

# Initial construction of a maladaptive personality trait model and inventory for DSM-5

R. F. Krueger<sup>1\*</sup>, J. Derringer<sup>1</sup>, K. E. Markon<sup>2</sup>, D. Watson<sup>3</sup> and A. E. Skodol<sup>4</sup>

<sup>1</sup> Department of Psychology, University of Minnesota, Minneapolis, MN, USA

<sup>2</sup> Department of Psychology, University of Iowa, Iowa City, IA, USA

<sup>3</sup> Department of Psychology, University of Notre Dame, Notre Dame, IN, USA

<sup>4</sup> Department of Psychiatry, University of Arizona School of Medicine, Phoenix, AZ, USA

**Background.** DSM-IV-TR suggests that clinicians should assess clinically relevant personality traits that do not necessarily constitute a formal personality disorder (PD), and should note these traits on Axis II, but DSM-IV-TR does not provide a trait model to guide the clinician. Our goal was to provide a provisional trait model and a preliminary corresponding assessment instrument, in our roles as members of the DSM-5 Personality and Personality Disorders Workgroup and workgroup advisors.

**Method.** An initial list of specific traits and domains (broader groups of traits) was derived from DSM-5 literature reviews and workgroup deliberations, with a focus on capturing maladaptive personality characteristics deemed clinically salient, including those related to the criteria for DSM-IV-TR PDs. The model and instrument were then developed iteratively using data from community samples of treatment-seeking participants. The analytic approach relied on tools of modern psychometrics (e.g. item response theory models).

**Results.** A total of 25 reliably measured core elements of personality description emerged that, together, delineate five broad domains of maladaptive personality variation: negative affect, detachment, antagonism, disinhibition, and psychoticism.

**Conclusions.** We developed a maladaptive personality trait model and corresponding instrument as a step on the path toward helping users of DSM-5 assess traits that may or may not constitute a formal PD. The inventory we developed is reprinted in its entirety in the Supplementary online material, with the goal of encouraging additional refinement and development by other investigators prior to the finalization of DSM-5. Continuing discussion should focus on various options for integrating personality traits into DSM-5.

Received 13 April 2011; Revised 27 July 2011; Accepted 25 October 2011; First published online 8 December 2011

**Key words:** Assessment, DSM, personality disorders, personality traits.

## Introduction

In DSM-IV-TR (Diagnostic and Statistical Manual of Mental Disorders, fourth edition, text revision), if maladaptive personality traits are part of the clinical picture but do not constitute a formal personality disorder (PD), the clinician is encouraged to record these traits on Axis II. Nevertheless, DSM-IV-TR does not provide a specific model for conceptualizing these traits, beyond their appearance as features of the 10 PDs. This is in spite of the fact that there has been much interest recently in the etiology (van den Oord *et al.* 2008; de Moor *et al.* 2010), treatment (Tang *et al.* 2009), clinical relevance (Hopwood *et al.* 2008) and

social costs (Cuijpers *et al.* 2010) of personality traits in the psychiatric literature.

Our aim in the current research was (a) to construct a preliminary maladaptive personality trait model for DSM-5 and then (b) to test and refine it through the creation of a provisional corresponding assessment instrument. This research was conducted under the auspices of the DSM-5 Personality and Personality Disorders Workgroup, by members of the workgroup and workgroup consultants. Nevertheless, it is critical to emphasize that no decisions have been formalized regarding the conceptualization of PDs in DSM-5, or regarding the ways in which the constructs of personality traits and PDs might best be represented in DSM-5. By disseminating our findings to date in this report, our hope is to stimulate additional research from other investigators that can inform DSM-5 and subsequent revisions.

\* Address for correspondence: R. F. Krueger, Ph.D., Department of Psychology, University of Minnesota, N414 Elliott Hall, Minneapolis, MN 55455, USA.

(Email: krueg038@umn.edu)

With regard to a personality trait model suitable for DSM-5, a variety of compelling models exist, instantiated in a corresponding variety of assessment instruments (for reviews, see Trull & Durrett, 2005; Clark, 2007). The DSM-5 workgroup and consultants (a number of whom are authors of published measures of clinically relevant personality constructs; Morey, 2003; Livesley & Jackson, 2009; Clark *et al.*, in press) began by reviewing these existing models and measures of maladaptive personality traits, with a particular focus on reviews completed as part of a 2004 pre-DSM-5 research planning meeting. In particular, Widiger & Simonsen (2005) provided evidence that four broad bipolar domains (i.e. domains with two opposite ends) of extraversion *v.* introversion, antagonism *v.* compliance, constraint *v.* impulsivity, and negative affect *v.* emotional stability could serve as an organizing framework for traits seen across 18 models that had been described in the literature. They also described a fifth potential domain, 'unconventionality *v.* closedness to experience', but noted that this domain was not well represented in the models they reviewed, although it is a major domain assessed by the Revised NEO Personality Inventory (Costa & McCrae, 1992). Moreover, a recent meta-analysis showed essentially zero correlation between this domain and DSM-IV PDs (Samuel & Widiger, 2008). Nevertheless, other research has identified a domain of peculiar or odd traits that provides coverage of features corresponding with some key components of schizotypal PD, i.e. 'cognitive or perceptual distortions and eccentricities of behavior' (DSM-IV-TR, APA, 1994; Harkness *et al.* 1995; Chmielewski & Watson, 2008). Hence, in addition to the four major domains identified by Widiger & Simonsen (2005), we also sought to identify and measure traits in a fifth domain of psychoticism, resulting in a model that, at the domain level, bears a strong resemblance to Harkness's Personality Psychopathology 5 model of clinically relevant personality variants (Harkness *et al.* 1995).

Our focus was initially on identifying and operationalizing specific maladaptive personality dimensions falling within five broad domains, with a focus on the poles of these domains that are associated with PD (i.e. introversion, antagonism, impulsivity, negative affect, and psychoticism). That is, the features of PD tend to be concentrated at specific poles of these domains. In a meta-analytic review of literature linking the Five Factor Model of personality (FFM; see Goldberg, 1993; Costa & Widiger, 2002) – which bears a strong resemblance to the model described by Widiger & Simonsen (2005) – with the DSM-IV PDs (Samuel & Widiger, 2008), DSM-IV PDs were associated with introversion (the absence of FFM extraversion), antagonism (the absence of FFM

agreeableness), impulsivity (the absence of FFM conscientiousness) and negative affect (FFM neuroticism). There were only two exceptions: an association between histrionic PD and FFM extraversion, and an association between obsessive-compulsive PD and FFM conscientiousness. Hence, we endeavored to ensure that our trait list also covered core features of histrionic PD and obsessive-compulsive PD. DSM-IV describes the core features of histrionic PD as 'excessive emotionality and attention seeking' (APA, 1994), so we ensured coverage of those two primary traits. DSM-IV describes the core features of obsessive-compulsive PD as 'preoccupation with orderliness, perfectionism, and control' (APA, 1994). These traits broadly define the constraint pole of impulsivity *v.* constraint in the Widiger & Simonsen (2005) model. Hence, we focused on the delineation and measurement of specific maladaptive traits in the domains of (I, high pole) introversion, (II, high pole) antagonism, (III, high pole) impulsivity *v.* (III, low pole) constraint, (IV, high pole) negative affect, and (V, high pole) psychoticism. As described below, we subsequently changed the name of the introversion domain to detachment and the name of the impulsivity domain to disinhibition, to better reflect the content of these domains, at least as that content emerged in our project.

In sum, our approach was to synthesize existing models to arrive at a model and assessment instrument that (a) encompass the four major domains of maladaptive personality variation identified by Widiger & Simonsen (2005), with explicit measurement of the poles of those domains associated with DSM-IV-TR PDs; (b) also contains an additional fifth domain of psychoticism; and (c) contains multiple specific maladaptive trait facets within all five domains, with a focus on covering the maladaptive trait features of DSM-IV-TR PDs. To our knowledge, no existing model and assessment instrument encompasses this complete set of features.

An additional consideration regarding existing models is that these models are not suited to being imported verbatim into the DSM because they are typically operationalized in specific, commercially available assessment instruments. In sum, our approach was to draw broadly on research on existing models, to frame the generation of an empirically based model and measure that is freely available and can be employed in research that can inform DSM-5 and beyond.

### Measure construction

We began with a hypothesized set of domains identified throughout the DSM-5 process to date as covering maladaptive personality variation in existing instruments and models. Work group members

**Table 1.** Original 37 facets, mapped to final 25 facets and five domains<sup>a</sup>

Original facet (prior to Round 1 data collection)	Restructured facet (after Round 2 analyses)	Final domain
<b>Negative Affect</b>		
Anxiousness	Anxiousness	
Emotional lability	Emotional Lability	
Hostility	Hostility	
Oppositionality		
Perseveration	Perseveration	
(lack of) Restricted Affectivity	(lack of) Restricted Affectivity	
Separation Insecurity	Separation Insecurity	
Submissiveness	Submissiveness	
<b>Detachment</b>		
Anhedonia	Anhedonia	
Depressivity		
Guilt and Shame		
Low Self-Esteem	Depressivity	
Pessimism		
Self-Harm		
Intimacy Avoidance	Intimacy Avoidance	
Suspiciousness	Suspiciousness	
Social Detachment		
Social Withdrawal	Withdrawal	
<b>Antagonism</b>		
Attention Seeking	Attention Seeking	
Aggression		
Callousness	Callousness	
Deceitfulness	Deceitfulness	
Grandiosity	Grandiosity	
Manipulativeness	Manipulativeness	
<b>Disinhibition</b>		
Distractibility	Distractibility	
Impulsivity	Impulsivity	
Irresponsibility	Irresponsibility	
(lack of) Orderliness		
(lack of) Perfectionism	(lack of) Rigid Perfectionism	
(lack of) Rigidity		
Recklessness		
(lack of) Risk Aversion	Risk Taking	
<b>Psychoticism</b>		
Cognitive Dysregulation	Eccentricity	
Eccentricity		
Dissociation Proneness	Cognitive and Perceptual Dysregulation	
Unusual Beliefs		
Unusual Perceptions	Unusual Beliefs and Experiences	

<sup>a</sup> Rigid perfectionism and Restricted affectivity are indicators of the opposite pole of their respective domains (see Table 3).

and consultants generated a list of 37 facets (specific personality traits, as opposed to broad domains containing multiple traits) as potential exemplars with the aim of covering all the domains. Each of the 11 members of the DSM-5 Personality and Personality Disorders Workgroup had the opportunity to contribute to this process and were given multiple

opportunities to review the resulting list of facets. Once the preliminary list of 37 facets was finalized, the authors of the current paper wrote brief definitions of each of the facets, and relied on these definitions in writing items designed to tap their content.

Table 1 presents a list of these domains and facets, using the terminology we arrived at ultimately.

Our initial measure construction took place over the course of two rounds of data collection, aimed at (a) measuring each proposed facet reliably, and (b) examining whether facets could be collapsed, or items reassigned among facets, within each of the domains. Both of these measure construction goals were completed by collecting data from an on-line panel of respondents specifically cultivated to ensure the ability to generalize to the US population (the Knowledge Networks Panel; Dennis, 2010). Samples were selected from among persons in the Knowledge Networks Panel who responded positively to the question 'have you ever seen a therapist for psychological or psychiatric counseling or therapy' in a previous survey. This sampling strategy was used in our initial rounds of data collection to help ensure that our respondents were more nationally representative than those who could be recruited from a specific clinic, but also more likely to show variation in maladaptive personality characteristics, relative to a sample not selected for seeking mental health services.

### Round 1

The first round of data collection was designed to examine our ability to measure the 37 initial facets. We began by writing eight specific personality items to measure each proposed facet. The reading level of all items was eighth grade or less, to ensure that the resulting inventory could be completed by persons with varying levels of education.

### Participants

A total of 762 persons participated in our initial round of data collection. Knowledge Networks collected all data via a web-based survey. Demographic characteristics for all three waves of data collection are presented in Supplementary Table S1. Sampling weights were applied in all analyses (Asparouhov, 2005) to adjust the current sample demographics to be representative of the US population (and, at Rounds 1 and 2, taking into account our study-specific inclusion criterion of having seen a psychiatrist or psychologist). The weighting took into account both sampling (e.g. under-sampling of telephone numbers not matched to a valid mailing address) and participant characteristics (i.e. sex, age, race/ethnicity, education, geographic location, living in a metropolitan area, and Internet access).

### Measure

The Round 1 item pool consisted of 296 items (eight per each of 37 facets). Item order was randomized, and the items were split into four sections of 74 items each.

These sections were then combined into six booklets. That is, if the 74-item sections are designated ABCD, the six booklets were: AB, AC, AD, BC, BD, and CD. Participants were then randomly assigned to receive one of the six booklets, such that each participant received 50% of the total measure and participant missingness across items was completely at random (see, e.g. Smits and Vorst, 2007). This reduced the response burden on each participant, while still ensuring that covariances were estimable among all items. Each item was displayed on a separate screen. All items presented the same four response options of 'Very false or often false', 'Sometimes or somewhat false', 'Sometimes or somewhat true' and 'Very true or often true'. The entire final measure, including respondent instructions, item content, and response options, is given in Appendix A of the Supplementary material.

### Analyses

Analyses were completed in the software package Mplus (Muthén & Muthén, 1998–2010) on raw data, weighted according to Knowledge Networks' determined sampling weights to adjust the data to be population representative, using a robust maximum-likelihood estimator to estimate model parameters and to model missingness directly (MLR), and treating categorical manifest variables (e.g. item-level responses) as categorical. Initially, Geomin-rotated exploratory factor analyses (EFAs) were run on all eight items within each facet, requesting factor results for one to four factors and treating the items as ordinal. Factor solutions were compared on their Bayesian information criterion (BIC; Schwarz, 1978), with a lower BIC indicating a better relative fit to the data. Because we were fitting directly to the raw response patterns, fit indices that can be obtained for maximum likelihood (ML) estimation applied to moment matrices (e.g. covariances) were not available in this scenario. We relied on BIC because of its emphasis on parsimony, which corresponded with our aim of identifying parcels of items that measured a single factor (i.e. the most parsimonious model among models containing one to k factors). Specifically, if BIC indicated that a one-factor solution fit the data best, all items were retained for subsequent analyses. If BIC indicated that the best-fitting solution included more than one factor, we selected items loading on the largest factor (in terms of the greatest number of items) and re-ran the EFA to confirm that a one-factor solution fit best in the retained items.

We then fit one-factor models to items within each of the facets, with the aim of further refining these initial facet measures. Specifically, if any items had

**Table 2.** Means and standard deviations for facet and domain scores (calculated as the average response across items) and classical test theory reliability estimates (Cronbach's  $\alpha$ )<sup>a</sup>

	Items	Selected sample, Round 2			Representative sample, Round 3		
		$\alpha$	Mean	(s.d.)	$\alpha$	Mean	(s.d.)
<b>Facet</b>							
Anhedonia	8	0.88	1.05	(0.65)	0.88	0.89	(0.64)
Anxiousness	9	0.91	1.25	(0.78)	0.91	1.02	(0.73)
Attention seeking	8	0.88	0.86	(0.63)	0.89	0.81	(0.65)
Callousness	14	0.88	0.46	(0.44)	0.91	0.40	(0.50)
Deceitfulness	10	0.89	0.56	(0.54)	0.85	0.52	(0.54)
Depressivity	14	0.94	0.66	(0.65)	0.95	0.53	(0.62)
Distractibility	9	0.92	0.99	(0.72)	0.91	0.82	(0.69)
Eccentricity	13	0.95	0.98	(0.77)	0.96	0.82	(0.76)
Emotional lability	7	0.90	1.06	(0.74)	0.89	0.94	(0.74)
Grandiosity	6	0.73	0.84	(0.55)	0.72	0.82	(0.58)
Hostility	10	0.88	1.05	(0.63)	0.89	0.91	(0.67)
Impulsivity	6	0.85	0.85	(0.67)	0.77	0.77	(0.57)
Intimacy avoidance	6	0.84	0.79	(0.69)	0.84	0.61	(0.65)
Irresponsibility	7	0.80	0.46	(0.50)	0.81	0.39	(0.49)
Manipulativeness	5	0.80	0.81	(0.61)	0.81	0.80	(0.67)
Perceptual dysregulation	12	0.89	0.54	(0.55)	0.86	0.44	(0.48)
Perseveration	9	0.86	0.88	(0.61)	0.88	0.82	(0.62)
Restricted affectivity	7	0.75	0.95	(0.56)	0.73	0.97	(0.56)
Rigid perfectionism	10	0.89	1.08	(0.65)	0.90	1.05	(0.68)
Risk taking	14	0.88	1.09	(0.53)	0.85	1.05	(0.51)
Separation insecurity	7	0.86	0.80	(0.68)	0.85	0.80	(0.68)
Submissiveness	4	0.80	1.20	(0.68)	0.78	1.17	(0.66)
Suspiciousness	7	0.83	1.04	(0.67)	0.73	0.95	(0.58)
Unusual beliefs and experiences	8	0.85	0.66	(0.62)	0.83	0.64	(0.63)
Withdrawal	10	0.93	1.12	(0.71)	0.93	1.01	(0.72)
<b>Domain</b>							
Negative affect	53	0.94	1.16	(0.44)	0.93	1.07	(0.44)
Detachment	45	0.96	0.91	(0.54)	0.96	0.78	(0.54)
Antagonism	43	0.94	0.65	(0.42)	0.95	0.61	(0.46)
Disinhibition	46	0.89	1.13	(0.35)	0.84	1.06	(0.30)
Psychoticism	33	0.95	0.74	(0.56)	0.96	0.64	(0.57)

<sup>a</sup> Item responses were coded from 0 to 3, with higher scores indicating greater pathology. Statistics are based on list-wise deletion; therefore, individuals from Round 1 are not included due to planned missingness. Items from Restricted affectivity and Rigid perfectionism were reverse-scored for their inclusion in the Negative affect and Disinhibition domain scores, respectively.

standardized loadings on their facet of less than 0.5, those items were dropped and the factor analyses were re-run with the retained items. We chose 0.5 as the minimum loading for retaining an item because this is a relatively conservative value, corresponding with high correlations among all items within a facet scale; we found that a cut-off of 0.5 ensured our ability to create scales with relatively good information curves and high levels of internal consistency (see Supplementary material and Table 2 for evidence that this strategy was ultimately effective).

Once each facet had been reduced to items that (1) fit a one-factor model, and (2) all loaded highly on the

facet, we examined the results in an item response theory (IRT) framework (for a thorough description of IRT-based theory and methods, see Embretson & Reise, 2000). Using a two-parameter logistic IRT model (the two parameters are threshold and slope, with three thresholds per item in our case because the items had four response options) allowed us to estimate reliability for individuals contingent upon their latent trait level ( $\theta$ ). We estimated information for both individual items and the overall facet (or test information, the sum of information curves for each of the included items). Facets were considered well measured if their test information was greater than 5



(equivalent to a standard error of measurement  $<0.45$ ) for at least 3 standard deviations of  $\theta$ ; this approach resulted in scales that would also be deemed reliable via evaluation using coefficient  $\alpha$  (see Table 2).

### Results

Test information curves (TICs) for each of the facets indicated that we were able to measure most, but not all, facets reliably. The following six facets were not well measured by the Round 1 items: submissiveness, separation insecurity, attention seeking, oppositionality, perseverance, and rigidity. A total of 65 items (22% of those fielded) were dropped following Round 1, either because they (a) did not fit a single-factor model within their designated facet, or (b) had low loadings on their designated facet.

### Round 2

In the second round of data collection, we included additional items to (a) enhance the measurement of facets that were not measured reliably in Round 1 and (b) refine the structure of scales via item-level factor analysis.

### Participants

A total of 366 participants responded to the second wave of data collection. Participant selection and weighting scheme were identical to that in the first wave of data collection.

### Measure

In the second wave of data collection, all participants were asked to respond to all items. For those facets that were well measured in Round 1, we wrote items to replace any dropped items so that each facet had at least eight items going into Round 2 data collection. For facets not well measured in Round 1, we wrote items so that each facet had at least 10 items. This resulted in 231 original items and 85 new items (for a total of 316 items) being fielded as part of the second wave of data collection. Participant instructions and response options were the same as those given in Round 1. Items were administered in the same relative random order as in Round 1, with replacement items inserted into the survey in place of removed items.

### Analyses

To maximize the information included in the Round 2 analyses, and because they were collected using the same sampling scheme, data from Rounds 1 and 2

were combined for those items that had been presented in both phases. All data were again weighted by Knowledge Networks' provided sampling weights to adjust sample demographics to be representative of the US population, taking into account our selection based on participants having seen a therapist. For each facet that was not well measured in Round 1, we again ran an EFA within that facet. As before, if BIC indicated a solution with two or more factors fit best, we dropped items loading on all but the largest factor and confirmed through a subsequent EFA that a one-factor solution now fit the data best. We then examined one-factor solutions within each of the facets, again dropping items with standardized loadings less than 0.5 on the facet. From these factor analysis results, we re-examined the IRT-parameterized item and TICs to ensure that all of the facets now were measured well by the remaining items.

We next conducted EFAs of all items within each of the proposed domains, estimating factor solutions up to one greater than the hypothesized number of facets (e.g. if six facets were hypothesized, we examined solutions up to and including seven factors). Due to the computational complexity of estimating EFA solutions with a large number of factors for numerous polytomous items, these analyses were run in the program IRTPRO (Cai, 2010; Scientific Software International, 2010) using the default Oblique CF-Quartimax rotation and Metropolitan-Hastings Robbins-Monro estimator (Cai, 2010). From these EFAs, the solution with the best relative fit to the raw data, defined as the lowest BIC value, was selected for item assignment to final facets as long as that solution resulted in all factors being interpretable (i.e. each factor needed more than one item to load most highly on it). If the best-fitting model did not present with all interpretable factors, we selected the factor solution with the next lowest BIC. Items were assigned to the facet on which they had the highest loading in the best-fitting factor solution. We reanalysed each of the newly defined facets using one-factor models in Mplus (Muthén & Muthén, 1998–2010). Again, if any items had standardized loadings less than 0.5 on their assigned facet, they were dropped from further analyses. Facets were then 'trimmed' to ensure that no facet had a disproportionate number of items assigned to it. Using a target maximum of 10 items per facet, we removed items ensuring that the retained items (1) were representative of the full range of items that had been assigned to that facet based on the item-level within-domain EFAs, both in terms of item model parameters and content coverage; and (2) continued to measure the restructured facet reliably (that is, test information greater than 5 for at least 3 standard deviations of the latent trait,  $\theta$ ).

## Results

Our results following Round 2 data collection confirmed that all 37 of the originally proposed facets were well measured by the current item set. The results of item-level within-domain EFAs, however, suggested that the 37 facets could be collapsed into a more parsimonious set of 25 final traits. In two cases, the results of the EFA suggested that two facets could be collapsed, but we kept these facets separate with the aim of retaining separate scales corresponding with both pairs of constructs. Specifically, for the single facets identified as (1) anxiousness combined with suspiciousness and (2) callousness combined with grandiosity, we re-ran EFAs including only these items (as opposed to all the items in the corresponding domains). In both cases the best-fitting model, based on the lowest BIC value, included two separable factors. We therefore split the items within these facets based on these results (i.e. separating anxiousness from suspiciousness and separating callousness from grandiosity). Similarly, the original facets of recklessness and risk aversion (reverse-scored) were collapsed into a single risk-taking facet, based on prior evidence indicating a single dimension and a resulting well-fitting single-factor factor analysis. A representation of how facets were collapsed as a result of these within-domain item-level EFAs is shown in Table 1.

One-factor models were then fit within each restructured facet identified in the EFA analyses, with items assigned to the facet on which they loaded most highly in the EFA. Following the removal of items that had standardized loadings of less than 0.5 on their assigned facet, as well as paring down items within several facets that now contained an unnecessarily large number of items (e.g. depressivity), we were left with a set of 220 items that reliably measured all 25 traits, with four to 14 items included within each facet.

## Summary

Following the completion of the second wave of data collection, we were able to identify a reduced set of 25 empirically derived reliable facets. The entire measure, along with facet-scoring documentation, is available in Appendices A and B of the Supplementary material. We turn now to examining the norms and reliability of the measure in a population-representative sample (that is, one that was not selected for seeking mental health services).

## Round 3

### Participants

The third round of data collection was designed to examine the 25 identified facets in a representative

sample ( $n = 264$ ), using Knowledge Networks' weights to adjust the sample to be US population representative. Participants responded to the entire final set of 220 items, presented in the same relative random order with the same instructions and response options as those given in previous rounds.

## Results

All facets continued to meet IRT reliability criteria. In addition, Cronbach's  $\alpha$  for each of the facets ranged from 0.72 to 0.96 in the normative (population-representative) sample data, with a median of 0.86. Table 2 presents  $\alpha$  estimates and sum-score means and standard deviations for each facet (with results from the Round 2 data presented as well, for comparison purposes). Full item-level parameters for the final measure are presented in Appendix C of the Supplementary material, along with IRT item and TICs (IICs and TICs; see Supplementary Fig. S1 and Fig. S2). To take advantage of both the selected and representative samples, latent trait levels, or factor scores, on each of the 25 facets were estimated by constraining the latent factors to have a mean of 0 and variance of 1 in the representative sample, with means and variances freely estimated in the selected sample. Consistent with the sampling strategy difference between Rounds 1 and 2 and Round 3, the latent means were higher in the selected sample compared with the representative sample for 24 of 25 scales (the one exception was that the restricted affectivity mean was slightly lower in the Rounds 1 and 2 data) and the variances were similar; the exact values are available upon request. Facet factor scores estimated from this final model are used in all structural analyses described subsequently, unless otherwise indicated.

## Preliminary structure

### Facet correlations

Correlations among all facets within both the selected (Rounds 1 and 2) and representative (round 3) samples are presented in Supplementary Table S2. In general, the facets tended to be positively inter-correlated. The median correlation among facets within the selected sample was 0.30, while the median correlation among facets within the representative sample was 0.45.

### EFAs

To evaluate the number of factors to extract and interpret, we focused primarily on substantive interpretability, but also took into account the minimum average partial (MAP) and parallel analysis criteria (Zwick & Velicer, 1986); in addition, we estimated

**Table 3.** Five-factor solution from facet factor score EFA in selected samples (Rounds 1 and 2,  $n = 1128$ )

	Negative affect	Detachment	Antagonism	Disinhibition	Psychoticism
Emotional lability	0.69 <sup>a</sup>	0.00	0.02	0.12	0.19
Anxiousness	0.56 <sup>a</sup>	0.30 <sup>b</sup>	0.03	0.10	0.13
Restricted affectivity	-0.54 <sup>a</sup>	0.48 <sup>b</sup>	0.20	-0.01	0.17
Separation insecurity	0.47 <sup>a</sup>	-0.06	0.16	0.14	0.04
Hostility	0.38 <sup>a</sup>	0.25	0.34 <sup>b</sup>	0.21	-0.02
Perseveration	0.34 <sup>a</sup>	0.12	0.16	0.07	0.33 <sup>b</sup>
Submissiveness	0.27 <sup>a</sup>	0.08	0.11	-0.11	-0.03
Withdrawal	-0.02	0.68 <sup>a</sup>	-0.02	0.01	0.21
Anhedonia	0.14	0.61 <sup>a</sup>	-0.07	0.38 <sup>b</sup>	-0.04
Depressivity	0.33 <sup>b</sup>	0.46 <sup>a</sup>	-0.05	0.34 <sup>b</sup>	0.12
Intimacy avoidance	-0.14	0.44 <sup>a</sup>	-0.07	0.04	0.22
Suspiciousness	0.30 <sup>b</sup>	0.37 <sup>a</sup>	0.09	0.15	0.21
Manipulativeness	0.00	-0.12	0.76 <sup>a</sup>	-0.02	0.05
Deceitfulness	0.05	0.04	0.59 <sup>a</sup>	0.30 <sup>b</sup>	0.11
Grandiosity	-0.11	0.06	0.55 <sup>a</sup>	-0.03	0.17
Attention seeking	0.13	-0.21	0.51 <sup>a</sup>	0.08	0.13
Callousness	-0.18	0.23	0.48 <sup>a</sup>	0.38 <sup>b</sup>	0.05
Irresponsibility	0.04	0.11	0.12	0.57 <sup>a</sup>	0.13
Impulsivity	0.11	-0.10	0.08	0.57 <sup>a</sup>	0.12
Rigid perfectionism	0.31 <sup>b</sup>	0.30 <sup>b</sup>	0.37 <sup>b</sup>	-0.38 <sup>a</sup>	0.28
Distractibility	0.24	0.14	-0.06	0.34 <sup>a</sup>	0.28
Risk taking	-0.19	-0.28	0.21	0.31 <sup>a</sup>	0.30 <sup>b</sup>
Unusual beliefs and experiences	-0.01	-0.05	0.08	-0.09	0.76 <sup>a</sup>
Eccentricity	-0.05	0.09	0.04	0.21	0.63 <sup>a</sup>
Perceptual dysregulation	0.15	0.07	0.01	0.24	0.58 <sup>a</sup>
Factor intercorrelations					
Detachment	0.22				
Antagonism	0.11	0.14			
Disinhibition	0.26	0.31	0.24		
Psychoticism	0.27	0.39	0.48	0.37	

EFA, Exploratory factor analysis.

<sup>a</sup> The largest loading in each row.

<sup>b</sup> Loadings  $\geq 0.30$ , but which are not the highest loading within that row.

bootstrap confidence intervals around eigenvalues of the facet score correlation matrix using the open-source statistical package R (Team RDC, 2010). MAP suggested three factors, parallel analysis suggested six, and eigenvalues for the first five factors were significantly greater than 1 (the values were 8.68, 2.42, 1.89, 1.38 and 1.02, all  $p < 0.02$ ;  $p$  value for the 6th factor = 0.58).

Focusing on the goal of extracting the maximum number of interpretable dimensions within this potential range of three to six factors, a five-factor solution produced five readily interpretable factors, whereas the sixth factor was under-identified, with only two salient loadings (emotional lability and hostility). We therefore focus here on a five-factor EFA, using the CF-Equamax oblique rotation in Mplus (Muthén & Muthén, 1998–2010). We chose EFA over confirmatory factor analysis (CFA) because strong

*a priori* hypotheses about the structure of these scales would be premature, and because simple structure CFA models are not realistic for personality data, in that many manifest personality variables tend to show meaningful factor cross loadings (Hopwood & Donnellan, 2010). We chose the CF-Equamax rotation because it aims to simplify both variable and factor complexity, spreading variances across the factors and combining features of both the Quartimax and Varimax criteria (Browne, 2001); we used the oblique version of Equamax because the general positive manifold among the facets (Supplementary Table S2) suggests that the factors are probably best conceived of as being correlated. The five-factor EFA loadings are presented in Table 3. The results correspond to domains that could be labeled negative affect, detachment, antagonism, disinhibition, and psychoticism. Of note, some facet scales are relatively 'pure' markers of



these domains (e.g. manipulateness, unusual beliefs and experiences), whereas others (e.g. hostility) are located 'in between' domains because they share features of more than one domain (see e.g. Markon *et al.* 2005).

To investigate the appropriateness of the five-factor solution identified in the selected samples when applied to a normative population (i.e. the representative sample in Round 3), we estimated a targeted-rotation EFA in Mplus. For each of the five factors, the targeted loading of the 25 facets was specified as the estimated loading in the selected samples. We then estimated factor congruences (using R, see <http://www.personality-project.org/R/html/factor.congruence.html>) between the five-factor solutions in the selected and representative samples. Factor loadings were very similar across the selected and representative samples for each of the domains. For negative affect, the congruence coefficient was 0.91; detachment = 0.96; antagonism = 0.97; disinhibition = 0.93; and psychoticism = 0.89.

## Discussion

The DSM-5 Personality and Personality Disorders Workgroup members and consultants pursued the creation of a maladaptive personality trait model and corresponding assessment instrument, drawing on and extending models in the existing literature (for recent reviews, see Trull & Durrett, 2005, Widiger & Simonsen, 2005, Clark, 2007; Krueger & Eaton, 2010). Our goal was to develop a model and corresponding assessment instrument (which can be referred to as 'a Personality Inventory for DSM-5' or PID-5) that is freely available to all interested clinicians and researchers; the inventory is given here as part of the online Supplement that accompanies this paper. In addition, other investigators are free to develop other assessment tools to assess DSM-5-relevant personality traits. One critical example along these lines is work on the Computerized Adaptive Assessment of Personality Disorder (CAT-PD) project (Simms *et al.* 2011). The CAT-PD project will result in a highly comprehensive multidimensional assessment instrument that can be administered efficiently via adaptive testing, and is thereby likely to have a fundamental impact on ongoing revisions of the DSM, as well as the broader field.

The major aim of the current project was realized. We created a publicly available inventory with scales to assess 25 maladaptive personality traits. Our preliminary factor analytic results suggest these traits delineate five higher-order domains that constitute a model synthesizing other empirically based maladaptive trait and disorder models in the literature (Harkness *et al.* 1995; Austin & Deary, 2000; Costa &

Widiger, 2002; De Clercq *et al.* 2006; Clark, 2007; Livesley, 2007; Nestadt *et al.* 2008; Watson *et al.* 2008; Piedmont *et al.* 2009; Widiger *et al.* 2009; Markon, 2010; Pincus *et al.* 2010; Kendler *et al.* 2011; Røysamb *et al.* 2011).

Nevertheless, there are a number of important limitations to this work. First, to date, our instrument development efforts have been limited to self-report. A thorough clinical assessment of maladaptive traits should rely on reporters in addition to the self (Oltmanns & Turkheimer, 2009), and future work should focus on developing assessment instruments for use by other informants. Moreover, the assessment of personality and the diagnosis of PD require careful clinical judgment, judgment that can be informed – but not replaced by – self-report trait assessments. It will be important to evaluate the use of the trait model by clinicians in field trials (Kraemer *et al.* 2010). Second, our work has focused on the assessment of personality traits in adults. Further work is needed to determine if our assessment instrument can be adapted for use with other age groups (e.g. adolescents).

Third, our assessment instrument is focused specifically on maladaptive personality traits, i.e. those associated with DSM-IV-TR PDs. Given the limited time available in many clinical encounters, maladaptive traits arguably represent the more pressing areas of clinical concern, relative to adaptive traits. In addition, maladaptive qualities are the traditional focus of the DSM. Nevertheless, adaptive traits also have clinical utility (Harkness & Lilienfeld, 1997), and the trait model and assessment instrument we delineate here could be expanded to encompass these characteristics in future work. For example, our inventory lacks measures of openness to experience because openness generally is not highly correlated with PD. Our assessment could be expanded to encompass traits in this domain, or alternatively, a more comprehensive clinical assessment could augment our instrument with another designed to assess openness or other adaptive traits.

Fourth, our starting point was workgroup discussion regarding elements of maladaptive personality variation. Numerous additional maladaptive personality variants could, in theory, be identified. We look forward to future research on how these additional dimensions may relate to the dimensions operationalized in our provisional inventory.

Fifth, we worked with statistical models that do not empirically estimate the form of the underlying latent trait distribution because these models are well-established psychometric models for initial scale creation in personality research (Simms & Watson, 2007). However, future research designed to explore the empirical latent trait distribution underlying the

scales we developed (e.g. via Johnson curves; van den Oord, 2005) would be valuable.

Sixth, additional validity research would be needed before these scales could be deemed appropriate for application in clinical settings (Cella et al. 2010). Our intent in conveying the current results at this juncture is to allow the broader field to explore the potential validity of these constructs and their corresponding scales, in the hopes that an empirical literature will coalesce, that can be referred to as the DSM continues to develop. We hope this can result in an approach to the development of the DSM that relies more on the accumulation of data, as opposed to the more traditional approach of investigators adopting DSM conventions, without pursuing research that questions those conventions and thereby aims to improve the reliability and validity of diagnosis. For example, although the provisional scales we provide here showed generally good internal consistency reliabilities (see Table 2), it may be desirable to further enhance some of the less reliable and shorter facet-level scales prior to contemplating their use in clinical settings. In addition, the preliminary factor structure we report here must be replicated in other samples before it could be considered a definitive account of the structure of our provisional scales, much less the broader domain of maladaptive personality variation.

## Note

Supplementary material accompanies this paper on the Journal's website (<http://journals.cambridge.org/psm>).

## Acknowledgements

The present study was supported by a grant from the American Psychiatric Institute for Research and Education (APIRE). J.D. is supported by National Institutes of Health (NIH) grant DA029377. R.F.K. is supported by NIH grant no. AA018111. We thank Darrel A. Regier for his helpful comments on a previous draft of this paper.

## Declaration of Interest

None.

## References

- APA (1994). *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn, revised, p. 685. American Psychiatric Association: Washington, DC.
- Asparouhov T (2005). Sampling weights in latent variable modeling. *Structural Equation Modeling* **12**, 411–434.

- Austin EJ, Deary IJ (2000). The 'four As': a common framework for normal and abnormal personality? *Personality and Individual Differences* **28**, 977–995.
- Browne MW (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research* **36**, 111–150.
- Cai L (2010). High-dimensional exploratory item factor analysis by a Metropolis–Hastings Robbins–Monro algorithm. *Psychometrika* **75**, 33–57.
- Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, Amtmann D, Bode R, Buysse DJ, Choi SW, Cook KF, DeVellis R, DeWalt D, Fries JF, Gershon R, Hahn E, Pilkonis P, Revicki D, Rose M, Weinfurt K, Hays RD, on behalf of the PROMIS Cooperative Group (2010). The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *Journal of Clinical Epidemiology* **63**, 1179–1194.
- Chmielewski M, Watson D (2008). The heterogeneous structure of schizotypal personality disorder: item-level factors of the schizotypal personality questionnaire and their associations with obsessive-compulsive disorder symptoms, dissociative tendencies, and normal personality. *Journal of Abnormal Psychology* **117**, 364–376.
- Clark LA (2007). Assessment and diagnosis of personality disorder: perennial issues and emerging conceptualization. *Annual Review of Psychology* **58**, 227–258.
- Clark LA, Simms LJ, Wu KD, Casillas A (in press). *Manual for the Schedule for Nonadaptive and Adaptive Personality – Second Edition (SNAP-2)*. University of Minnesota Press: Minneapolis, MN.
- Costa Jr. PT, McCrae RR (1992). *Revised NEO Personality Inventory NEO-PI-R and NEO Five-Factor Inventory NEO-FFI Professional Manual*. Psychological Assessment Resources: Odessa, FL.
- Costa Jr. PT, Widiger TA (editors) (2002). *Personality Disorders and the Five-Factor Model of Personality*, 2nd edn. American Psychological Association: Washington, DC.
- Cuijpers P, Smit F, Penninx BWJH, de Graaf R, ten Have M, Beekman ATF (2010). Economic costs of neuroticism: a population-based study. *Archives of General Psychiatry* **67**, 1086–1093.
- De Clercq B, De Fruyt F, Van Leeuwen K, Mervielde I (2006). The structure of maladaptive personality traits in childhood: a step toward an integrative developmental perspective for DSM-5. *Journal of Abnormal Psychology* **115**, 639–657.
- de Moor MHM, Costa PT, Terracciano A, Krueger RF, de Geus EJC, Toshiko T, Penninx BW, Esko T, Madden PA, Derringer J, Amin N, Willemsen G, Hottenga JJ, Distel MA, Uda M, Sanna S, Spinhoven P, Hartman CA, Sullivan P, Realo A, Allik J, Heath AC, Pergadia ML, Agrawal A, Lin P, Gruzza R, Nutile T, Cuiullo M, Rujescu D, Giegling I, Konte B, Widen E, Cousminer DL, Eriksson JG, Palotie A, Peltonen L, Luciano M, Tenesa A, Davies G, Lopez LM, Hansell NK, Medland SE, Ferrucci L, Schlessinger D, Montgomery GW, Wright MJ, Aulchenko YS, Janssens AC, Oostra BA, Metspalu A, Abecasis GR, Deary IJ, Rääkkönen K, Bierut LJ, Martin NG, van Duijn CM, Boomsma DI

- (2010). Meta-analysis of genome-wide association studies for personality. *Molecular Psychiatry*. Published online: 21 December 2010. doi:10.1038/mp.2010.128.
- Dennis JM** (2010). Summary of Knowledge Panel Design (<http://www.knowledgenetworks.com/knpanel/docs/KnowledgePanel%28R%29-Design-Summary-Description.pdf>). Accessed 11 November 2011.
- Embretson SE, Reise SP** (2000). *Item Response Theory for Psychologists*. Lawrence Erlbaum Associates, Inc.: Mahwah, NJ.
- Goldberg LR** (1993). The structure of phenotypic personality traits. *American Psychologist* **48**, 26–34.
- Harkness AR, Lilienfeld SO** (1997). Individual differences science for treatment planning: personality traits. *Psychological Assessment* **9**, 349–360.
- Harkness AR, McNulty JL, Ben-Porath YS** (1995). The personality psychopathology five (PSY-5): constructs and MMPI-2 scales. *Psychological Assessment* **7**, 104–114.
- Hopwood CJ, Donnellan MB** (2010). How should the internal structure of personality inventories be evaluated? *Personality and Social Psychology Review* **14**, 332–346.
- Hopwood CJ, Quigley BD, Morey LC, Bender DS, Skodol AE, Stout RL, McGlashan TH, Grilo CM, Gunderson JG** (2008). Personality traits and mental health treatment utilization. *Personality and Mental Health* **2**, 207–217.
- Kendler KS, Aggen SH, Knudsen GP, Røysamb E, Neale MC, Reichborn-Kjennerud T** (2011). The structure of genetic and environmental risk factors for syndromal and subsyndromal common DSM-IV axis I and all axis II disorders. *American Journal of Psychiatry* **168**, 29–39.
- Kraemer HC, Kupfer DJ, Narrow WE, Clarke DE, Regier DA** (2010). Moving toward DSM-5: the field trials. *American Journal of Psychiatry* **167**, 1158–1160.
- Krueger RF, Eaton NR** (2010). Personality traits and the classification of mental disorders: toward a more complete integration in DSM-5 and an empirical model of psychopathology. *Personality Disorders: Theory, Research, and Treatment* **1**, 97–118.
- Livesley WJ** (2007). A framework for integrating dimensional and categorical classifications of personality disorder. *Journal of Personality Disorders* **21**, 199–224.
- Livesley WJ, Jackson DN** (2009). *Manual for the Dimensional Assessment of Personality Pathology*. Sigma Press: Port Huron, MI.
- Markon KE** (2010). Modeling psychopathology structure: a symptom-level analysis of axis I and II disorders. *Psychological Medicine* **40**, 273–288.
- Markon KE, Krueger RF, Watson D** (2005). Delineating the structure of normal and abnormal personality: an integrative hierarchical approach. *Journal of Personality and Social Psychology* **88**, 139–157.
- Morey LC** (2003). *Essentials of PAI Assessment*. John Wiley: New York.
- Muthén LK, Muthén BO** (1998–2010). *MPlus Users Guide*, 6th edn. Muthén & Muthén: Los Angeles, CA.
- Nestadt G, Costa Jr. PT, Hsu F-C, Samuels J, Bienvenu OJ, Eaton WW** (2008). The relationship between the five-factor model and latent *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* personality disorder dimensions. *Comprehensive Psychiatry* **49**, 98–105.
- Oltmanns TF, Turkheimer E** (2009). Person perception and personality pathology. *Current Directions in Psychological Science* **18**, 32–36.
- Piedmont RL, Sherman MF, Sherman NC, Dy-Liacco GS, Williams JE** (2009). Using the five-factor model to identify a new personality disorder domain: the case for experiential permeability. *Journal of Personality and Social Psychology* **96**, 1245–1258.
- Pincus AL, Lukowitsky MR, Wright AGC** (2010). The interpersonal nexus of personality and psychopathology. In *Contemporary Directions in Psychopathology: Scientific Foundations for DSM-V and ICD-11* (ed. T. Millon, R. Krueger and E. Simonsen), pp. 523–552. Guilford: New York.
- Røysamb E, Kendler KS, Tambs K, Orstavik RE, Neale MC, Aggen SH, Torgersen S, Reichborn-Kjennerud T** (2011). The joint structure of DSM-IV axis I and axis II disorders. *Journal of Abnormal Psychology* **120**, 198–209.
- Samuel DB, Widiger TA** (2008). A meta-analytic review of the relationships between the five-factor model and DSM-IV-TR personality disorders: a facet level analysis. *Clinical Psychology Review* **28**, 1326–1342.
- Schwarz GE** (1978). Estimating the dimension of a model. *Annals of Statistics* **6**, 461–464.
- Scientific Software International, Inc.** (2010). *IRTPRO Application*, version 1.1.20215.1 BETA (<http://www.ssicentral.com>). Accessed August 2010.
- Simms L, Goldberg LR, Roberts JE, Watson D, Welte J, Rotterman JH** (2011). Computerized Adaptive Assessment of Personality Disorder: introducing the CAT-PD project. *Journal of Personality Assessment* **93**, 380–389.
- Simms L, Watson D** (2007). The construct validation approach to personality scale construction. In *Handbook of Research Methods in Personality Psychology* (ed. R. W. Robins, R. C. Fraley and R. F. Krueger), pp. 240–258. Guilford: New York.
- Smits N, Vorst HCM** (2007). Reducing the length of questionnaires through structurally incomplete designs: an illustration. *Learning and Individual Differences* **17**, 25–34.
- Tang TZ, DeRubeis RJ, Hollon SD, Amsterdam J, Shelton R, Schalet B** (2009). Personality change during depression treatment: a placebo-controlled trial. *Archives of General Psychiatry* **66**, 1322–1330.
- Team RDC** (2010). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna.
- Trull TJ, Durrett CA** (2005). Categorical and dimensional models of personality disorder. *Annual Review of Clinical Psychology* **1**, 355–380.
- van den Oord EJCG** (2005). Estimating Johnson curve population distributions in MULTILOG. *Applied Psychological Measurement* **29**, 45–64.
- van den Oord EJCG, Kuo P-H, Hartmann AM, Webb BT, Moller H-J, Hetteima JM, Giegling I, Bukszár J, Rujescu D** (2008). Genomewide association analysis followed by a replication study implicates a novel candidate gene for neuroticism. *Archives of General Psychiatry* **65**, 1062–1071.
- Watson D, Clark LA, Chmielewski M** (2008). Structures of personality and their relevance to

psychopathology: II. Further articulation of a comprehensive unified trait structure. *Journal of Personality* **76**, 1545–1586.

**Widiger TA, Livesley WJ, Clark LA** (2009). An integrative dimensional classification of personality disorder. *Psychological Assessment* **21**, 243–255.

**Widiger TA, Simonsen E** (2005). Alternative dimensional models of personality disorder: finding a common ground. *Journal of Personality Disorders* **19**, 110–130.

**Zwick WF, Velicer WF** (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin* **99**, 432–442.