of the five factor model. *Psychological Assessment*, *30*(1), 1–9.

Bartholomew, K., & Horowitz, L. M. (1991). Attachment styles among young adults: A test of a four-category model. *Journal of Personality and Social Psychology*, *61*(2), 226–244.

Bender, D. S., Morey, L. C., & Skodol, A. E. (2011). Toward a model for assessing level of personality functioning in *DSM–5*, part I: A review of theory and methods. *Journal of Personality Assessment*, *93*(4), 332–346.

Blais, M., Lenderking, W., Baer, L., DeLorell, A., Peets, K., Leahy, L., & Burns, C. (1999). Development and initial validation of a brief mental health outcome measure. *Journal of Personality Assessment*, *73*(3), 359–373.

Bornstein, R. F., & Hopwood, C. J. (2014). Introduction to multi-method clinical assessment. In C. J. Hopwood & R. F. Bornstein (Eds.), *Multi-method clinical assessment* (pp. 1–20). New York, NY: Guilford Press.

Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. New York, NY: Routledge.

Christensen, T. B., Papp, M. C. S., Arnesen, M., Koritzinsky, K., Nysaeter, T.-E., Eikenaeas, I., … Skodol, A. (2018). Interrater reliability of the structured clinical interview for the DSM-5 alternative model of personality disorders module I: Level of personality functioning scale. *Journal of Personality Assessment*, *100*(6), 630–641. doi:10.1080/00223891.2018.1483377

Cloninger, C. R. (2006). The science of well-being: An integrated approach to mental health and its disorders. *World Psychiatry*, *5*, 71–76.

Cramer, V., Torgersen, S., & Kringlen, E. (2006). Personality disorders and quality of life. A population study. *Comprehensive Psychiatric*, *47*(3), 178–184.

Dowgwillo, E. A., Roche, M. J., & Pincus, A. L. (2018). Examining the interpersonal nature of criterion A of the *DSM-5* section III alternative model for personality disorders using bootstrapped confidence intervals for the interpersonal circumplex. *Journal of Personality Assessment*, *100*(6), 581–592. doi:10.1080/00223891.2018.1464016

Few, L. R., Miller, J. D., Rothbaum, A., Meller, S., Maples, J., Terry, D. P., … 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. doi: 10.1037/a0034878

Gunderson, J. G., & Lyons-Ruth, K. (2008). BPD's interpersonal hypersensitivity phenotype: A gene-environment-developmental model. *Journal of Personality Disorders*, *22*(1), 22–41.

Haggerty, G., Blake, M., Naraine, M., Siefert, C., & Blais, M. A. (2010). Construct validity of the Schwartz Outcome Scale-10: Comparisons to interpersonal distress, adult attachment, alexithymia, the five-factor model, romantic relationship length, and ratings of childhood memories. *Clinical Psychology & Psychotherapy*, *17*, 44–50.

Hopwood, C. J., Good, E. W., & Morey, L. C. (2018). Validity of the *DSM-5* levels of personality functioning scale-self-report. *Journal of Personality Assessment*, *100*(6), 650–659. doi:10.1080/00223891.2017.1420660.

Hopwood, C. J., Pincus, A. L., DeMoor, R. M., & Koonce, E. A. (2008). Psychometric characteristics of the inventory of interpersonal problems—Short circumplex (IIP-SC) with college students. *Journal of Personality Assessment*, *90*(6), 615–618.

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

Horowitz, L. M., Alden, L. E., Wiggins, J. S., & Pincus, A. L. (2000). *Inventory of interpersonal problems*. London, UK: Psychological Corporation.

Huprich, S. K., Nelson, S. M., Meehan, K. B., Siefert, C. J., Haggerty, G., Sexton, J., … Baade, L. (2018). Introduction of the *DSM-5* levels of personality functioning questionnaire. *Personality Disorders: Theory, Research, and Treatment*, *9*(6), 553–563. Advanced online publication. doi:10.1037/per0000264

Hutsebaut, J., Feenstra, D. J., & Kamphuis, J. H. (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.

Hutsebaut, J., Kamphuis, J. H., Feenstra, D. J., Weekers, L. C., & De Saeger, H. (2017). Assessing *DSM-5* oriented LPF: Development and psychometric evaluation of the semi-structured interview for personality functioning DSM-5 (STIP-5). *Personality Disorders*, *8*(1), 94–101.

Jackson, D. L., Gillaspy, J. A., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: An overview and some recommendations. *Psychological Methods*, *14*(1), 6–21.

Keeley, J. W., Flanagan, E. H., & McCluskey, D. L. (2014). Functional impairment and the *DSM-5* dimensional system for personality disorder. *Journal of Personality Disorders*, *28*(5), 657–674.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford.

Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., & Skodol, A. E. (2012). Initial construction of maladaptive personality trait model and inventory for *DSM-5*. *Psychological Medicine*, *42*(09), 1879–1890.

Krueger, R. F., Hopwood, C. J., Wright, A. G. C., & Markon, K. E. (2014). *DSM–5* and the path toward empirically based and clinically useful conceptualization of personality and psychopathology. *Clinical Psychology: Science and Practice*, *21*, 245–261.

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

Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., … Reed, G. M. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist*, *56*(2), 128–165.

Morey, L. C. (2017). Development and initial evaluation of a self-report form of the *DSM-5* levels of personality functioning scale. *Psychological Assessment*, *29*(10), 1302–1308. Advance online publication. doi:10.1037/pas0000450.

Morey, L. C., Berghuis, H., Bender, D. S., Verheul, R., Krueger, R. F., & Skodol, A. E. (2011). Toward a model of 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.

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers*, 32, 369–402.

Osborne, J. W. (2014). *Best practices in exploratory factor analysis*. Scotts Valley, CA: CreateSpace Independent Publishing.

Roche, M. J., Jacobson, N. C., & Pincus, A. L. (2016). Using repeated daily assessments to uncover oscillating patterns and temporally-dynamic triggers in structures of psychopathology: Applications to the *DSM–5* alternative model of personality disorders. *Journal of Abnormal Psychology*, 125(8), 1090–1102.

Ruiz, M. A., Pincus, A. L., Borkovec, T. D., Echemendia, R. J., Castonguay, L. G., & Ragusea, S. A. (2004). Validity of the inventory of interpersonal problems for predicting treatment outcome: An investigation with the Pennsylvania practice research network. *Journal of Personality Assessment*, 83(3), 213–222.

Scharfe, E., & Bartholomew, K. (1994). Reliability and stability of adult attachment patterns. *Personal Relationships*, 1(1), 23–43.

Schmitt, T. A., Sass, D. A., Chappelle, W., & Thompson, W. (2018). Selecting the "best" factor structure and moving measurement validation forward: An illustration. *Journal of Personality Assessment*, 100(4), 345–362. doi:10.1080/00223891.2018.1449116.

Sexton, J., Siefert, C. J., Nelson, S., Meehan, K., Dauphin, B., & Huprich, S. (2019). The test-retest reliability and incremental validity of the DSM-5 levels of personality functioning questionnaire. Unpublished Manuscript.

Skodol, A. E., Clark, L. A., Bender, D. S., Krueger, R. F., Morey, L. C., Verheul, R., … Oldham, J. M. (2011). Proposed changes in personality and personality disorder assessment and diagnosis for *DSM-5* part I: Description and rationale. *Personality Disorders*, 2(1), 4–22. doi:10.1037/a0021891

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson.

Thylstrup, B., Simonsen, S., Nemery, C., Simonsen, E., Noll, J. F., Myatt, M. W., & Hesse, M. (2016). Assessment of personality-related levels of functioning: A pilot study of clinical assessment of the *DSM-5* level of personality functioning based on a semi-structured interview. *BMC Psychiatry*, 16, 298. doi:10.1186/s12888-016-1011-6.

Trull, T. J., & Widiger, T. A. (2013). Dimensional models of personality: The five factor model and the *DSM-5*. *Dialogues in Clinical Neuroscience*, 15, 135–146.

Wilson, S., Stroud, C. B., & Durbin, C. E. (2017). Interpersonal dysfunction in personality disorders: A meta-analytic review. *Psychological Bulletin*, 143(7), 677–734.

Wright, A. G. C. (2017). The current state and future of factor analysis in personality disorder research. *Personality Disorders, Theory Research and Treatment*, 8(1), 14–25.

Zimmerman, J., Böhnke, J. R., Eschstruth, R., Mathews, A., Wenzel, K., & Leising, D. (2015). The latent structure of personality functioning: Investigating criterion A of the alternative model for personality disorders in *DSM-5*. *Journal of Abnormal Psychology*, 124, 532–548.

Appendix

Identity problems scale items

6. I wish I could feel confident and comfortable in my own skin more often.
7. I struggle to describe how I usually am.
9. There are days when I just don't think I will ever figure out who I really am.
10. Sometimes I feel so messed up that I don't think anyone could understand me.
15. Unlike other people, I don't have a sense of what my future plans are.
25. I find myself in the same dilemmas over and over again even though I keep trying to make things better.

Self-direction scale items

4. When someone provides me with "constructive criticism," I act like it is OK but deep down I am very worried.
17. If I'm in a new situation, I really need someone to help me find my way to get comfortable.
21. When starting a new activity, it is very important I have someone there to help me.
27. I find it safer to follow other people's ideas than risk following my own and failing.
28. Even when I know what to do, I like to get reassurance from others about my choice before acting on it.
30. I feel it is better to follow what other people are doing than do what I want to do because I might stand out.

Empathy problems scale items

8. If one of my friends or coworkers becomes angry at me, I start to think about getting a new friend since I don't need someone to get me that upset.
22. Because others are not able to give me what I want, I have to take what I can from them anytime it comes my way.
32. Even with people I know, if they show signs of sadness or anger, I try to get away as soon as I can.
41. It is a waste of time to think about my or others' personal reactions.
43. Others say I don't listen to them very well.
61. People would say that I often overestimate the amount of very close friends I have.

Intimacy problems scale

12. Others might say that I am more motivated to accomplish things than spend time with others.[a]
48. I feel I can do a better job on my own without the demands of others.
49. Other people say I am hard to get to know.
51. When people ask me a personal question, I often feel uncomfortable.
52. Unlike most, I don't want to depend on another or have someone depending on me.
54. It takes a very long time for me to tell others about my personal history.

---

[a]This item had a poor loading in confirmatory factor analyses. Our construct validity data were virtually identical regardless of if we included this item or not. Future studies might be required to determine if this item should be included in the short form. At this time, we suggest administering the item but not using it in calculating a scale score.