

Testing Whether the *DSM-5* Personality Disorder Trait Model Can Be Measured With a Reduced Set of Items: An Item Response Theory Investigation of the Personality Inventory for *DSM-5*

Jessica L. Maples and Nathan T. Carter
University of Georgia

Lauren R. Few
Washington University School of Medicine

Cristina Crego and Whitney L. Gore
University of Kentucky

Douglas B. Samuel
Purdue University

Rachel L. Williamson
University of Georgia

Donald R. Lynam
Purdue University

Thomas A. Widiger
University of Kentucky

Kristian E. Markon
University of Iowa

Robert F. Krueger
University of Minnesota

Joshua D. Miller
University of Georgia

The fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)* includes an alternative model of personality disorders (PDs) in Section III, consisting in part of a pathological personality trait model. To date, the 220-item Personality Inventory for *DSM-5* (PID-5; Krueger, Derringer, Markon, Watson, & Skodol, 2012) is the only extant self-report instrument explicitly developed to measure this pathological trait model. The present study used item response theory-based analyses in a large sample ($n = 1,417$) to investigate whether a reduced set of 100 items could be identified from the PID-5 that could measure the 25 traits and 5 domains. This reduced set of PID-5 items was then tested in a community sample of adults currently receiving psychological treatment ($n = 109$). Across a wide range of criterion variables including NEO PI-R domains and facets, *DSM-5* Section II PD scores, and externalizing and internalizing outcomes, the correlational profiles of the original and reduced versions of the PID-5 were nearly identical ($r_{ICC} = .995$). These results provide strong support for the hypothesis that an abbreviated set of PID-5 items can be used to reliably, validly, and efficiently assess these personality disorder traits. The ability to assess the *DSM-5* Section III traits using only 100 items has important implications in that it suggests these traits could still be measured in settings in which assessment-related resources (e.g., time, compensation) are limited.

Keywords: *DSM-5*, assessment, personality, personality disorders

Supplemental materials: <http://dx.doi.org/10.1037/pas0000120.supp>

The transition from a categorical to a dimensional model of personality disorders (PDs) has been regarded as inevitable for over 20 years (Frances, 1993) due to widespread acknowledgment of the many limitations associated with categorical approaches

(Clark, 2007). For instance, categorical diagnoses of PD are highly heterogeneous, demonstrate a high degree of comorbidity, and are not generally reliable (e.g., Trull & Durrett, 2005). Significant empirical support exists for the transition toward dimensional

This article was published Online First April 6, 2015.

Jessica L. Maples and Nathan T. Carter, Department of Psychology, University of Georgia; Lauren R. Few, Department of Psychiatry, Washington University School of Medicine; Cristina Crego and Whitney L. Gore, Department of Psychology, University of Kentucky; Douglas B. Samuel, Department of Psychology, Purdue University; Rachel L. Williamson, Department of Psychology, University of Georgia; Donald R. Lynam, Department of Psychology, Purdue University; Thomas A. Widiger,

Department of Psychology, University of Kentucky; Kristian E. Markon, Department of Psychology, University of Iowa; Robert F. Krueger, Department of Psychology, University of Minnesota; Joshua D. Miller, Department of Psychology, University of Georgia.

Correspondence concerning this article should be addressed to Jessica L. Maples or Joshua D. Miller, Department of Psychology, University of Georgia, Athens, GA 30602-3013. E-mail: jessmaples@gmail.com or jdmiller@uga.edu

models (Widiger & Trull, 2007) as they offer the ability to help resolve a number of important issues including coverage, patterns of comorbidity, and gender differences in diagnostic rates, to name just a few important considerations (e.g., Krueger & Tackett, 2003; Lynam & Widiger, 2007).

Consistent with the research demonstrating the validity of trait based approaches to the conceptualization, assessment, and diagnosis of PDs (e.g., Widiger & Costa, 2012), the *DSM-5* Personality & Personality Disorders Work Group (*DSM-5* PPD WG) proposed a diagnostic model of PDs for *DSM-5* that included an assessment of general impairment in personality functioning (i.e., Criterion A) and a dimensional model of pathological personality traits (i.e., Criterion B). Specifically, Criterion A requires “moderate or greater impairment in personality (self/interpersonal) functioning.” Criterion B involves the presence of one or more pathological personality traits from five broad pathological trait domains, including negative affectivity, detachment, antagonism, disinhibition, and psychoticism, and 25 specific facets (e.g., emotional lability, eccentricity) nested within these trait domains. These domains are noted within the *DSM-5* to be “maladaptive variants of the extensively validated and replicated model of personality known as the ‘Big Five,’ or Five-Factor Model of personality (FFM), and are also similar to the domains of the Personality Psychology Five (PSY-5)” (APA, 2013, p. 773).

In the *DSM-5* Section III alternative diagnostic model, PD diagnoses are based on an individual’s match to the pattern of personality impairment and elevations on maladaptive traits that are specified for six PD types retained from the categorical diagnostic system. This model also includes the diagnosis of Personality Disorder-Trait Specified (PD-TS), which would be used for when a PD is present but criteria for one of the six specific PDs are not fully met. The trait-based model was included in Section III of *DSM-5* for “emerging measures and models” in order to encourage further study and evaluation of its reliability, validity, and clinical utility. The inclusion of this proposal in Section III suggests that it may be incorporated as the only system in future iterations of the *DSM*, depending on its performance in empirical investigations over the next several years.

A crucial instrument in the testing of the *DSM-5* Section III PD model is the inventory created to measure this new five-factor trait model of PDs, the Personality Inventory for *DSM-5* (PID-5; Krueger, Derringer, Markon, Watson, & Skodol, 2012). In order to construct the PID-5, the *DSM-5* PPD WG first reviewed existing personality models (e.g., Widiger & Simonsen, 2005) and generated a list of 37 facets for which brief definitions were created, followed by writing items for each facet. Using an iterative approach across three rounds of data collection, item response theory (IRT) models were used to select PID-5 items and exploratory factor analyses were used to refine the factor structure. In the final round of data collection, 220 items and 25 facets were investigated and results supported the reliability of all facets and existence of a five-factor structure (Krueger et al., 2012).

A number of studies on the reliability and validity of the PID-5 have now been published (Krueger & Markon, 2014), including those replicating the five-factor structure of the PID-5 alone (e.g., Wright, Thomas, et al., 2012) and in concert with other FFM measures (Gore & Widiger, 2013; Thomas et al., 2013). In a sample of psychiatric patients gathered as part of the *DSM-5* field trial, PID-5 domains and facets manifested acceptable reliabilities

and the domains converged with the corresponding FFM domains (Quilty, Ayeart, Chmielewski, Pollock, & Bagby, 2013). A recent IRT based investigation of the PID-5 and the International Personality Item Pool-NEO (IPIP-NEO; Goldberg, 1999), a measure of the FFM, found that both measures demonstrates largely overlapping information curves of four of the five domains, providing evidence that the PID-5 represents maladaptive versions of the FFM (Suzuki, Samuel, Pahlen, & Krueger, 2015). In a sample of community adults receiving psychological treatment, the PID-5 demonstrated good convergence with interviewer ratings of the *DSM-5* pathological traits, accounted for substantial variance in semistructured interview based ratings of *DSM-5* Section II PDs, and manifested expected relations with external criteria including FFM domains, anxiety and depression symptomatology, alcohol and drug use, and antisocial behaviors (Few et al., 2013). Finally, several recent studies have found that PID-5 traits proposed for use in the diagnosis of the six retained PD types explained a substantial portion of the variance in self-report *DSM-IV* PD scores (Samuel, Hopwood, Krueger, Thomas, & Ruggero, 2013; Hopwood, Thomas, Markon, Wright, & Krueger, 2012; Yam & Simms, 2014).

Another important area of investigation is clarifying the relations between the PID-5 scales and those from measures of other prominent personality models. Research in this area is rapidly expanding as the PID-5 has been investigated in relation to the Personality Assessment Inventory (Morey, 1991), Schedule for Nonadaptive and Adaptive Personality (Clark, 1993), HEXACO Personality Inventory (Ashton & Lee, 2007), Computerized Adaptive Test–Personality Disorder (Simms et al., 2011), interpersonal circumplex (Wiggins & Pincus, 2002), and the Personality Psychopathology-5 (Harkness, McNulty, & Ben-Porath, 1995; e.g., Anderson et al., 2013; Ashton et al., 2012; Hopwood et al., 2013; Thomas et al., 2013; Watson, Stasik, Ro, & Clark, 2013; Wright & Simms, 2014; Wright, Pincus, et al., 2012). The PID-5 has also been shown to demonstrate expected relations with measures of individual forms of personality pathology, such as narcissism (e.g., Miller, Gentile, Wilson, & Campbell, 2013) and psychopathy (e.g., Strickland, Drislane, Lucy, Krueger, & Patrick, 2013).

To date, the PID-5 is the only self-report instrument explicitly developed to measure the Section III PD traits and, as such, has had a significant effect on the literature. One potential constraint to the development of a robust literature on the *DSM-5* Section III PD model, particularly in clinical settings, is the length of the instrument. The PID-5 uses 220 items to assess the five domains and 25 more specific facets, which may be prohibitively long in certain circumstances, particularly because much research with the PID-5 will require that it is only one assessment in a much larger protocol. As such, there may be concerns regarding participant fatigue and cost (in terms of compensation for participation) associated with its length. Although a brief version of the PID-5 exists (i.e., the PID-5-Brief Form; 25-items www.psychiatry.org/File%20Library/Practice/DSM/DSM-5/ThePersonalityInventoryForDSM5BriefFormAdult.pdf), its utility is more limited as it targets only the five higher order domains. Ultimately, the transition from the *DSM-IV/5* categorical approach to a dimensional model like that contained in Section III of the *DSM-5* will require the accumulation of a substantial research literature supporting such a transition. To ensure that this occurs, it is worth investigating whether a

reduced set of PID-5 items can be identified that can measure the *DSM-5* traits in an efficient yet reliable and valid manner. In the current study, IRT analyses were used to test whether it was possible to identify an abridged set of PID-5 items with the aforementioned properties.

Previous studies provide support for the possibility of identifying a smaller set of items from longer and well-validated measures that do not result in a significant reduction of reliability or validity. For example, a previous study identified 136 of the original 290 items from the Dimensional Assessment of Personality Pathology-Basic Questionnaire (DAPP-BQ; Livesley & Jackson, 2009) and used these to form abbreviated DAPP-BQ scales that demonstrated appropriate reliability, convergence with original DAPP-BQ scales, and replicated the factor structure of the longer scales (van Kampen, de Beurs, & Andrea, 2008). A review of different approaches used to abbreviate the time and length required to administer the Minnesota Multiphasic Personality Inventory (MMPI; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989) identified 14 studies that developed short form versions of this measure (Butcher & Hostetler, 1990). Recently, IRT methodology was used to create a 120 item version of the 300-item IPIP-NEO (Goldberg, 1999) with the resultant scales demonstrating a high degree of similarity with regard to their correlational profiles across a wide range of criterion variables (Maples, Guan, Carter, & Miller, 2014).

IRT methodologies are well suited for such purposes (e.g., Edelen & Reeve, 2007), as the detailed item-level information that is provided allows for a more in-depth examination of the properties of items and their relations with latent traits. As such, a subset of items can be identified that maximize reliability while measuring individuals similarly well across a range of trait levels. Given that larger sample sizes result in more stable results for IRT parameter estimates, the present study used a large sample ($n = 1,417$) in order to investigate which PID-5 items most optimally discriminate between people at different trait levels (i.e., provide measurement precision). Results from these analyses were then used to identify four items per facet, or 100 items in total, that provide adequate content coverage of the traits along with maximum precision in their measurement.¹

Next, this set of 100 items was cross-validated by comparing it with the original version in an independent sample of community sample of individuals currently receiving psychological treatment ($n = 109$; see Few et al., 2013 for details). The factor structure of the abbreviated PID-5 scales was first tested in relation to the longer scales in a sample that combined the aforementioned derivation and validation samples. The correlations between the full and reduced length scales across both the derivation and validation samples were also considered. Following these analyses, the performance of the reduced PID-5 scales in relation to the original scales in this independent validation sample was considered, using a wide variety of outcomes, including NEO Personality Inventory-R domains, clinician rated PID-5 facets and domains, *DSM-5* Section II PD scores, and internalizing and externalizing outcomes. The convergent validity of the reduced PID-5 scales with the full length PID-5 scales and clinician-rated *DSM-5* traits was investigated, as was the discriminant validity of the reduced PID-5 compared with the other PID-5 domains. Previous research has shown that the average discriminant validity correlations across the PID-5 domains were .46 (Quilty et al., 2013), as such it

was expected that the discriminant validity for both the full length and reduced PID-5 scales would demonstrate correlations in this range as well.

The criterion validity of the reduced PID-5 scales was investigated in relation to the FFM trait domains, *DSM-IV/5* categorical PDs, and internalizing and externalizing outcomes. Given that the PID-5 trait model represents maladaptive variants of the FFM traits, it was expected that each domain would demonstrate its strongest absolute correlation with its counterpart (e.g., FFM antagonism and PID-5 Antagonism). The Section III trait model conceptualizes the categorical PDs as representing configurations of the PID-5 traits, as such it was expected that the PID-5 traits would demonstrate significant correlations with many of the *DSM-IV/5* PDs. Finally, based on previous FFM research, it was expected that negative affectivity would demonstrate the strongest relations with internalizing outcomes (e.g., Kotov, Gamez, Schmidt, & Watson, 2010), and that antagonism and disinhibition would demonstrate the strongest relations with externalizing outcomes (e.g., Miller & Lynam, 2001).

Method

Participants and Procedure

Derivation sample. The derivation sample consisted of 1,417 adults (59.4% female; *Mean* age = 25.6; *SD* = 10.7; 71% White; 5.8% Black; 18.6% Asian) composed of individuals from four different samples. The first sample consisted of adults ($n = 306$; see Miller, Gentile, Wilson, & Campbell, 2013 for details) who were recruited via Amazon's Mechanical Turk (MTurk) Web site and compensated \$2 for completion of online surveys. This sample was 57% male, 49% Asian, 46% Caucasian, with a mean age of 29.7 (*SD* = 10.2). Twenty-one individuals were removed from the original data ($n = 327$) due to extensive missing data or obviously invalid responses. The second sample consisted of undergraduate introductory psychology students ($n = 445$; see Gore & Widiger, 2013 for details) who received research credit for participation. The majority of this sample was female (67%), the mean age was 19.23 years, and 83% identified as Caucasian, 10% as African American, 2% as Asian, 1% as Hispanic, and 3% as other. From the original sample ($n = 585$), 100 (17%) participants were removed because they did not adequately complete the measures and 40 (7%) were deleted due to elevated scores on a validity scale. The third sample consisted of community adults ($n = 330$; see Rojas & Widiger, 2014 for details) recruited from MTurk who

¹ The subsmissiveness facet only contains four items in the original version of the PID-5, as such all of the items were retained for the reduced version. IRT analyses were still conducted on the original four items of this facet and RMSEA was .05, indicating adequate model data fit and suggesting that the item parameters were interpretable. Additionally, given that four items have been shown to successfully measure the subsmissiveness facet, we initially chose four items as the threshold for the remainder of the scales in order to match the shortest existing and validated facet scale. IRT analyses also suggested that four items per scale provided a good balance between appropriate breadth and instrument efficiency. Additionally, we compared three-, four-, and five-item versions of scales for five randomly selected traits in order to ensure that four items were the most appropriate. The RMSEA values and TIFs for these analyses are provided in supplemental materials; these results provide support for the use of four-item scales.

completed online questionnaires. Thirty-three individuals were removed from the original sample due to either not completing at least 80% of the administered questionnaires, or due to elevated scores on a validity scale. This majority of this sample was female (63%), had a mean age of 35.1 ($SD = 12.9$), and 77.3% identified as White/Caucasian, 6.4% as Black/African American, 6.1% Asian, 4.5% Hispanic/Latino, 1.2% American Indian/Alaskan Native, .6% Native Hawaiian/Pacific Island, and 3.3% Other. The fourth sample consisted of undergraduate students ($n = 336$; see Griffin & Samuel, 2014) who received class credit for participation. Fifty-two participants were removed from the original sample ($n = 388$) due to excessive missing data and/or nonvalid responses, such as long strings of identical answers. The mean age was 19.4 ($SD = 2$), 61% identified as female, and 65% identified as White, 20% Asian, and 4% African American.

Although the current data were collected from four different samples, previous research has demonstrated measurement invariance across undergraduate and MTurk samples (Behrend, Sharek, Meade, & Wiebe, 2011) as well as comparable results from MTurk and undergraduate samples (Johnson & Borden, 2012; Paolacci, Chandler, & Ipeiritos, 2010; Sprouse, 2011). Additionally, all participants were drawn from the same subpopulation (i.e., Adult, American). In general, no substantive rationale existed to suspect that persons in these samples had large enough differences in psychosocial experience to suspect that their interpretations of, or responses to, item content would be different (i.e., noninvariant; see Carter, Kotrba, & Lake, 2014).²

Validation sample. Participants were 109 adults currently receiving psychological or psychiatric treatment (71% female; Mean age = 35.9; $SD = 12.7$; 91% White; 6% Black; 3% Asian; see Few et al., 2013). Data from 110 participants were collected; one participant was removed from the final sample due to non-compliance. Participants were recruited via advertisements placed in local newspapers and mental health treatment facilities which called for individuals “currently receiving mental health treatment.” Interested individuals contacted the laboratory via telephone, provided verbal consent, and were administered a screening questionnaire. Individuals could not participate if they were experiencing psychotic symptoms or were currently receiving inpatient treatment. During the 3-hr laboratory assessment session, a trained graduate student interviewer conducted a videotaped, semistructured Axis II diagnostic assessment and completed interview ratings of pathological personality traits and impairment. Observer ratings of pathological personality traits and impairment were completed using the videotaped interview. Following the interview, participants completed several self-report assessments, were debriefed, and compensated \$30.

Measures Included in Both the Derivation and Validation Sample

Demographic form. A brief demographic questionnaire was administered to all participants assessing race, sex, and age.

Personality inventory for DSM-5 (PID-5; Krueger et al., 2012). The PID-5 is a 220-item measure of the 25 personality traits of DSM-5 PD trait model. The facets each contain between four and 14 items. For each item, participants rate how much a statement describes themselves on a Likert scale that ranges from 0 (*very false or often false*) to 3 (*very true or often true*). Section III of the DSM-5

specifies that one or more of the 25 traits must be elevated in order for a personality disorder to be present. Additionally, Section III retains six of the 10 DSM-IV PDs (i.e., schizotypal, antisocial, borderline, narcissistic, avoidant, obsessive-compulsive), and specifies which traits should be elevated in order to meet criteria for diagnosis of each disorder. For instance, antisocial personality disorder requires that six of seven traits must be elevated, including manipulativeness, callousness, deceitfulness, hostility, risk-taking, impulsivity, and irresponsibility, and self and interpersonal dysfunction must be present. Section III also includes a diagnosis of Personality Disorder-Trait Specified, in which the criteria for one of the six specific PDs is not met, but impairment in personality functioning and elevations on one or more specific domains or facets are present. In the present study, PID-5 domain scores were computed based on recent American Psychiatric Association guidelines (www.psychiatry.org/File%20Library/Practice/DSM/DSM-5/ThePersonalityInventoryForDSM5FullVersionAdult.pdf) in which each domain represents the average of three facet scales.

Measures Included in the Validation Sample Only

Structured Clinical Interview for DSM-IV Axis II Personality Disorders. (SCID-II; First et al., 1997). The SCID-II is a semistructured interview that assesses the 10 DSM-IV/5 PDs. Each PD criterion is scored using a 0 (i.e., absent), 1 (i.e., subclinical), or 2 (i.e., present) rating. A clinical psychology doctoral student conducted each interview, and subsequently the videotape of the interview was watched and rated by another doctoral student. Intraclass correlations (ICCs) were computed using the interviewer ratings and observer ratings to assess the interrater reliability of the SCID-II ratings; ICCs ranged from .79 (schizotypal PD) to .92 (avoidant and borderline PDs). Alphas for the DSM-IV PD scores ranged from .68 (obsessive-compulsive PD) to .84 (antisocial PD) with a median of .74. In terms of categorical diagnoses, 37.6% of the sample met criteria for at least one DSM-IV PD; the most common diagnoses were avoidant (19.3%) and borderline PD (11%). The mean dimensional counts ranged from 1.78 (histrionic PD) to 4.91 (borderline PD) with a median of 3.32.

DSM-5 clinicians' personality trait rating form. (PTRF; American Psychiatric Association, 2011). This clinician rating form uses a single-item to assess each of the 25 proposed traits subsumed by five trait domains: negative affectivity, detachment, antagonism, disinhibition, and psychoticism. Clinicians provided a 0 (i.e., *very little or not at all descriptive*) to 3 (i.e., *extremely descriptive*) rating based on their perceived presence of a given trait, and facet ratings were summed to provide each domain score. Alphas for the domains ranged from .76 to .89. Double-entry

²To ensure our assumptions were correct, we conducted IRT-based differential item functioning analyses on five randomly selected (original) facet-level PID-5 scales to investigate whether item responses of the groups (i.e., undergraduates vs. MTurk users) were comparable. Specifically, we utilized the Langer-improved Wald's chi-square test (Cai, Thissen, & du Toit, 2011; Langer, 2008) to determine whether item parameters were significantly different between these groups. Eight of the 41 items tested demonstrated DIF, although only one of these eight items had been retained in our abbreviated versions of these scales. That is, out of the 20 items that were retained in the reduced scales, only one (5%) of them demonstrated DIF, consistent with the nominal Type I error rate, as well as simulations mimicking similar data analytic conditions (i.e., sample size, number of groups; see Woods, Cai, & Wang, 2013).

intraclass correlations were computed in order to assess interrater reliabilities for the domains and facets of the clinician-rated PTRF. Double-entry ICCs reflect similarity in shape as well as elevation and scatter (see McCrae, 2008 for a review). For the facet ratings, ICCs ranged from .12 (perseveration) to .83 (impulsivity) with a median of .55. For the single-item domain ratings, ICCs ranged from .50 (negative affectivity) to .82 (disinhibition).

NEO Personality Inventory–Revised. (NEO PI-R; Costa & McCrae, 1992). The NEO PI-R is a 240-item measure of the Five-Factor Model of personality that assesses five domains (i.e., neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness) and 30 facets. Coefficient alphas for the facets ranged from .58 (tendermindedness) to .90 (trust) with a median of .81. Alphas for the domains ranged from .89 to .95.

Crime and Analogous Behavior Scale. (CAB; see Miller & Lynam, 2003). The CAB is a self-report inventory that assesses a variety of externalizing behaviors. A lifetime drug use variety and antisocial behavior variety score was created by giving participants a 1 for every different drug or act endorsed (drug use: eight items; e.g., cocaine; $\alpha = .80$; $M = 2.82$, $SD = 2.20$; antisocial behavior: 10 items; e.g., stealing; $\alpha = .76$; $M = 2.25$, $SD = 2.14$).

Alcohol Use Disorders Identification Test. (AUDIT; Saunders, Aasland, Babor, de la Fuente, & Grant, 1993). The AUDIT is a 10-item self-report measure of problematic alcohol consumption. In the current study, only the total score was used ($\alpha = .90$); scores ranged from 0–36, with a mean of 8.29 ($SD = 8.26$).

Patient-Reported Outcomes Measurement Information System (PROMIS)–Emotional Distress—Anxiety, Depression, Anger—Short Forms (Pilkonis, Choi, Reise, Stover, Riley, & Cella, 2011). The PROMIS scales are brief self-report questionnaires designed to assess the experience of a particular emotion over the past 7 days. The mean anxiety, depression, and anger scores were 22.15 ($SD = 6.50$; $\alpha = .94$), 22.20 ($SD = 9.17$; $\alpha = .97$), and 19.94 ($SD = 6.97$, $\alpha = .90$), respectively.

Data Analysis

IRT analyses were conducted using IRTPRO software program (Cai, Thissen, & du Toit, 2011) to estimate IRT parameters assuming Samejima's (1969) graded response model (GRM; Zickar, 1998). Item parameters were estimated using marginal maximum likelihood (MML) estimation. The GRM provides a model of how person properties (e.g., personality traits) and item properties (e.g., item discrimination) interact to determine individuals' responses to those items. In the GRM, three types of parameters are used to explain responses to items. First, the GRM posits that the latent trait of the individual, θ , partially determines item responses. For example, a personality item is most likely to be endorsed by those with higher levels of θ and less likely to be endorsed by those with moderate to low levels of θ . Second, the GRM includes b_{ik} , or *threshold* parameters. There are $C-1$ thresholds for a given item, where C is the number of response options. Each $b_{i,k}$ parameter indicates the level of the trait needed for a person to choose a particular option instead of the option below it (e.g., the trait level needed to be more likely to choose *strongly agree* instead of *agree*). Items with higher b parameters are more extreme (e.g., reflect more unusual or severe psychopathology) than those with lower b values. Finally, the extent to which that item is able to differentiate or *discriminate* between those with lower and higher

levels of θ is expressed as the discrimination parameter, a . Similar to a factor loading, the a parameter describes the strength of relationship between θ and responses to the item. A high a value indicates that the item is highly sensitive to differences in the relevant trait, whereas a low a value indicates that the item has less value in differentiating respondents on that latent trait. There was an average of .05% missing data across the 220 PID-5 items. These missing values were handled internally through IRTPRO, which simulates the response pattern and fills in isolated missing values based on what would be expected given a participant's theta value.

The unidimensional GRM was first fitted to all PID-5 items for each facet. We then selected the four PID-5 items that had the highest discrimination parameters while also closely monitoring another important aspect of IRT for developing short forms, the *test information function* (TIF). Whereas reliability quantifies measurement precision as an overall property of the test and/or item in a given population or sample, the information function quantifies how measurement precision varies by the trait level, θ , of the respondent. By plotting the degree of information at various levels of θ it is possible to determine whether abbreviated scales are consistent with longer scales in terms of how much information is provided at various levels of the trait. For example, Figure 1 shows an overlay of the TIFs for the original PID-5 grandiosity scale and the reduced scale created by selecting highly discriminating items. As can be seen, both versions have higher information for those with higher levels of grandiosity than those with lower levels. In this way, after selecting items, TIFs from the original and reduced PID-5 scales were compared with ensure that similar measurement properties were obtained in the former; the idea was to maintain the same relative level of information in regard to θ -levels. The TIFs for original and reduced PID-5 scales and IRT parameters are included in the supplemental materials. Additionally, the root mean square error of approximation (RMSEA) of the PID-5 and reduced PID-5 scales was investigated in order to assess model-data fit.

Results

Item Response Theory Investigation of PID-5 Items

The final sets of the four selected Personality Inventory for DSM-5 (PID-5) items for each facet were used to evaluate model-data fit

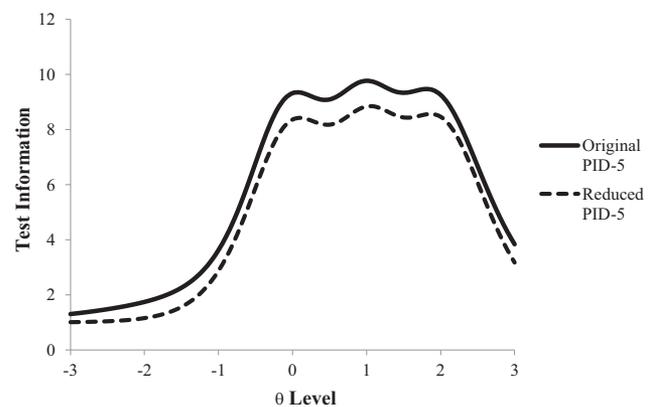


Figure 1. Test information functions (TIFs) for the original PID-5 grandiosity scale and the reduced version.

compared with the fit obtained in the original form. The RMSEAs for these models ranged from .036 to .079 for the reduced item scales and from .034 to .083 for the longer scales (see Table 1), which indicates adequate model-data fit (MacCallum, Browne, & Sugawara, 1996) for all scales, or how well a unidimensional model fits the data. These analyses resulted in four item PID-5 scales for each of the facets that consisted of those PID-5 items that were most representative of the underlying trait measured by original PID-5 measure. An examination of the overlap of the 90% confidence intervals of the RMSEAs revealed some minor differences in model-data fit between the versions, with six PID-5 scales showing better model-data fit for the full length scales and two showing better model-data fit for the reduced scales. However, these differences were relatively small with a mean difference of .003. Most importantly, however, no RMSEA values were out of the range of reasonable model-data fit.

Reliability and Mean Inter-Item Correlations of the PID-5 and Reduced PID-5 Scales

Coefficient alphas and mean interitem correlations (MIC) were calculated for the original and reduced PID-5 scales (see Table 2). Alphas for the PID-5 domain scores ranged from .92 to .96 with a mean of .93 for the derivation sample, and .91 to .96

with a mean of .94 for the validation sample. Alphas for the reduced PID-5 domain scores ranged from .89 to .91 with a mean of .90 for the derivation sample, and .87 to .91 with a mean of .89 for the validation sample. For the PID-5 facet scores, coefficient alphas ranged from .72 to .95 with a mean of .86 in the derivation sample, and .78 to .95 with a mean of .88 for the validation sample. For the reduced PID-5 facet scores, alphas ranged from .74 to .88 with a mean of .83 in the derivation sample, and .67 to .91 with a mean of .83 in the validation sample. In the derivation sample, the mean MICs for scores on the PID-5 and reduced PID-5 were .36 and .41 for the domains and .43 and .54 for the facets, respectively. In the validation sample, the mean MICs for scores on the PID-5 and reduced PID-5 were .41 and .41 for the domains and .53 and .66 for the facets, respectively. Descriptive statistics for the PID-5 and reduced PID-5 scales, both combined and by gender, and coefficient alphas for the scales in each individual sample are included in the supplemental materials.

Factor Analysis of the Original and Reduced PID-5 Scales

In order to compare the factor structure of the original and reduced PID-5 scales, exploratory factor analyses using principal

Table 1
Fit Indices for Long and Reduced PID-5 Scales

	PID-5			Reduced PID-5		
	RMSEA	90% CIs		RMSEA	90% CIs	
		Low	High		Low	High
Negative affectivity						
Anxiousness	.050	.044	.055	.047	.041	.053
Emotional lability	.052	.046	.057	.057	.050	.062
Hostility	.044	.038	.049	.046	.040	.051
Perseveration*	.054	.048	.059	.068	.061	.075
(Lack of) Restricted affectivity	.044	.039	.049	.047	.042	.053
Separation insecurity	.056	.050	.062	.048	.042	.053
Submissiveness	.049	.043	.054	.049	.043	.054
Detachment						
Anhedonia ⁺	.048	.043	.054	.036	.031	.041
Depressivity	.043	.037	.048	.048	.042	.053
Intimacy avoidance	.058	.052	.064	.060	.054	.066
Suspiciousness	.051	.046	.057	.047	.041	.052
Withdrawal	.045	.039	.050	.051	.045	.056
Psychoticism						
Eccentricity*	.044	.038	.049	.061	.055	.067
Perceptual dysregulation	.053	.047	.058	.062	.055	.068
Unusual beliefs and experiences	.054	.048	.059	.062	.056	.068
Antagonism						
Attention seeking*	.054	.048	.059	.071	.064	.078
Callousness*	.034	.028	.039	.050	.044	.056
Deceitfulness	.045	.039	.050	.039	.034	.044
Grandiosity	.048	.042	.053	.048	.042	.054
Manipulativeness	.072	.065	.078	.075	.068	.082
Disinhibition						
Distractibility*	.047	.041	.052	.076	.068	.083
Impulsivity	.057	.051	.063	.058	.051	.063
Irresponsibility	.045	.039	.050	.050	.044	.056
(Lack of) Rigid perfectionism*	.055	.050	.061	.072	.065	.079
Risk taking ⁺	.083	.075	.091	.057	.051	.063

⁺ Indicates the model-data fit was superior for the reduced form compared with the original. * Indicates that model-data fit was superior for the original form compared with the reduced.

Table 2
Coefficient Alphas and MIC for the PID-5 and IRT Selected PID-5 Items

	Alpha				MIC			
	Original		Reduced		Original		Reduced	
	DS	VS	DS	VS	DS	VS	DS	VS
Negative affectivity	.92	.94	.90	.91	.35	.41	.42	.46
Anxiousness	.88	.93	.88	.89	.46	.59	.58	.67
Emotional lability	.88	.92	.86	.89	.52	.61	.60	.66
Hostility	.86	.86	.82	.82	.38	.40	.53	.54
Perseveration	.86	.83	.81	.78	.40	.35	.51	.47
(Lack of) Restricted affectivity	.81	.79	.76	.74	.38	.35	.44	.42
Separation insecurity	.82	.87	.78	.85	.39	.49	.46	.59
Submissiveness	.78	.78	.78	.78	.47	.47	.47	.47
Detachment	.92	.95	.89	.90	.34	.42	.42	.42
Anhedonia	.81	.92	.82	.91	.35	.60	.54	.71
Depressivity	.94	.95	.88	.88	.52	.57	.64	.65
Intimacy avoidance	.79	.86	.82	.86	.39	.50	.54	.61
Suspiciousness	.72	.82	.74	.74	.27	.41	.42	.43
Withdrawal	.91	.94	.85	.85	.51	.60	.58	.60
Psychoticism	.96	.96	.90	.89	.39	.41	.44	.40
Eccentricity	.95	.92	.88	.90	.60	.62	.66	.68
Perceptual dysregulation	.89	.86	.82	.72	.40	.36	.54	.44
Unusual beliefs and experiences	.85	.86	.79	.77	.42	.43	.48	.45
Antagonism	.92	.91	.91	.87	.36	.33	.45	.37
Attention seeking	.89	.90	.85	.88	.49	.52	.59	.64
Callousness	.90	.87	.85	.79	.42	.34	.59	.51
Deceitfulness	.87	.89	.81	.82	.40	.46	.53	.54
Grandiosity	.82	.80	.84	.76	.44	.40	.56	.45
Manipulativeness	.82	.81	.80	.76	.47	.45	.50	.44
Disinhibition	.92	.94	.89	.90	.34	.40	.39	.42
Distractibility	.90	.92	.86	.90	.49	.56	.60	.69
Impulsivity	.83	.90	.85	.86	.43	.59	.59	.61
Irresponsibility	.79	.80	.77	.67	.35	.37	.46	.34
(Lack of) Rigid perfectionism	.88	.91	.81	.84	.42	.51	.52	.56
Risk taking	.85	.92	.82	.88	.28	.46	.54	.65
Mean	.88	.89	.84	.84	.42	.47	.52	.54

Note. MIC = Mean inter-item correlation; DS = Derivation sample; VS = Validation sample.

axis factoring with an equamax rotation were performed in a dataset combining the validation and derivation samples. Given that the Section III model is presented as a five-factor model and the bulk of the research investigating the PID-5 has investigated a five factor higher order structure, the results from analyses in which five factors were extracted are presented (see Table 3). Congruence coefficients (Tucker, 1951) for the original and reduced scale factor loadings for the five factors were .99, .99, .99, .98, and .93. Given the hierarchical nature of personality structure in which a range of meaningful factors can be extracted (e.g., Krueger & Markon, 2014; Markon, Krueger, & Watson, 2005), the factor analyses were also performed in which two, three, and four factors were extracted. The congruence coefficients for the original and reduced PID-5 scales were .99, .99, .98, and .98 for the four factor solution, .99, .99, and .99 for the three factor solution, and .98 and .97 for the two factor solution.

Convergent Validity

We examined the convergent validity correlations manifested by the PID-5 domain and facet scores with the scores on the domains and facets from the reduced set of PID-5 items in the derivation and validation samples (see Table 4). For all mean correlation

calculations, individual correlations were first transformed using the Fisher's-Z transformation before being averaged and transformed back into Pearson correlations. At the domain level in the derivation sample, convergent correlations ranged from .96 to .98 with a mean of .97; facet level correlations ranged from .89 to 1.0 with a mean of .94. The overall mean convergent correlation for the reduced PID-5 scores with the PID-5 scores was .95 in the derivation sample. In the validation sample, convergent validity correlations for the domains ranged from .96 to .98 with a mean of .98; facet level convergent correlations ranged from .82 to 1.0 with a mean of .95. The overall mean convergent correlation for the reduced PID-5 scores with the PID-5 scores was .96 in the validation sample.

Discriminant Validity

Next the discriminant validity correlations manifested by scores from the reduced PID-5 scores were compared with those generated by the full-length PID-5 scales (see Table 5). Discriminant validity correlations are within form correlations, such that each PID-5 domain was correlated with the other four domains from within the same scale (i.e., full or reduced). In the derivation sample, PID-5 discriminant validity correlations ranged from .41

Table 3
Factor Loadings for Original and Reduced PID-5 Scales

	Factor 1		Factor 2		Factor 3		Factor 4		Factor 5	
	O	R	O	R	O	R	O	R	O	R
Negative affectivity										
Anxiousness	.74	.74	.32	.23	.01	.01	.02	.11	.26	.23
Emotional lability	.73	.65	.11	.11	.10	.07	.25	.29	.29	.40
Hostility	.39	.41	.38	.33	.47	.29	.26	.29	.21	.23
Perseveration	.53	.56	.28	.26	.19	.17	.26	.33	.50	.29
(Lack of) Restricted affectivity	-.13	-.12	.55	.54	.32	.33	.14	.13	.32	.04
Separation insecurity	.64	.66	.00	-.06	.24	.18	.20	.19	.17	.16
Submissiveness	.43	.47	.08	.08	.13	.15	.06	.14	.19	.05
Detachment										
Anhedonia	.46	.43	.71	.60	.04	.08	.11	.21	.15	.31
Depressivity	.56	.40	.61	.56	.07	.08	.24	.21	.22	.40
Intimacy avoidance	.06	.07	.54	.54	.16	.13	.11	.09	.27	.32
Suspiciousness	.38	.38	.41	.39	.28	.34	.08	.19	.35	.41
Withdrawal	.17	.18	.80	.78	.09	.03	.02	.11	.32	.23
Psychoticism										
Eccentricity	.21	.22	.25	.27	.10	.14	.40	.37	.57	.38
Cog. and perceptual dysreg.	.32	.13	.33	.26	.26	.35	.40	.31	.59	.67
Unusual beliefs and experiences	.15	.14	.22	.16	.31	.24	.28	.28	.63	.77
Antagonism										
Attention seeking	.21	.24	-.20	-.22	.60	.56	.38	.29	.18	.10
Callousness	.04	.03	.50	.48	.60	.53	.33	.28	.20	.25
Deceitfulness	.20	.19	.33	.30	.61	.60	.41	.39	.18	.26
Grandiosity	.05	.09	.13	.24	.70	.70	.15	.15	.31	.29
Manipulativeness	.08	.07	.13	.16	.69	.66	.32	.32	.17	.21
Disinhibition										
Distractibility	.46	.47	.33	.20	.03	.04	.52	.52	.29	.17
Impulsivity	.22	.15	.10	.07	.21	.23	.74	.74	.19	.18
Irresponsibility	.24	.21	.43	.37	.34	.37	.55	.48	.15	.34
(Lack of) Rigid perfectionism	.35	.37	.10	.16	.35	.31	-.14	-.02	.47	.29
Risk taking	-.14	.01	-.10	.07	.26	.39	.58	.59	.07	.22
Congruence coefficients	.99		.99		.99		.98		.93	

Note. O = original, full length PID-5 scales; R = reduced item PID-5 scales.

to .70 with a mean of .56, whereas the reduced length scales produced discriminant validity correlations that ranged from .41 to .70 with a mean of .57. The similarity of the correlation profiles generated across scores on the PID-5 and the reduced PID-5 scales was tested using a double entry ICC that measures absolute agreement in the correlation profiles (see McCrae, 2008). The ICC across correlational profiles for the PID-5 scores with the reduced PID-5 scores was .98 in the derivation sample. In the validation sample, PID-5 discriminant validity correlations ranged from .26 to .59 with a mean of .46, whereas reduced length PID-5 scales manifested discriminant validity correlations that ranged from .27 to .58 with a mean of .45. The ICC across correlational profiles for the PID-5 scores with the reduced PID-5 scores was .95 in the validation sample.

Criterion Validity (Validation Sample Only)

Relations between the PID-5 domains and NEO PI-R domains. The correlations between the PID-5 domains, as measured by the full measure and the set of reduced items, and self-reported NEO PI-R domains were tested next (see Table 6).³ In general, the domain level scores from both the full and reduced length PID-5 scales manifested reasonable convergent and discriminant validity correlations with their counterpart NEO PI-R domains, aside from psychoticism, with which both the PID-5 and

reduced PID-5 demonstrated null relations with openness ($r = .06$ and $.00$, respectively). For both the full length and reduced PID-5 scales, negative affectivity manifested its largest correlation with neuroticism ($r = .83$ and $.81$, respectively), Detachment manifested its largest correlation with extraversion ($r = -.75$ and $-.70$, respectively), Antagonism manifested its largest correlation with agreeableness ($r = -.69$ and $-.69$, respectively), and disinhibition manifested its largest correlation with conscientiousness ($r = -.76$ and $-.74$, respectively). The ICC across their correlational profiles was .98.

Convergent validity with interviewer rated DSM-5 Section III trait scores. The convergent validity correlations manifested by scores on both the full and reduced length PID-5 scales were examined in relation to interviewer rated DSM-5 trait scores (see Table 7). At the domain level, convergent correlations ranged from .50 to .72 with a mean of .63 for the PID-5 and .50 to .70 with a mean of .61 for the reduced length PID-5 scales. At the facet level,

³ The criterion relations with the original PID-5 measure were previously reported in Few et al., 2013. They are reproduced here for sake of comparison with the reduced PID-5. Correlations between the PID-5 domains and criterion variables are at times different than reported in the original manuscript, as the present manuscript utilizes a different domain scoring technique for the PID-5, consistent with current APA recommendations.

Table 4
 Convergent Correlations Among the Reduced Length PID-5 Scales With the Original PID-5

	Reduced PID-5	
	DS	VS
Negative affectivity	.98	.98
Anxiousness	.95	.97
Emotional lability	.95	.97
Hostility	.91	.91
Perseveration	.92	.90
(Lack of) Restricted affectivity	.93	.94
Separation insecurity	.95	.97
Submissiveness	—	—
Detachment	.96	.98
Anhedonia	.91	.96
Depressivity	.93	.94
Intimacy avoidance	.95	.98
Suspiciousness	.89	.95
Withdrawal	.94	.95
Psychoticism	.96	.96
Eccentricity	.96	.95
Perceptual dysregulation	.89	.82
Unusual beliefs and experiences	.93	.93
Antagonism	.98	.97
Attention seeking	.95	.94
Callousness	.93	.89
Deceitfulness	.93	.93
Grandiosity	.95	.94
Manipulativeness	.97	.97
Disinhibition	.97	.98
Distractibility	.94	.96
Impulsivity	.96	.97
Irresponsibility	.94	.94
(Lack of) Rigid perfectionism	.92	.95
Risk taking	.82	.90
Overall mean r	.95	.96
Domain level mean r	.97	.98
Facet level mean r	.94	.95

Note. In the derivation sample, correlations are significant at $p \leq .01$ at .075, $p \leq .001$ at .09. In the validation sample, correlations are significant at $p \leq .01$ at .25, $p \leq .001$ at .30. DS = derivation sample; VS = validation sample.

convergent correlations ranged from .32 to .69 with a mean of .52 for the full length PID-5, and .18 to .68 with a mean of .51 for the reduced length PID-5 scales. The overall mean convergent correlations between the interviewer rated *DSM-5* trait scores and the full and reduced length PID-5 scales were .54 and .53, respectively, and the ICC across their correlational profiles was .95.

Relations among PID-5 domains and interviewer rated *DSM-5* Section II PDs. Next, the correlations between the scores on the full and reduced length PID-5 scales with the *DSM-5* Section II PD scores were examined (see Table 8). Across both the full and reduced length PID-5 scales, negative affectivity demonstrated significant relations with seven of the 10 PDs, all except for schizoid, histrionic, and narcissistic. Across both the full and reduced length PID-5 scales, detachment demonstrated significant relations with seven of the 10 PDs, all except for histrionic, narcissistic, and obsessive–compulsive. Across both the full and reduced length PID-5 scales, psychoticism demonstrated significant relations with eight of the 10 PDs. Both the full and reduced PID-5 scales demonstrated a null relation between psychoticism

and histrionic PD, whereas the full length PID-5 psychoticism scale demonstrated a significant relation with obsessive–compulsive PD ($r = .25$) whereas the correlation manifested by the reduced PID-5 scale did not reach significance ($r = .19$). Across both the full and reduced length PID-5, antagonism demonstrated significant relations with five of the 10 PDs, including paranoid, schizotypal, antisocial, borderline, and narcissistic PD. Across both the full and reduced length PID-5, disinhibition demonstrated significant relations with eight of 10 PDs. Both versions of the PID-5 disinhibition scale demonstrated a null relation with schizoid PD, whereas the reduced length PID-5 manifested a significant relation with obsessive–compulsive PD ($r = .25$) but the relation manifested by the full length PID-5 scale was not significant ($r = .24$). The intraclass correlation across their correlational profiles was .99.

Relations between PID-5 domains and externalizing behaviors and internalizing symptoms. The correlations between the PID-5 domains, as measured by scores on the full and reduced length PID-5 scales, and a range of externalizing and internalizing outcomes were examined (see Table 9). Across the full and reduced length PID-5 scales, negative affectivity demonstrated null relations with the three externalizing outcomes and significant positive relations with all internalizing outcomes. Across both the full and reduced length PID-5 scales, detachment demonstrated a significant relation with drug use, anxiety, anger, and depression. Psychoticism, as measured by the full and reduced length PID-5 scales, demonstrated significant relations with all externalizing and internalizing outcomes. Similarly, antagonism demonstrated significant relations with all externalizing and internalizing outcomes except for drug use across both the full and reduced length scales. Likewise, both the full and reduced length PID-5 disinhibition scales produced significant relations with all externalizing and internalizing outcomes. The ICC across their correlational profiles was .99.

Overall Similarity

The similarity in the correlation profiles for the full and reduced PID-5 scales was computed across all correlations provided in Tables 4 through 9. The overall similarity of the correlation profiles across the previous analyses was very high, demonstrating that the PID-5 scales using the full and reduced item sets generated a pattern of nearly identical correlations with important external criteria ($r_{ICC} = .995$).

Discussion

The inclusion of the alternative model of PDs in Section III of the *DSM-5* represents an important step toward the potential adoption of a dimensional model of personality pathology that is, at least partly, predicated on the use of a trait personality model as the primary approach in the next iterations of the *DSM*. The future adoption of this approach, however, would be facilitated by the accumulation of a substantial research literature supporting such a transition. The PID-5 is the only current measure created explicitly to assess this model and has been the subject of a growing number of empirical studies reporting on its reliability and validity. In the hopes of encouraging further research investigating the trait portion of the alternative diagnostic model, the present study used IRT

Table 5

Discriminant Validity Correlations Among PID-5 Domains in the Derivation and Validation Samples

	Derivation sample									
	Negative affectivity		Detachment		Psychoticism		Antagonism		Disinhibition	
	O	R	O	R	O	R	O	R	O	R
Negative affectivity	—	—	.47	.47	.59	.55	.41	.42	.58	.59
Detachment	.43	.47	—	—	.58	.58	.44	.48	.55	.54
Psychoticism	.59	.58	.58	.61	—	—	.63	.64	.70	.70
Antagonism	.41	.41	.44	.46	.63	.63	—	—	.62	.62
Disinhibition	.58	.57	.55	.56	.70	.69	.62	.64	—	—

	Validation sample									
	Negative affectivity		Detachment		Psychoticism		Antagonism		Disinhibition	
	O	R	O	R	O	R	O	R	O	R
Negative affectivity	—	—	.43	.43	.49	.41	.35	.33	.58	.59
Detachment	.43	.45	—	—	.46	.43	.26	.28	.40	.40
Psychoticism	.49	.50	.46	.48	—	—	.46	.41	.59	.58
Antagonism	.35	.37	.26	.27	.46	.43	—	—	.51	.53
Disinhibition	.58	.58	.40	.39	.59	.53	.51	.53	—	—

Note. In the derivation sample, correlations are significant at $p \leq .01$ at .075, $p \leq .001$ at .09. In the validation sample, correlations are significant at $p \leq .01$ at .25, $p \leq .001$ at .30. O = original, full length PID-5 scales; R = reduced item PID-5 scales.

analyses to test whether a reduced set of PID-5 items can be identified that can measure these traits in a reliable, valid, and efficient manner.

Internal Consistency

Scores from both the full and reduced PID-5 manifested good internal consistency, as demonstrated by both coefficient alphas and mean interitem correlations. Scores from the PID-5 and reduced PID-5 demonstrated mean coefficient alphas of .93 and .90 for the domains and .86 and .83 for the facets, respectively. These findings are consistent with previous data on the PID-5, as in the initial development study coefficient alpha for the PID-5 facets ranged from .72 to .96 with a median of .86 (Krueger et al., 2012). Given that item number is an important component in the calculation of coefficient alpha, it is notable that scores from the reduced PID-5 demonstrated only a small reduction in coefficient alpha despite representing a nearly 55% reduction in the total number of items, from 220 to 100. Mean interitem correlations were also calculated in order to compare internal consistency in a way that does not depend on the number of items in each scale. Mean interitem correlations should generally fall between .15 and .50 (Clark & Watson, 1995), and broader constructs, such as the PID-5 domains, should manifest lower mean interitem correla-

tions, whereas more narrow constructs, such as the PID-5 facets, should demonstrate higher values. In the present study, across both samples, the average mean interitem correlations for the PID-5 and reduced PID-5 were .38 and .42 for the domains and .46 and .55 for the facets, respectively. The coefficient alpha and mean interitem correlation data from the current study suggest that both sets of scores demonstrate strong internal consistency, despite the reduced set of PID-5 items representing a substantial decrease in total item number.

Internal Structure

The Section III PD pathological trait model is a hierarchical structure, in which five higher order domains subsume 25 traits. As such, it was important to test the similarity of the internal structure of the original and reduced PID-5 scales. Congruence coefficients for the five factor loadings were .99, .99, .99, .98, and .93. Congruence coefficients above .95 indicates that the factors can be considered equal (Lorenzo-Seva & ten Berge, 2006), as such the results suggest that four of the five factors are identical across the original and reduced PID-5 scales, with the fifth factor demonstrating strong similarity. Given the possibility of alternative structures of the PID-5, two-, three-, and four-factor structure solutions were also investigated, and congruence coefficients across all three

Table 6

Relations Among PID-5 Domains and NEO PI-R Domains in the Validation Sample

	Validation sample									
	Negative affectivity		Detachment		Psychoticism		Antagonism		Disinhibition	
	O	R	O	R	O	R	O	R	O	R
Neuroticism	.83	.81	.55	.53	.45	.36	.30	.31	.60	.60
Extraversion	-.29	-.28	-.75	-.70	-.25	-.24	-.03	-.06	-.15	-.17
Openness	-.10	-.12	-.36	-.35	.06	.00	-.15	-.18	.02	.04
Agreeableness	-.30	-.30	-.26	-.23	-.30	-.28	-.69	-.69	-.40	-.40
Conscientiousness	-.44	-.42	-.32	-.28	-.34	-.31	-.26	-.30	-.76	-.74

Note. Correlations are significant at $p \leq .01$ at .25, $p \leq .001$ at .30. O = original, full length PID-5 scales; R = reduced item PID-5 scales.

Table 7
Convergent Correlations Between the Original and Reduced Set of PID-5 Items With Clinician-Rated Traits in the Validation Sample

	Interview ratings	
	O	R
Negative affectivity	.65	.63
Anxiousness	.47	.49
Emotional lability	.55	.52
Hostility	.62	.62
Perseveration	.32	.18
(Lack of) Restricted affectivity	.36	.43
Separation insecurity	.60	.61
Submissiveness	.40	.40
Detachment	.72	.70
Anhedonia	.66	.67
Depressivity	.62	.60
Intimacy avoidance	.39	.40
Suspiciousness	.67	.68
Withdrawal	.68	.67
Psychoticism	.53	.51
Eccentricity	.35	.32
Perceptual dysregulation	.45	.43
Unusual beliefs and experiences	.60	.55
Antagonism	.50	.51
Attention seeking	.52	.45
Callousness	.53	.51
Deceitfulness	.43	.49
Grandiosity	.36	.33
Manipulativeness	.41	.39
Disinhibition	.69	.68
Distractibility	.48	.47
Impulsivity	.64	.67
Irresponsibility	.61	.59
(Lack of) Rigid perfectionism	.39	.36
Risk taking	.55	.57
Overall mean r	.54	.53
Domain level mean r	.63	.61
Facet level mean r	.52	.51

Note. Correlations are significant at $p \leq .01$ at .25, $p \leq .001$ at .30.
 O = original, full length PID-5 scales; R = reduced item PID-5 scales.

factor solutions ranged from .97 to .99, indicating that the factor loadings across the original and reduced PID-5 scales were essentially equal.

Convergent Validity

A growing empirical literature has begun to accumulate that provides initial support for the validity of the PID-5 measure (see Krueger & Markon, 2014, for a review). As such, investigating the convergent validity of scores of the reduced set of items in comparison with the original measure was important. Scores from the reduced set of PID-5 items demonstrated strong convergent validity with scores from the PID-5, with the mean convergent validity correlation being .97 at the domain level and .94 at the facet level. Notably, these results were replicated in a second independent sample, with a mean convergent validity correlation of .98 at the domain level and .95 at the facet level. Ultimately, the reduced PID-5 scales were almost perfectly correlated with the longer scales from which they were derived.

Discriminant Validity

Discriminant validity is another important component of measure validation (Campbell & Fiske, 1959). In the initial sample, the mean discriminant validity correlations for the PID-5 and reduced PID-5 domains ranged from .41 to .70 with a mean of .56, and .41 to .70 with a mean of .57, respectively. These discriminant validity correlations are higher than is typically found for personality domains. For instance, in a study comparing four different measures of the FFM, median discriminant validity correlations at the domain level ranged from .16 to .20 (Maples et al., 2014). However, discriminant validity relies on an understanding of how strongly constructs should or should not relate to each other, and it may be that these pathological trait domains should relate more strongly, compared with relations found between normal trait domains, due to the strong emphasis on impairment associated with each domain. In this regard, substantial correlations between measures of different forms of psychopathology are common (e.g., Lahey et al., 2012; Markon, 2010). Additionally, the current findings are consistent with previous research in which the mean discriminant validity among the PID-5 domains was .46 (Quilty et al., 2013). Most important for the purposes of the present study, discriminant validity correlations were highly similar across the scales created using the full and reduced item sets. The intraclass correlation of their correlation profiles was .97, suggesting that the scales built from the reduced set of PID-5 items measured these constructs in a manner that is nearly identical to the original full length PID-5.

Criterion Validity

Criterion validity was investigated in order to test if the reduced set of PID-5 items could successfully maintain the nomological network of the full item scales. Given that the proposed DSM-5 pathological trait model is purported to represent maladaptive variants of the FFM (APA, 2013), the convergent validity between the PID-5 domains as assessed via both versions of this measure with the FFM domains was investigated. Consistent with previous research (e.g., Quilty et al., 2013), domain level scores from both the full and reduced length PID-5 scales demonstrated reasonable convergent validity, as all demonstrated their strongest absolute correlation with their counterpart NEO PI-R domains, aside from the relation between psychoticism and openness, which demonstrated null relations ($r = .06$ and $.00$, respectively). The lack of convergence between psychoticism, as assessed via scores on the PID-5 openness is not entirely surprising, as in the development of the PID-5 the psychoticism factor was modeled to more closely resemble the psychoticism domain from Harkness' Personality Psychopathology 5 (PSY-5; Harkness et al., 1995). The relation between openness and psychoticism is the subject of substantial ongoing investigation (e.g., Chmielewski, Bagby, Markon, Ring, & Ryder, 2014; Edmundson, Lynam, Miller, Gore, & Widiger, 2011; Gore & Widiger, 2013; Suzuki et al., 2015). With regard to the focus of the present investigation, the relations between the full and reduced length PID-5 domains and the FFM domains were very similar, as demonstrated by an intraclass correlation of .95 across their correlations with the FFM domains.

Scores from both sets of PID-5 scales also demonstrated good convergent validity with interviewer rated DSM-5 trait scores, as the overall mean convergent correlation between the interviewer

Table 8
PID-5 Domains and Clinician Rated DSM-IV/5 PDs in the Validation Sample

DSM-IV/5 PDs	Negative affectivity		Detachment		Psychoticism		Antagonism		Disinhibition	
	O	R	O	R	O	R	O	R	O	R
Paranoid	.40	.39	.32	.31	.38	.35	.46	.43	.38	.40
Schizoid	.20	.20	.64	.64	.29	.30	.22	.21	.16	.17
Schizotypal	.28	.29	.37	.38	.60	.59	.30	.27	.28	.29
Antisocial	.29	.27	.25	.26	.41	.39	.49	.51	.62	.61
Borderline	.58	.56	.34	.35	.46	.43	.40	.40	.61	.64
Histrionic	.12	.10	-.15	-.12	.12	.13	.16	.17	.29	.27
Narcissistic	.22	.21	.07	.08	.27	.27	.56	.53	.35	.37
Avoidant	.51	.51	.55	.54	.43	.39	.12	.11	.37	.39
Dependent	.60	.60	.41	.39	.43	.39	.22	.24	.49	.50
OCPD	.25	.25	.22	.25	.25	.19	.21	.17	.24	.25

Note. Correlations are significant at $p \leq .01$ at .25, $p \leq .001$ at .30. OCPD = Obsessive-Compulsive PD; O = original, full length PID-5 scales; R = reduced item PID-5 scales.

rated PID-5 scores and the full and reduced length PID-5 scales was .54 and .53, respectively. The intraclass correlation between their correlation profiles was .95. Previous research on the traditional, categorical models of PDs that compares clinician assessment of PD with self-report data has suggested poor agreement between these methods (e.g., Wilberg, Dammen, & Friis, 2000), so the convergence between self-report PID-5 scores, using both the full and reduced set of items, and interviewer rated pathological traits is promising.

Coverage is another important component of the clinical utility of the dimensional PD model (e.g., Verheul, 2006), including coverage of the existing categorical *DSM-IV/5* PD constructs. As such, the relations between scores on both the PID-5 and the reduced length PID-5 scales and the existing diagnoses, as assessed via semistructured interview, were investigated. Across PID-5 scores built from the full and reduced item sets, the domains demonstrated significant relations with between five and eight of the *DSM-5* Section II PD scores. As with the previous external criteria, the pattern of relations manifested by the two sets of PID-5 items was very similar ($r_{ICC} = .99$).

Given the clinical nature of the traits included in the *DSM-5* trait model, measures of these traits should manifest expected correla-

tions with clinically relevant outcomes; therefore, the relations between scores on the full and reduced length PID-5 scales with a range of externalizing and internalizing outcomes were investigated. As expected, the domains from both the PID-5 and reduced length PID-5 scales demonstrated variable associations with externalizing behaviors and internalizing symptoms, largely in a manner consistent with the literature on their NEO PI-R counterparts. For instance, previous meta-analytic work indicated the existence of a strong relation between neuroticism and symptoms of anxiety and depression (Kotov, Gamez, Schmidt, & Watson, 2010), and in the present study scores on both the full and abbreviated PID-5 negative affectivity scales demonstrated their strongest relation with anxiety and depression. Similarly, the empirical literature surrounding agreeableness and conscientiousness suggests robust relations between these domains and externalizing outcomes (e.g., Miller & Lynam, 2001), and scores on both the full and reduced PID-5 antagonism and disinhibition domains demonstrated significant relations with the externalizing outcomes. Overall, the full and reduced PID-5 scales manifested very similar sets of correlations with the externalizing behaviors and internalizing symptoms ($r_{ICC} = .99$).

Table 9
PID-5 Domains in Relation to Externalizing Outcomes and Internalizing Symptoms in the Validation Sample

	Negative affectivity		Detachment		Psychoticism		Antagonism		Disinhibition	
	O	R	O	R	O	R	O	R	O	R
Externalizing										
Alcohol	.21	.19	.11	.10	.28	.26	.33	.33	.35	.34
Drug count	.24	.24	.30	.30	.39	.33	.23	.21	.46	.45
ASB	.11	.12	.15	.15	.33	.32	.30	.30	.29	.29
Internalizing										
Anxiety	.70	.71	.47	.46	.50	.43	.31	.26	.44	.44
Depression	.68	.69	.59	.59	.48	.43	.30	.28	.45	.45
Anger	.54	.56	.33	.31	.43	.40	.35	.34	.48	.50

Note. Correlations are significant at $p \leq .01$ at .25, $p \leq .001$ at .30; ASB = Antisocial behavior; O = original, full length PID-5 scales; R = reduced item PID-5 scales.

Limitations and Conclusions

One limitation of the current study is that the initial development sample was composed of approximately half undergraduates, with the other half consisting of community adults collected from MTurk. Given that the PID-5 is intended for use in clinical samples, one concern would be that the IRT development of the reduced PID-5 form in a nonclinical sample may not contain enough respondents representing extreme levels of the trait, thus may result in precise parameter estimates only for a limited range of the traits. However, in order to mitigate this concern the results were cross-validated in a community sample of individuals, all of whom were currently receiving psychological or psychiatric treatment. The PID-5 scales derived from the full and reduced items sets performed very similarly in this sample, as they had in the original development sample. As is necessary with all empirical endeavors, the current findings require replication using a variety of sampling approaches and in relation to a host of relevant criteria. While IRT parameters are theoretically subpopulation invariant, it is feasible that other samples might suggest other items as the best four possible items within the facet scales.

We believe that the inclusion of a dimensional alternative model of PDs in Section III of the *DSM-5* represents a significant advancement in the classification, diagnosis, and treatment of personality pathology. The likelihood that this model or some variant of it will become the sole approach in future editions of the *DSM* will be facilitated by the development of a large and robust empirical literature that documents its performance in a wide array of samples and in relation to a wide array of relevant outcomes. The present study suggests that it is possible to identify and use a significantly reduced set of PID-5 items to score the *DSM-5* domains and facets that result in nearly identical reliability and validity and similar nomological networks with the PID-5. The availability of a much shorter version of the PID-5 that maintains the central features of this promising measure may enable researchers to embed it in a greater number of studies of psychiatric disorders and functioning, which would serve to expedite the process by which a substantial empirical body of research is accrued on the *DSM-5* Section III model of personality disorders.

References

- American Psychiatric Association. (2011). *DSM-5 clinicians' personality trait rating form*. Retrieved from <http://www.dsm5.org/ProposedRevisions/Pages/PersonalityandPersonalityDisorders.aspx>
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Association.
- Anderson, J. L., Sellbom, M., Bagby, R. M., Quilty, L. C., Veltri, C. O., Markon, K. E., & Krueger, R. F. (2013). On the convergence between PSY-5 domains and PID-5 domains and facets: Implications for assessment of *DSM-5* personality traits. *Assessment, 20*, 286–294. <http://dx.doi.org/10.1177/1073191112471141>
- Ashton, M. C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review, 11*, 150–166. <http://dx.doi.org/10.1177/1088868306294907>
- Ashton, M. C., Lee, K., de Vries, R. E., Hendrickse, J., & Born, M. P. (2012). The maladaptive personality traits of the Personality Inventory for *DSM-5* (PID-5) in relation to the HEXACO personality factors and schizotypy/dissociation. *Journal of Personality Disorders, 26*, 641–659. <http://dx.doi.org/10.1521/pedi.2012.26.5.641>
- Behrend, T. S., Sharek, D. J., Meade, A. W., & Wiebe, E. N. (2011). The viability of crowdsourcing for survey research. *Behavior Research Methods, 43*, 800–813. <http://dx.doi.org/10.3758/s13428-011-0081-0>
- Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A. M., & Kaemmer, B. (1989). *Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration and scoring*. Minneapolis, MN: University of Minnesota Press.
- Butcher, J. N., & Hostetler, K. (1990). Abbreviating MMPI item administration: What can be learned from the MMPI for the MMPI—2? *Psychological Assessment: Journal of Consulting and Clinical Psychology, 2*, 12–21. <http://dx.doi.org/10.1037/1040-3590.2.1.12>
- Cai, L., Thissen, D., & du Toit, S. H. C. (2011). *IRTPRO 2.1 for Windows* [Software program]. Chicago, IL: Scientific Software International.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81–105. <http://dx.doi.org/10.1037/h0046016>
- Carter, N. T., Kotrba, L. M., & Lake, C. J. (2014). Null results in assessing survey score comparability: Illustrating measurement invariance using item response theory. *Journal of Business and Psychology, 29*, 205–220. <http://dx.doi.org/10.1007/s10869-012-9283-4>
- Chmielewski, M., Bagby, R. M., Markon, K., Ring, A. J., & Ryder, A. G. (2014). Openness to experience, intellect, schizotypal personality disorder, and psychoticism: Resolving the controversy. *Journal of Personality Disorders, 28*, 483–499. http://dx.doi.org/10.1521/pedi_2014_28_128
- Clark, L. A. (1993). *SNAP, Schedule for Nonadaptive and Adaptive Personality: Manual for administration, scoring, and interpretation*. Minneapolis, MN: University of Minnesota Press.
- Clark, L. A. (2007). Assessment and diagnosis of personality disorder: Perennial issues and an emerging reconceptualization. *Annual Review of Psychology, 58*, 227–257. <http://dx.doi.org/10.1146/annurev.psych.57.102904.190200>
- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*, 309–319. <http://dx.doi.org/10.1037/1040-3590.7.3.309>
- Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.
- Edelen, M. O., & Reeve, B. B. (2007). Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. *Quality of Life Research, 16*, 5–18. Retrieved from <http://dx.doi.org/10.1007/s11136-007-9198-0>
- Edmundson, M., Lynam, D. R., Miller, J. D., Gore, W. L., & Widiger, T. A. (2011). A five-factor measure of schizotypal personality traits. *Assessment, 18*, 321–334. <http://dx.doi.org/10.1177/1073191111408228>
- Few, L. R., Miller, J. D., Rothbaum, A. O., 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*, 1057–1069. <http://dx.doi.org/10.1037/a0034878>
- First, M. B., Gibbon, M., Spitzer, R. L., Williams, J. B. W., & Benjamin, L. S. (1997). *Structured Clinical Interview for DSM-IV Axis I disorders (SCID-I)*. Washington, DC: American Psychiatric Press.
- Frances, A. (1993). Dimensional diagnosis of personality—Not whether, but when and which. *Psychological Inquiry, 4*, 110–111. http://dx.doi.org/10.1207/s15327965pli0402_7
- Goldberg, L. R. (1999). A broad-bandwidth, public-domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. J. Deary, F. De Fruyt, and F. Ostendorf (Eds.), *Personality Psychology in Europe* (pp. 7–28). Tilburg: University Press.

- Gore, W. L., & Widiger, T. A. (2013). The *DSM-5* dimensional trait model and five-factor models of general personality. *Journal of Abnormal Psychology, 122*, 816–821. <http://dx.doi.org/10.1037/a0032822>
- Griffin, S. A., & Samuel, D. B. (2014). A closer look at the lower-order structure of the Personality Inventory for *DSM-5*: Comparison with the Five-Factor Model. *Personality Disorders, 5*, 406–412. <http://dx.doi.org/10.1037/per0000074>
- Harkness, A. R., McNulty, J. L., & Ben-Porath, Y. S. (1995). The Personality Psychopathology Five (PSY-5): Constructs and MMPI-2 scales. *Psychological Assessment, 7*, 104–114. <http://dx.doi.org/10.1037/1040-3590.7.1.104>
- Hopwood, C. J., Thomas, K. M., Markon, K. E., Wright, A. G., & Krueger, R. F. (2012). *DSM-5* personality traits and *DSM-IV* personality disorders. *Journal of Abnormal Psychology, 121*, 424–432. <http://dx.doi.org/10.1037/a0026656>
- Hopwood, C. J., Wright, A. G., Krueger, R. F., Schade, N., Markon, K. E., & Morey, L. C. (2013). *DSM-5* pathological personality traits and the personality assessment inventory. *Assessment, 20*, 269–285. <http://dx.doi.org/10.1177/1073191113486286>
- Johnson, D. R., & Borden, L. A. (2012). Participants at your fingertips: Using Amazon's Mechanical Turk to increase student-faculty collaborative research. *Teaching of Psychology, 39*, 245–251. <http://dx.doi.org/10.1177/0098628312456615>
- Kotov, R., Gamez, W., Schmidt, F., & Watson, D. (2010). Linking “big” personality traits to anxiety, depressive, and substance use disorders: A meta-analysis. *Psychological Bulletin, 136*, 768–821. <http://dx.doi.org/10.1037/a0020327>
- Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., & Skodol, A. E. (2012). Initial construction of a maladaptive personality trait model and inventory for *DSM-5*. *Psychological Medicine, 42*, 1879–1890. Retrieved from <http://dx.doi.org/10.1017/S0033291711002674>
- Krueger, R. F., & Markon, K. E. (2014). The role of the *DSM-5* personality trait model in moving toward a quantitative and empirically based approach to classifying personality and psychopathology. *Annual Review of Clinical Psychology, 10*, 477–501. <http://dx.doi.org/10.1146/annurev-clinpsy-032813-153732>
- Krueger, R. F., & Tackett, J. L. (2003). Personality and psychopathology: Working toward the bigger picture. *Journal of Personality Disorders, 17*, 109–128. <http://dx.doi.org/10.1521/pedi.17.2.109.23986>
- Lahey, B. B., Applegate, B., Hakes, J. K., Zald, D. H., Hariri, A. R., & Rathouz, P. J. (2012). Is there a general factor of prevalent psychopathology during adulthood? *Journal of Abnormal Psychology, 121*, 971–977. <http://dx.doi.org/10.1037/a0028355>
- Langer, M. (2008). *A reexamination of Lords Wald test for differential item functioning using item response theory and modern error estimation* (Unpublished doctoral dissertation). University of North Carolina, Chapel Hill, North Carolina.
- Livesley, W. J., & Jackson, D. (2009). *Manual for the dimensional assessment of personality pathology—basic questionnaire*. Port Huron, MI: Sigma.
- Lorenzo-Seva, U., & ten Berge, J. M. F. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 2*, 57–64.
- Lynam, D. R., & Widiger, T. A. (2007). Using a general model of personality to understand sex differences in the personality disorders. *Journal of Personality Disorders, 21*, 583–602. <http://dx.doi.org/10.1521/pedi.2007.21.6.583>
- MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130–149. <http://dx.doi.org/10.1037/1082-989X.1.2.130>
- Maples, J. L., Guan, L., Carter, N. T., & Miller, J. D. (2014). A test of the International Personality Item Pool representation of the Revised NEO Personality Inventory and development of a 120-item IPIP-based measure of the five-factor model. *Psychological Assessment, 26*, 1070–1084. <http://dx.doi.org/10.1037/pas0000004>
- Markon, K. E. (2010). Modeling psychopathology structure: A symptom-level analysis of Axis I and II disorders. *Psychological Medicine, 40*, 273–288. <http://dx.doi.org/10.1017/S0033291709990183>
- Markon, K. E., Krueger, R. F., & Watson, D. (2005). Delineating the structure of normal and abnormal personality: An integrative hierarchical approach. *Journal of Personality and Social Psychology, 88*, 139–157. <http://dx.doi.org/10.1037/0022-3514.88.1.139>
- McCrae, R. R. (2008). A note on some measures of profile agreement. *Journal of Personality Assessment, 90*, 105–109. <http://dx.doi.org/10.1080/00223890701845104>
- Miller, J. D., Gentile, B., Wilson, L., & Campbell, W. K. (2013). Grandiose and vulnerable narcissism and the *DSM-5* pathological personality trait model. *Journal of Personality Assessment, 95*, 284–290. <http://dx.doi.org/10.1080/00223891.2012.685907>
- Miller, J. D., & Lynam, D. R. (2001). Structural models of personality and their relation to antisocial behavior: A meta-analysis. *Criminology, 39*, 765–798. <http://dx.doi.org/10.1111/j.1745-9125.2001.tb00940.x>
- Miller, J. D., & Lynam, D. R. (2003). Psychopathy and the Five-factor model of personality: A replication and extension. *Journal of Personality Assessment, 81*, 168–178. http://dx.doi.org/10.1207/S15327752JPA8102_08
- Morey, L. C. (1991). *Personality Assessment Inventory professional manual*. Odessa, FL: Psychological Assessment Resources.
- Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making, 5*, 411–419.
- Pilkonis, P. A., Choi, S. W., Reise, S. P., Stover, A. M., Riley, W. T., & Cella, D. (2011). Item banks for measuring emotional distress from the Patient-Reported Outcomes Measurement Information System (PROMIS®): Depression, anxiety, and anger. *Assessment, 18*, 263–283. <http://dx.doi.org/10.1177/1073191111411667>
- Quilty, L. C., Ayeart, L., Chmielewski, M., Pollock, B. G., & Bagby, R. M. (2013). The psychometric properties of the Personality Inventory for *DSM-5* in an APA *DSM-5* field trial sample. *Assessment, 20*, 362–369.
- Rojas, S. L., & Widiger, T. A. (2014). Convergent and discriminant validity of the Five Factor Form. *Assessment, 21*, 143–157. <http://dx.doi.org/10.1177/1073191113517260>
- Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Suppl.*, No. 17.
- Samuel, D. B., Hopwood, C. J., Krueger, R. F., Thomas, K. M., & Ruggero, C. J. (2013). Comparing methods for scoring personality disorder types using maladaptive traits in *DSM-5*. *Assessment, 20*, 353–361. <http://dx.doi.org/10.1177/1073191113486182>
- Saunders, J. B., Aasland, O. G., Babor, T. F., de la Fuente, J. R., & Grant, M. (1993). Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption. II. *Addiction, 88*, 791–804. <http://dx.doi.org/10.1111/j.1360-0443.1993.tb02093.x>
- Simms, L. J., Goldberg, L. R., Roberts, J. E., Watson, D., Welte, J., & Rotterman, J. H. (2011). Computerized adaptive assessment of personality disorder: Introducing the CAT-PD project. *Journal of Personality Assessment, 93*, 380–389. <http://dx.doi.org/10.1080/00223891.2011.577475>
- Sprouse, J. (2011). A validation of Amazon Mechanical Turk for the collection of acceptability judgments in linguistic theory. *Behavior Research Methods, 43*, 155–167. <http://dx.doi.org/10.3758/s13428-010-0039-7>
- Strickland, C. M., Drislane, L. E., Lucy, M., Krueger, R. F., & Patrick, C. J. (2013). Characterizing psychopathy using *DSM-5* personality traits. *Assessment, 20*, 327–338. <http://dx.doi.org/10.1177/1073191113486691>

- Suzuki, T., Samuel, D. B., Pahlen, S., & Krueger, R. F. (2015). *DSM-5 alternative personality disorder model traits as maladaptive extreme variants of the Five-Factor Model: An item-response theory analysis. Journal of Abnormal Psychology*. Advance online publication. <http://dx.doi.org/10.1037/abn0000035>
- Thomas, K. M., Yalch, M. M., Krueger, R. F., Wright, A. G., Markon, K. E., & Hopwood, C. J. (2013). The convergent structure of *DSM-5* personality trait facets and five-factor model trait domains. *Assessment, 20*, 308–311. <http://dx.doi.org/10.1177/1073191112457589>
- Trull, T. J., & Durrett, C. A. (2005). Categorical and dimensional models of personality disorder. *Annual Review of Clinical Psychology, 1*, 355–380. <http://dx.doi.org/10.1146/annurev.clinpsy.1.102803.144009>
- Tucker, L. R. (1951). *A method for synthesis of factor analysis studies* (Personnel Research Section Report No. 984). Washington, DC: Department of the Army.
- van Kampen, D., de Beurs, E., & Andrea, H. (2008). A short form of the Dimensional Assessment of Personality Pathology-Basic Questionnaire (DAPP-BQ): The DAPP-SF. *Psychiatry Research, 160*, 115–128. <http://dx.doi.org/10.1016/j.psychres.2007.05.002>
- Verheul, R. (2006). Clinical utility of dimensional models for personality pathology. Dimensional models of personality disorders: Refining the research agenda for *DSM-V*. *Journal of Personality Disorders, 19*, 203–218.
- Watson, D., Stasik, S. M., Ro, E., & Clark, L. A. (2013). Integrating normal and pathological personality: Relating the *DSM-5* trait-dimensional model to general traits of personality. *Assessment, 20*, 312–326. <http://dx.doi.org/10.1177/1073191113485810>
- Widiger, T. A., & Costa, P. T., Jr. (2012). Integrating normal and abnormal personality structure: The Five-Factor Model. *Journal of Personality, 80*, 1471–1506. <http://dx.doi.org/10.1111/j.1467-6494.2012.00776.x>
- Widiger, T. A., & Simonsen, E. (2005). Alternative dimensional models of personality disorder: Finding a common ground. *Journal of Personality Disorders, 19*, 110–130. <http://dx.doi.org/10.1521/pedi.19.2.110.62628>
- Widiger, T. A., & Trull, T. J. (2007). Plate tectonics in the classification of personality disorder: Shifting to a dimensional model. *American Psychologist, 62*, 71–83. <http://dx.doi.org/10.1037/0003-066X.62.2.71>
- Wiggins, J. S., & Pincus, A. L. (2002). Personality structure and the structure of personality disorders. In P. Costa, Jr. & T. Widiger (Eds.), *Personality disorders and the five-factor model of personality* (2nd ed., pp. 103–124). Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/10423-007>
- Wilberg, T., Dammen, T., & Friis, S. (2000). Comparing Personality Diagnostic questionnaire-4+ with Longitudinal, Expert, All Data (LEAD) standard diagnoses in a sample with a high prevalence of axis I and axis II disorders. *Comprehensive Psychiatry, 41*, 295–302. <http://dx.doi.org/10.1053/comp.2000.0410295>
- Woods, C. M., Cai, L., & Wang, M. (2013). The Langer-improved Wald test for DIF testing with multiple groups: Evaluation and comparison to two-group IRT. *Educational and Psychological Measurement, 73*, 532–547. <http://dx.doi.org/10.1177/0013164412464875>
- Wright, A. G., Pincus, A. L., Hopwood, C. J., Thomas, K. M., Markon, K. E., & Krueger, R. F. (2012). An interpersonal analysis of pathological personality traits in *DSM-5*. *Assessment, 19*, 263–275. <http://dx.doi.org/10.1177/1073191112446657>
- Wright, A. G., & Simms, L. J. (2014). On the structure of personality disorder traits: Conjoint analyses of the CAT-PD, PID-5, and NEO-PI-3 trait models. *Personality Disorders, 5*, 43–54. <http://dx.doi.org/10.1037/per0000037>
- Wright, A. G., Thomas, K. M., Hopwood, C. J., Markon, K. E., Pincus, A. L., & Krueger, R. F. (2012). The hierarchical structure of *DSM-5* pathological personality traits. *Journal of Abnormal Psychology, 121*, 951–957. <http://dx.doi.org/10.1037/a0027669>
- Yam, W. H., & Simms, L. J. (2014). Comparing criterion- and trait-based personality disorder diagnoses in *DSM-5*. *Journal of Abnormal Psychology, 123*, 802–808. <http://dx.doi.org/10.1037/a0037633>
- Zickar, M. J. (1998). Modeling item-level data with item response theory. *Current Directions in Psychological Science, 7*, 104–109. <http://dx.doi.org/10.1111/1467-8721.ep10774739>

(Appendix follows)

Appendix
IRT-Based Reduced Set of PID-5 Items

Negative affectivity

Anxiousness: 79, 109, 130, 174
 Emotional lability: 122, 138, 165, 181
 Hostility: 38, 92, 158, 170
 Perseveration: 60, 80, 100, 128
 (Lack of) Restricted affectivity: 84r, 91r, 167r, 184r
 Separation insecurity: 50, 127, 149, 175
 Submissiveness: 9, 15, 63, 202

Detachment

Anhedonia: 23, 26, 124, 157
 Depressivity: 81, 151, 163, 169
 Intimacy avoidance: 89, 120, 145, 203
 Suspiciousness: 2, 117, 133, 190
 Withdrawal: 82, 136, 146, 186

Psychoticism

Eccentricity: 25, 70, 152, 205
 Perceptual dysregulation: 44, 154, 192, 217
 Unusual beliefs and experiences: 106, 139, 150, 209

Antagonism

Attention seeking: 74, 173, 191, 211
 Callousness: 19, 153, 166, 183
 Deceitfulness: 53, 134, 206, 218
 Grandiosity: 40, 114, 187, 197
 Manipulativeness: 107, 125, 162, 219

Disinhibition

Distractibility: 118, 132, 144, 199
 Impulsivity: 4, 16, 17, 22
 Irresponsibility: 129, 156, 160, 171
 (Lack of) Rigid perfectionism: 105r, 123r, 176r, 196r
 Risk taking: 39, 48, 67, 159

Received August 5, 2014
 Revision received February 12, 2015
 Accepted February 18, 2015 ■