OPEN

Research Article

The Developmental Profile Inventory: Constructing a Clinically Useful Self-Report for Levels of Psychodynamic Personality Functioning

MARIKE G. POLAK, PhD
LAURA VAN RIEL, MD, MSc
THEO J.M. INGENHOVEN, MD, PhD
HENRICUS L. VAN, MD, PhD

Objective: **The Developmental Profile Inventory (DPI) was constructed to assess psychodynamic personality functioning by self-report.**

Method: **On the basis of the frame of reference of the Development Profile interview method, a self-report was developed covering 3 domains, self, interpersonal functioning, and problem-solving strategies, which represent 6 maladaptive and 3 adaptive developmental levels of psychodynamic functioning. The DPI was administered to patients with personality disorders who were receiving psychotherapy (N = 179) and to normal controls (N = 228).**

Results: **The internal reliabilities of the subscales were in general in the fair to good range in the patient sample, (α = 0.67 to 0.88, ω_h = 0.52 to 0.87) and adequate to good in the healthy controls (α = 0.71 to 0.91, ω_h = 0.71 to 0.90). Mean item-rest correlations were adequate (0.30 to 0.50). Test-retest reliability was good (intraclass correlation = 0.73 to 0.91). The hypothesized factorial structure of the DPI with 9 subscales organized in 3 clusters was partly confirmed by confirmatory factor analysis χ^2/df = 2.37, root mean square error of approximation = 0.060, root mean residual square = 0.078, and comparative fit index = 0.630, with each factor showing over 80% standardized loadings > 0.30, and at least 75% loadings > 0.40. The DPI discriminated patients and healthy controls in a meaningful way. Correlations among the DPI and other self-report measures of global personality pathology and psychological complaints showed satisfactory convergent and discriminant validity.**

Conclusions: **The DPI is a promising self-report measure for assessing both adaptive and maladaptive patterns of psychodynamic personality functioning. The appropriate initial psychometric properties justify proceeding with more formal tests of construct validity and predictive performance in broader mental health settings.**

(*Journal of Psychiatric Practice 2018;24;239–252*)

KEY WORDS: psychodynamic personality functioning, assessment, DSM-5 Section III, alternative model, personality disorder, Developmental Profile

In recent decades, clinicians have expressed a need to diagnose personality pathology in a more clinically meaningful way than was done in the fourth and earlier editions of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV),[1] by utilizing psychodynamic concepts and offering a more balanced perspective on both healthy and maladaptive patterns of personality functioning.[2–7] Above all,

POLAK: Department of Psychology, Education & Child Studies (DPECS), Erasmus University Rotterdam, Rotterdam, The Netherlands; VAN RIEL: Novarum Centre for Eating Disorders & Obesity, Arkin Institute of Mental Health, Amsterdam, The Netherlands; INGENHOVEN: NPI, Arkin Institute of Mental Health, Amsterdam, The Netherlands; VAN: Arkin Institute of Mental Health, Amsterdam, The Netherlands

Please send correspondence to: Marike G. Polak, PhD, Department of Psychology, Education & Child Studies (DPECS), Erasmus University Rotterdam, Woudestein, T13-17, P.O. Box 1738, Rotterdam 3000 DR, The Netherlands (e-mail: polak@essb.eur.nl).

Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website, www.psychiatricpractice.com.

Acknowledgment: The authors would like to honor the late Professor Robert E. Abraham, founding father of the Developmental Profile. The authors thank him for his important contributions to the current paper and to designing the self-report instrument, the Developmental Profile Inventory, which is discussed in this article.

The authors declare no conflicts of interest.

DOI: 10.1097/PRA.0000000000000323

measurement of personality pathology on a dimensional scale requires a developmental perspective and addressing both pathologic and healthy aspects of personality functioning, which was often missing in clinical practice.[2,5] In addition, there was limited evidence for the clinical utility of the DSM-IV personality disorders (PDs) for individual case formulations and clinical decision-making, such as in planning appropriate psychotherapeutic treatment.[8] The goal of the introduction of the Alternative Model for PDs in Section III (Emerging Measures and Models) of the DSM-5,[9] which describes problems with personality functioning (ie, self-functioning and interpersonal-functioning) as well as maladaptive personality traits as the core of a PD, was to fulfill this need.

The developmental approach to personality functioning is both theoretically and practically most strongly elaborated in the psychodynamic perspective. Although measuring psychodynamic concepts in a reliable way has long been considered complex, in recent decades considerable progress has been made, which has resulted in the development of several assessment instruments. Methods that have been developed to assess and measure ego strength and psychodynamic personality functioning for clinical practice are mostly based on clinical interviewing, such as *Kernberg's Structural Interview*,[10,11] the *Developmental Profile*,[12] or more recently, the *Semi-Structured Interview for Personality Functioning in DSM-5* (STiP-5.1).[13,14] Available self-report screeners that give some indication of impairments in these structural and psychodynamic personality characteristics include the *General Assessment of Personality Disorder* (GAPD),[15,16] the *Inventory of Personality Organization* (IPO),[17] the *Severity Indices of Personality Problems* (SIPP-118),[18] the *Level of Personality Functioning Self Report* (LPFS-SR),[19] and *the LPFS Brief Form*.[20]

However, a validated self-report instrument measuring personality functioning from a more comprehensive developmental psychodynamic perspective has not yet been available. In this article, we introduce such a self-report measure, the *Developmental Profile Inventory* (DPI),[21] which is based on the *Developmental Profile* (DP) interview assessment procedure.[12,22,23]

The original DP provides an overview of strengths and vulnerabilities of personality functioning along various developmental lines, such as *social attitudes*, *object relations*, *self-esteem*, *norms*, *needs*, and *problem-solving strategies* (*defense mechanisms* and *coping styles*). Patients are scored on 9 hierarchically organized *developmental levels* that indicate the degree of maturity on every developmental line (see Appendix, Supplemental Digital Content 1, http://links.lww.com/JPP/A33, for an illustration of the DP model). The DP model distinguishes 3 "Primitive" levels of functioning (*Lack of Structure*, *Fragmentation*, *Egocentricity*), 3 "Neurotic" levels of functioning (*Dependency, Resistance*, *Rivalry*), and 3 Adaptive levels of functioning (*Individuation*, *Solidarity*, *Generativity*). The clinical usefulness of the DP interview relies on standardizing psychodynamic personality diagnostics to make it more convenient for diagnosis and treatment planning[22] as well as empirical research.

Multiple studies have examined the psychometric properties of the DP interview and scoring method.[23–29] These studies showed sufficient (interrater) reliability and internal consistency. Predictive validity with respect to the process and outcome of treatment was also established. Finally, empirical support was found for an underlying bipolar continuum, representing the hierarchical organization of the developmental levels ranging from primitive maladaptive behavioral patterns up to more neurotic functioning and adaptive, mature capabilities.

In line with the DP interview assessment procedure, the DPI operationalizes not only maladaptive behavioral patterns but also healthy personality characteristics. Within the DPI these characteristics are not mutually exclusive—that is, both types of functioning are measured on separate levels, next to each other.

Adult behavior can best be understood as the result of an interaction between these maladaptive and adaptive levels of functioning. Therefore, and in contrast to the Alternative DSM-5 Model for PDs, adaptive and maladaptive patterns of functioning can be assessed separately, next to each other in a nonexclusive way. In line with the DP, the DPI measures psychodynamic personality functioning within 3 domains: *self*, *interpersonal functioning*, and *problem-solving strategies* which includes both defense mechanisms and coping strategies. In this way, we believe the DPI captures 3 major personality concepts for formulating an appropriate case conceptualization, for treatment selection, and for tailoring treatment to the patient's capabilities,

needs, and individual goals. For instance, it might enable the therapist to differentiate between an insight-oriented approach or a more structured symptom-oriented or problem-oriented approach (eg, Van Manen et al[30]).

Moreover, as an easily applicable instrument, the DPI may serve to monitor processes during psychotherapeutic treatment and to measure outcomes of treatment and at follow-up. It could also facilitate research into psychodynamic concepts in larger populations. Finally, the DPI could be used as a screener for personality pathology to determine whether a more comprehensive psychodynamic assessment, such as administration of the full DP interview version, is indicated.

The DPI consists of 108 items that reflect the 9 *developmental levels* of psychodynamic functioning, each covering the domains of *self*, *interpersonal functioning*, and *problem-solving strategies*. Consistent with the DP, the DPI consists of 6 maladaptive developmental levels, *Lack of Structure*, *Fragmentation*, *Egocentricity*, *Dependence*, *Resistance*, and *Rivalry*, with the first 3 levels organized in a Primitive cluster and the second 3 levels organized in a Neurotic cluster. In addition, there is a cluster of 3 Adaptive developmental levels: *Individuation*, *Solidarity*, and *Generativity*.

OBJECTIVE AND STUDY DESIGN

The goal of this article is to introduce the DPI and to present results concerning its reliability and fundamental aspects of its validity. The following hypotheses were tested or explored. First, in terms of reliability, we expected the DPI scales to show both sufficient internal consistency and short-term test-retest reliability (ie, resistance to the influence of temporary emotional states). Second, in terms of validity, the study was designed to investigate construct validity by means of a confirmatory factor analysis (CFA); given the organization of items in the DPI, we expected to fit a correlated 3-second order factor model incorporating 9-first order factors. This model has the following structure: the first second order factor represents the Primitive cluster, with first order factors *Lack of Structure*, *Fragmentation*, and *Egocentricity*; the next second order factor represents the Neurotic cluster, with first order factors *Dependence*, *Resistance*, and *Rivalry*, the final second order factor represents the Adaptive cluster with first order factors *Individuation*, *Solidarity*, and *Generativity*. Furthermore, we expected the DPI scales to differentiate between normal controls and clinical patients with PDs (concurrent validity). Finally, we expected to find moderate, significant correlations among the subscales of the DPI and other self-report measures of global personality pathology (convergent validity) as well as significant, but lower in magnitude, correlations with self-reported psychological complaints (discriminant validity).

This study employed a cross-sectional design, with the exception of the test-retest measurements in subsamples, and explored pooled samples of patients and normal controls in the Netherlands.

METHOD

Participants

The sample of patients with PDs (PD sample; N = 179) represented referrals from 2012 through 2014 to 2 mental health care institutions in the Netherlands: Pro Persona Lunteren (n = 142) and GGz Central Amersfoort (n = 37). These institutions offer a variety of outpatient, day hospital, and inpatient services for the diagnosis and treatment of PDs, such as schema-focused therapy and psychodynamic psychotherapy.

Patients completed the DPI as part of the standard intake assessment procedures. Of the 179 patients, 30% were men and 70% women, mean age was 32.9 years (SD = 9.4 y, range = 19 to 56 y). Educational level was low (primary school/lower vocational education) in 19.0%, intermediate (secondary school/intermediate vocational education) in 49.7%, and high (upper vocational education/university) in 31.3%. All of the patients in the sample met criteria for at least one DSM-IV PD [as assessed by the LEAD[31] (longitudinal expert evaluation that uses all data)] procedure.

The healthy controls (control sample; N = 228) were comprised of 3 samples: a sample of health care professionals in training as psychiatrists or clinical psychologists (n = 96), a community sample of volunteers living in the proximity of Erasmus University, Rotterdam (n = 95), and a sample of health care workers and non–health care employees (eg, receptionists, office managers) of Arkin Mental Health Care Institution, Amsterdam (n = 37). Data were collected between 2012 and 2015. Normal

controls who reported having received psychological help at least once in the preceding 6 months were excluded from the healthy control sample, as were participants who were 75 years of age or older. In the control sample, 24.8% were men, 75.2% were women, and the mean age was 47.2 years (SD = 14.6 y, range = 22 to 74 y). Educational level was low in 6.6%, intermediate in 12.8%, and high in 80.5%. Informed consent was obtained from all participants in the study.

The concurrent validity analysis, in which the PD sample was compared with the control sample with respect to mean scores on the various subscales of the DPI, was conducted in a subsample from the control sample. To rule out demographic characteristics as confounders, we took a subsample that matched, as closely as possible, the PD sample with respect to age, sex, and educational level. We used a stratified random sampling procedure in which we randomly sampled participants from strata based on age, sex, and educational level. The main objective was to obtain a ratio of low versus intermediate/high educational level that matched the ratio in the PD sample (ie, 19% vs. 81%) while at the same time matching the ratio of men to women in the PD sample (ie, 30% vs. 70%).

In the resulting control subsample (n = 79), 30.4% were men and 69.6% were women, and the mean age was 46.5 years (SD = 16.2 y, range = 22 to 74 y). Educational level was low in 19.0%, intermediate in 36.7%, and high in 44.3%.

Test-retest reliability was examined in a subset of the original control sample (n = 101) and in a subset of the PD sample (n = 48). The participants completed the DPI self-report twice within a 13- to 37-day interval (mean = 17 d) and, in the case of the patients, before treatment.

Convergent and discriminant validity were examined in a subset of the PD sample (n = 98) for whom data were available on related self-report measures, namely the SIPP-118,[18] the PDQ-4+,[32] and the Outcome Questionnaire (OQ-45).[33]

Measures

SIPP-118

The SIPP-118[18] is a self-report tool measuring the severity of the generic and changeable components of PDs. The SIPP-118 consists of 118 items rated on a 4-point Likert scale, covering 16 facets of personality functioning, organized in 5 higher order domains: self-control, identity integration, relational capacities, responsibility, and social concordance. The SIPP-118 has been extensively used in both clinical and academic fields and its psychometric properties have been well established.[18,34]

The PDQ-4+

The PDQ-4+[32] (Dutch translation[35]) is a 99-item true-false self-report measure that corresponds directly to the criteria for DSM-IV PDs. The PDQ-4+ contains 1 item for each DSM-IV PD criterion, which can be summed to a PD scale total score. In the current study, the PDQ total score was calculated as an indication of overall personality disturbance. Furthermore, PD symptom counts were computed by summing the items for each PD. Psychometric properties of the current version of the PDQ-4+ justify using this instrument as a screener, in particular for the presence or absence of specific PDs.[36,37]

The OQ-45

The OQ-45[33] is a self-report instrument consisting of 45 items scored on a 5-point Likert scale. It is widely used to monitor clinical change in patients in mental health care settings. The instrument consists of 3 dimensions considered to be related to clinical change, namely, levels of psychiatric symptoms, performance in various social roles, and interpersonal functioning. In this study, the total score for the OQ-45 was calculated as a general indicator of psychological complaints. In earlier research (cross-cultural) validity was established, and reliabilities were found to be sufficient for most of the subscales and the total sum score.[33,38]

Constructing the DPI

The DPI was derived from the DP clinical assessment procedure and scoring protocol. To develop relevant items, the authors selected item descriptions from the DP scoring protocol, and prototypical statements made by patients during DP interviews, which were known to the authors of the DPI. The authors then selected the most relevant items from this aggregated list of potential items on the basis of an expert, intuitive, clinical point of view.

The original 9 developmental lines of the DP were merged into 3 DPI domains: *self*, *interpersonal functioning*, and *problem-solving strategies*. The first 2 domains seemed most relevant to retain in light of

TABLE 1. Examples of Maladaptive and Adaptive Items in the Developmental Profile Inventory

	Developmental Level	
Domain	***Dependence (maladaptive functioning)***	***Individuation (adaptive functioning)***
Self	I have an ongoing urgent need for the warmth or involvement of others	I have clear goals in my life, and I work them out systematically
Interpersonal functioning	I feel uncertain about my decisions, unless they are verified by others	I feel free to give my opinion, even when others donot agree with my point of view
Problem solving	When things turn out wrong, I quickly get discouraged	Whenever necessary, I will find an adequate way to stand up for myself

the formulation of the Alternative DSM-5 Model for PDs, and the third domain of problem-solving behavior was retained since it explicitly focuses on habitual defense mechanisms and coping strategies, and thus seems important as a treatment focus. Within these domains, we distinguished 6 maladaptive levels (ie, 3 Primitive levels, *Lack of Structure*, *Fragmentation*, and *Egocentricity*, and 3 Neurotic levels, *Dependence*, *Resistance*, and *Rivalry*), and 3 Adaptive levels (ie, *Individuation*, *Solidarity*, and *Generativity*). Within each developmental level, 4 items refer to 1 of the 3 domains, as a consequence each developmental level consists of 12 items. As in the original DP model, the developmental levels are organized in a hierarchical order.

Items in the DPI are presented as statements about habitual functioning, such as, *I can only relax when I have everything under control* (here, the level is *Resistance*, and the domain is *problem-solving strategies*). Table 1 shows some examples of items in the DPI. Respondents are asked to consider the degree to which statements habitually apply to them, thus how they *usually* experience themselves. The response format for each item is a 4-point Likert scale: not applicable at all (0), partly applicable (1), clearly applicable (2), and (almost) completely applicable (3). After the questionnaire is completed, item scores are summed to compute the relevant subscale (ie, developmental level) scores. These sum scores for the 9 developmental levels (ie, *Lack of Structure*, *Fragmentation*, *Egocentricity*, *Dependence*, *Resistance*, *Rivalry*, *Individuation*, *Solidarity*, and *Generativity*) compose the patient's DP.

In the next phase of constructing the DPI, 8 clinicians (clinical psychologists, psychotherapists, and psychiatrists) with extensive experience with the DP interview were asked to indicate, blindly from each other, the developmental level to which they believed each item referred. In addition, these clinicians could comment or make suggestions for improving the items. As a guideline, items that were correctly assigned to the corresponding developmental level by fewer than 6 expert judges (<75%) were revised. This criterion resulted in adjustments being made in 19 items.

Subsequently, the 8 clinicians judged the revised list of 108 items in a second round. For all levels, at least 75% of the judges assigned each item to the appropriate level. The average correct assignment among judges was 87.8% (SD = 16.0). The interrater reliability among judges was excellent with an intraclass correlation of 0.95 (95% confidence interval, 0.94-0.96). These results indicate good face validity.

Statistical Analyses

Internal Consistency/Composite Reliability

Internal consistency/composite reliability was determined using the Cronbach α coefficient,[39] the McDonald omega$_h$ (ω_h, where h relates to the hierarchical or general factor),[40,41] and item-rest correlations (*rir*, *which indicates, for each level, the correlation between each item and the total score, with that item excluded*), which were determined separately for the PD sample (N = 179), the control sample (N = 228), and the matched control subsample (n = 79). The coefficient ω_h indicates how well a common factor explains the variance in the scale scores for each subscale in the questionnaire.

We regarded 0.70 as the threshold for acceptable reliability and 0.20 as the threshold for acceptable average item-rest correlations.[42,43]

Factor Structure

Construct validity was examined using CFA with Robust Maximum Likelihood estimation to account for non-normality of item scores. Because a ratio of subjects-to-variables of 4:1 or larger in CFA is advised,[44] we studied the combined healthy and PD sample with N = 384 after list-wise deletion of missing values. Computations were performed using the R packages "lavaan"[45] to conduct CFA and "semTools"[46] to compute omega coefficients. We tested whether the assumed correlated 3-second order factor model incorporating 9-first order factors of the DPI could be confirmed. Each first order factor encompassed 12 items.

Goodness of Fit

We report the absolute fit measures χ^2, root mean residual square (RMR), and root mean square error of approximation (RMSEA) that represent the discrepancy between the observed and model implied covariance matrices. We included the χ^2/df ratio (ie, normed χ^2) because the χ^2 statistic is sensitive to sample size, whereas the normed χ^2 is adjusted for the number of degrees of freedom. Furthermore, we included the comparative fit index (CFI), which represents the proportion of improvement in fit compared with a null model, where the covariances among the observed variables are all assumed to be 0. We included CFI as it is an often reported measure of fit, although the appropriateness of using null models as comparative baselines has been questioned and RMSEA seems to be preferred over CFI (eg, Rigdon[47]). General guidelines indicate that, for reasonable model fit, the χ^2/df ratio should be in the range of 2 to 3 or less, RMR and RMSEA should be <0.08 and CFI > 0.90.[48,49] For the individual factor loadings, we used cut-off values of 0.30 and 0.40 as the minimum level of practical significance.[50,51]

Test-retest Reliability

Test-retest reliability was expressed by intraclass correlation coefficients (ICC), using a 2-way mixed model, single measure, absolute agreement. We interpreted ICC according to the rules of thumb of Landis and Koch[52]: ICC < 0.00 = poor, 0.00 to 0.20 = slight, 0.21 to 0.40 = fair, 0.41 to 0.60 = moderate, 0.61 to 0.80 = substantial, and 0.81 to 1.00 = almost perfect.

Concurrent Validity

Concurrent validity was investigated by testing univariate mean differences with *t* tests on all DPI scales and clusters of scales between the PD patient sample and the matched subsample of healthy controls. Effect size was expressed as Cohen *d*[53] and interpreted according to the standard rules of thumb with 0.20 = small, 0.50 = medium, and 0.80 = large.

Convergent and Discriminant Validity

Finally, convergent and discriminant validity were examined using Pearson correlations between all subscales of the DPI and related self-report measures, namely, the SIPP-118, the PDQ-4+, and the OQ-45. Strength of the linear relationship was interpreted according to Cohen's criteria,[53] with 0.10 = small, 0.30 = moderate, and 0.50 = large. All of the analyses, except for the CFA and omega coefficient, were conducted with IBM SPSS (version 23) statistical software.

RESULTS

Internal Consistency and Test-retest Reliability

Table 2 shows internal consistency reliability of the 9 hierarchically ordered DP levels and the 3 clusters of DP levels (Primitive, Neurotic, and Adaptive). In the healthy control sample (N = 228) and in the matched healthy control subsample (n = 79), all developmental levels showed acceptable (> 0.70) to good (> 0.80) internal consistency. All 3 clusters showed high values of α, around 0.90. In the PD sample (N = 179), internal consistency was acceptable to good for almost all levels and clusters of levels. The level of *Resistance* was the only exception, with an α just below the threshold of 0.70.

Mean item-rest correlations (*rir*) were good for all levels and clusters in all samples with values between 0.30 and 0.50. Most items showed *rir* > 0.20, not more than 2 items per developmental level showed values of *rir* < 0.20. In the PD sample, four items showed *rir* < 0.10, which showed fair to good *rir* in the control sample (range: 0.17 to 0.38, median = 0.23). In the control subsample, one item

TABLE 2. Cronbach α, McDonald Omega (ω_h) and Item-Rest Correlations (*rir*) for the Developmental Profile Inventory Scales in the Healthy Control Sample, the Matched Control Subsample, and the Sample With Personality Disorders (PD Sample)

	Control Sample (N = 228)			*Matched Control Subsample (n = 79)*			*PD Sample (N = 179)*		
Scales	***α***	***ω_h***	***rir [Mean (Range)]***	***α***	***ω_h***	***rir [Mean (Range)]***	***α***	***ω_h***	***rir [Mean (Range)]***
Adaptive	0.91	0.90	0.47 (0.25-0.63)	0.90	0.88	0.43 (0.14-0.59)	0.86	0.84	0.36 (0.06-0.57)
Generativity	0.78	0.78	0.43 (0.36-0.55)	0.72	0.69	0.36 (0.22-0.49)	0.74	0.69	0.37 (0.17-0.63)
Solidarity	0.85	0.85	0.54 (0.20-0.66)	0.83	0.82	0.50 (0.04-0.63)	0.80	0.80	0.45 (0.20-0.64)
Individuation	0.80	0.80	0.45 (0.27-0.57)	0.80	0.81	0.45 (0.24-0.62)	0.73	0.72	0.37 (0.18-0.49)
Neurotic	0.89	0.87	0.41 (0.18-0.62)	0.90	0.88	0.43 (0.14-0.69)	0.86	0.82	0.35 (0.03-0.60)
Rivalry	0.75	0.74	0.40 (0.17-0.49)	0.81	0.81	0.47 (0.21-0.70)	0.73	0.71	0.37 (0.05-0.59)
Resistance	0.71	0.71	0.35 (0.23-0.54)	0.70	0.70	0.34 (0.19-0.54)	0.67	0.52	0.31 (0.09-0.44)
Dependence	0.80	0.80	0.45 (0.29-0.58)	0.81	0.80	0.46 (0.27-0.73)	0.75	0.74	0.39 (0.18-0.60)
Primitive	0.90	0.90	0.43 (0.23-0.56)	0.90	0.90	0.43 (0.16-0.65)	0.88	0.87	0.38 (0.06-0.65)
Egocentricity	0.77	0.76	0.41 (0.24-0.57)	0.79	0.79	0.44 (0.19-0.64)	0.77	0.77	0.41 (0.08-0.58)
Fragmentation	0.79	0.77	0.44 (0.34-0.56)	0.79	0.78	0.44 (0.31-0.57)	0.80	0.79	0.44 (0.14-0.68)
Lack of structure	0.75	0.76	0.40 (0.27-0.52)	0.75	0.73	0.40 (0.29-0.60)	0.70	0.69	0.33 (0.06-0.51)

Scores for each developmental level were based on 12 items; higher order cluster scores (for Adaptive, Neurotic, and Primitive) were based on 36 items.

(question 19) had an unsatisfactory *rir* (0.04), whereas in the PD sample the *rir* for this item was fair (0.20).

Omega was closely in line with alpha, with the exception of *Resistance*, which showed an omega value in the PD sample that was far below the threshold of 0.70. In general, omega values were acceptable to good, which indicates the DPI scales reflect homogeneous constructs, and furthermore, that each of the 3 clusters (ie, Primitive, Neurotic, and Adaptive), can be considered underlying hierarchical factors. These results indicate adequate internal consistency and composite reliability, and can be interpreted as a first indication of adequate construct validity.

The theory underlying the DPI items, that there are 3 clusters, with 3 levels in each cluster, was tested by means of CFA. A correlated 3-second order factor model incorporating 9-first order factors was defined, in which each first order factor encompassed 12 items. The fit of the CFA model was $\chi^2_{5662} = 13401.74$, $P < 0.001$; $\chi^2/df = 2.37$; RMSEA = 0.060, 90% confidence interval = 0.058-0.061; RMR = 0.078; and CFI = 0.630, indicating that DPI is acceptable with respect to absolute fit, but is far below threshold with respect to comparative fit. This means that the hypothesized organization of items in the DPI is empirically demonstrable, although the outcomes are mixed. Standardized factor loadings are summarized in Table 3.

All of the factors had over 80% of items with standardized factor loadings > 0.30, and at least 75% loadings > 0.40. This is indicative of an adequate construct validity. The 3 most ill-fitting items were "I've done bad things, that just happens, you can't do anything about it" (*Lack of Structure*, loading = 0.07), "I put criticism to one side" (*Egocentricity*, loading = 0.10), and "As soon as I notice any form of injustice, I revolt" (*Resistance*, loading = 0.17). Modification indices showed that fit for these items could be improved by allowing cross loadings on, respectively, *Egocentricity*, *Resistance*, and again *Egocentricity*, meaning that these items show overlap with these constructs.

Both the Primitive and the Neurotic factors were, as expected, negatively correlated with the Adaptive factor ($r = -0.74$, $P < 0.001$, and $r = -0.71$, $P < 0.001$,

TABLE 3. Summary of Standardized Factor Loadings for the Correlated 3-Second Order 9-Factor Confirmatory Factor Analysis Model in the Combined Personality Disorder and Control Sample (N = 384)

Factors	*No. Factors/ Items*	*Mean of Factor Loadings*	*Range*	*No. Loadings > 0.30 (%)*	*No. Loadings > 0.40 (%)*
Adaptive	3	0.85	0.81-0.90	3 (100)	3 (100)
Generativity	12	0.49	0.30-0.63	12 (100)	9 (75)
Solidarity	12	0.60	0.39-0.75	12 (100)	10 (83.3)
Individuation	12	0.62	0.44-0.77	12 (100)	12 (100)
Neurotic	3	0.92	0.86-0.96	3 (100)	3 (100)
Rivalry	12	0.53	0.25-0.80	10 (83.3)	10 (83.3)
Resistance	12	0.57	0.17-0.75	10 (83.3)	10 (83.3)
Dependence	12	0.60	0.28-0.82	10 (83.3)	9 (75)
Primitive	3	0.77	0.47-0.93	3 (100)	3 (100)
Egocentricity	12	0.47	0.10-0.64	10 (83.3)	9 (75)
Fragmentation	12	0.63	0.29-0.76	11 (91.7)	10 (83.3)
Lack of structure	12	0.52	0.07-0.81	10 (83.3)	10 (83.3)

The first order factors Generativity, Solidarity, and Individuation load on the second order factor Adaptive (bold); the first order factors Rivalry, Resistance, and Dependence load on the second order factor Neurotic (bold); the first order factors Egocentricity, Fragmentation, and Lack of Structure load on the second order factor Primitive (bold). Each first order factor has 12 corresponding item factor loadings.

respectively). The Neurotic and Primitive factors showed a high positive intercorrelation, $r = 0.95$, $P < 0.001$, which indicates that the distinction between these 2 maladaptive clusters is less clear.

Table 4 shows the test-retest reliability for all DPI scales within an interval of 13 to 37 days in a control sample and a PD sample. In both samples, test-retest reliability of all subscales and clusters was good to excellent, with median ICC levels of 0.86 in the control sample and 0.81 in the PD sample, indicating robustness and no short-term state pollution.

Concurrent Validity

The mean DP level scores and standard deviations for the PD sample and the matched control sample are reported in Table 5. All between-group differences were significant ($P < 0.001$), except for the level of *Egocentricity* on which patients with PD were not distinguished from healthy controls. As hypothesized, healthy controls reported higher levels of adaptive characteristics and lower levels of maladaptive characteristics compared with patients with PD. The median effect size (d) for the difference between the PD sample and the control sample was 1.39, indicating large effect sizes and good concurrent validity.

Convergent and Discriminant Validity

Convergent validity was explored with Pearson correlation coefficients between developmental levels and the PDQ-4+ total score and the specific DSM PD symptom counts (Table 6). Correlations with the 5 higher-order domains of the SIPP-118 were also computed and are reported in the text below. Finally, discriminant validity was examined with the correlations between the developmental levels and the OQ-45 total score (Table 6).

First, convergent validity coefficients regarding the PDQ-4 symptom counts were inspected. As can be seen, all PD symptom counts were significantly positively correlated with both maladaptive cluster scores (ie, Primitive cluster scores and Neurotic cluster scores). In contrast, some but not all of the PD symptom counts were significantly negatively correlated with the Adaptive cluster scores. More specifically, severe cluster A PDs (especially Schizotypal

TABLE 4. Test-retest Reliability, ICC, for the Developmental Profile Inventory Scales in the Healthy Control Sample and the PD Sample

	Control Sample (n = 101)	*PD Sample (n = 48)*
Scales	***ICC (95% CI)***	***ICC (95% CI)***
Adaptive	0.80 (0.72-0.86)	0.84 (0.72-0.91)
Generativity	0.73 (0.62-0.81)	0.78 (0.64-0.87)
Solidarity	0.83 (0.76-0.88)	0.85 (0.74-0.93)
Individuation	0.79 (0.70-0.85)	0.87 (0.78-0.93)
Neurotic	0.85 (0.79-0.90)	0.85 (0.75-0.92)
Rivalry	0.81 (0.73-0.87)	0.88 (0.80-0.93)
Resistance	0.75 (0.65-0.83)	0.76 (0.63-0.87)
Dependence	0.83 (0.76-0.88)	0.86 (0.76-0.92)
Primitive	0.89 (0.84-0.93)	0.91 (0.84-0.95)
Egocentricity	0.77 (0.67-0.84)	0.86 (0.76-0.92)
Fragmentation	0.86 (0.80-0.91)	0.87 (0.77-0.97)
Lack of structure	0.81 (0.74-0.87)	0.81 (0.68-0.89)

ICC was based on a 2-way mixed model, single measure, absolute agreement; ICC <0.00 = poor, 0.00 to 0.20 = slight, 0.21 to 0.40 = fair, 0.41 to 0.60 = moderate, 0.61 to 0.80 = substantial, and 0.81 to 1.00 = almost perfect. ICC indicates intraclass correlation; PD, personality disorder.

and Paranoid) showed strong correlations with the Primitive cluster. The same applies to the severe cluster B PDs (especially Borderline and Narcissistic), which also showed strong correlations with the Primitive cluster. The strongest correlations with the Neurotic cluster were present for PDQ-4+ Dependent and Avoidant PD. The DPI level scores showed correlations with the PD symptom counts that further differentiate between the PD syndromes. For example, Borderline PD demonstrated its strongest correlations (> 0.40) with *Lack of Structure* and *Fragmentation*; Dependent PD showed the strongest correlation (0.60) with *Dependence*; and Narcissistic PD showed the strongest correlations (> 0.47) with the levels of *Rivalry* and *Egocentricity*. These results are largely consistent with our expectations and theoretical understanding, although not all PD symptom counts showed significant negative correlations with the Adaptive cluster.

Second, most levels and clusters of levels showed, as predicted, moderate to high correlations with the PDQ-4+ total score, indicating that higher scores for all maladaptive levels were substantially related to a higher number of criteria for PD as measured with the PDQ-4+. In addition, higher scores for all adaptive levels were substantially related to a lower number of PDQ-4+ criteria for PD. In general, the predicted convergent validity coefficients with overall personality disturbance were confirmed, with moderate to large effect sizes, with the exception of the level *Generativity*.

Third, convergent validity coefficients regarding the SIPP-118 were calculated. All maladaptive levels showed negative correlations with all SIPP domains, with r ranging from −0.13 to −0.62, median = −0.39. Adaptive levels correlated positively with SIPP domains, with r ranging from 0.15 to 0.61, median = 0.23), with, for example, strong correlations between Individuation and Identity Integration, $r_{96} = 0.52$, $P < 0.01$, and Solidarity showing strong correlations with both Relationship capacities, $r_{96} = 0.61$, $P < 0.01$ and Social Concordance $r_{96} = 0.47$, $P < 0.01$. Overall, the predicted convergent validity coefficients with the SIPP domains were confirmed, although the level *Generativity* showed the weakest correlations with the SIPP domains (r ranging from 0.15 to 0.23, median = 0.20).

Finally, most DPI developmental levels showed small to moderate, but significant, correlations with the OQ-45 total score (Table 6), indicating that higher scores for all maladaptive levels were related to higher levels of psychological problems and symptomatic distress, and higher scores for all adaptive levels were substantially related to a lower level of psychological problems. The fact that the absolute discriminant validity coefficients as indexed by the OQ-45 total score (median = 0.29) were smaller in magnitude than the absolute convergent validity coefficients as indexed by the PDQ4+ total score (median = 0.52) is indicative of satisfactory discriminant validity. As predicted, the DPI assesses characteristics that transcend mere symptom distress.

DISCUSSION

This study investigated reliability and aspects of validity of the DPI, a 108-item self-report measure designed to assess adaptive as well as maladaptive

TABLE 5. Average Scores Per DPI Scale in the Matched Healthy Control Subsample and the PD Sample

	Matched Control Subsample (N = 79)	*PD Sample (N = 179)*	*Comparison of Group Means*	
DPI Scales	***M (SD)***	***M (SD)***	***|t|***	***Cohen d***
Adaptive	73.9 (13.7)	55.2 (13.3)	10.26*	1.39
Generativity	23.5 (5.0)	20.0 (5.4)	4.98*	0.67
Solidarity	25.6 (5.5)	20.4 (6.0)	6.62*	0.89
Individuation	24.7 (5.4)	14.8 (5.4)	13.44*	1.82
Neurotic	28.1 (13.2)	52.2 (14.2)	12.83*	1.73
Rivalry	8.0 (5.3)	14.8 (6.0)	9.08*	1.72
Resistance	10.5 (4.5)	18.2 (5.5)	10.88*	1.47
Dependence	9.5 (5.2)	19.2 (5.8)	12.79*	1.73
Primitive	21.7 (12.1)	35.9 (13.8)	7.92*	1.07
Egocentricity	9.1 (5.0)	8.6 (4.9)	0.74	0.10
Fragmentation	6.3 (4.7)	14.7 (6.5)	11.54*	1.39
Lack of structure	6.3 (4.4)	12.7 (5.2)	10.29*	1.30

Level scores were based on 12 items, with the sum score potentially ranging from 0 to 36; higher order cluster scores were based on 36 items, with the sum potentially ranging from 0 to 108; Cohen d[53] 0.20 = small, 0.50 = medium, 0.80 = large.
DPI indicates Developmental Profile Inventory; PD, personality disorder.
**P < 0.001.*

patterns of psychodynamic personality functioning. The 9 subscales showed fair to good internal consistency (Cronbach α) in patients with PDs and sufficient to good reliability in normal controls. Moreover, all 3 clusters of subscales (Adaptive, Neurotic, and Primitive) showed good internal reliability. Furthermore, mean item-rest correlations (*rir*) were adequate in all samples for all levels and clusters. Values of the composite reliability (coefficient omega) were comparable to Cronbach α in both samples, with the exception of the level of Resistance in the PD sample. These results indicate adequate internal reliability, and are a preliminary indication of construct validity, although caution is warranted when interpreting individual scores on the DPI scales, in particular for the level of Resistance. In addition, short-term test-retest reliability was good in both samples, indicating the robustness of the DPI and limited state-pollution.

The hypothesized factorial structure of the DPI, with its 9 subscales organized in 3 clusters, as tested with a CFA correlated 3-second order factor model incorporating 9-first order factors, was partly confirmed. The outcomes of model fit were mixed: absolute fit was adequate, whereas comparative fit was far below threshold, with, on the other hand, an overall satisfactory pattern of factor loadings.

In addition, subsequent levels and clusters of levels discriminated significantly between patients with PDs and normal controls, except for the level of *Egocentricity*. As hypothesized, individuals from the general population reported higher levels of adaptive characteristics and lower levels of maladaptive characteristics compared with patients with PDs. The median effect size for the difference between the PD sample and the control sample was large, which is a first indication of concurrent validity, although more research is necessary to determine whether the DPI distinguishes patients with PDs from non-PD patients with psychological problems.

The results of the current study provide support for the discriminant and convergent validity of the DPI. In general, the developmental levels and clusters of levels were shown to converge adequately with related self-report measures, namely, the PDQ-4+ and the SIPP-118. An exception was the level of *Generativity*, which showed relatively small correlations with both instruments. *Generativity* refers to care for others and society in general. An explanation of the relatively weak

TABLE 6. Correlations Among the DPI, the PDQ-4+ Syndromes and Total Score and the OQ-45 Total Score (N = 98)

		Personality Disorder Symptom Counts								*PDQ-4+*	*OQ-45*
DPI Scales	***SZT***	***PAR***	***SZD***	***BPD***	***NAR***	***ANT***	***HIS***	***DEP***	***AVD***	***OCD***	***Total Score***	***Total Score***
Adaptive	−0.26*	−0.22*	−0.50**	−0.28**	0.05	−0.20	0.17	−0.17	−0.44**	−0.08	−0.34**	−0.30*
Generativity	−0.04	−0.05	−0.26*	−0.11	0.00	−0.14	0.08	0.06	−0.10	0.13	−0.07	−0.14
Solidarity	−0.29**	−0.27**	−0.55**	−0.25*	0.03	−0.19	0.27*	0.00	−0.37**	−0.22*	−0.32**	−0.24
Individuation	−0.27**	−0.19	−0.35**	−0.28**	0.07	−0.14	0.03	−0.46**	−0.56**	−0.07	−0.39**	−0.32**
Neurotic	0.37**	0.31**	0.21*	0.30**	0.39**	0.26*	0.43**	0.61**	0.55**	0.35**	0.65**	0.31**
Rivalry	0.25*	0.18	0.16	0.28**	0.49**	0.29**	0.44**	0.48**	0.37**	0.25*	0.57**	0.33**
Resistance	0.44**	0.35**	0.21*	0.26*	0.20	0.19	0.20	0.44**	0.55**	0.35**	0.55**	0.25*
Dependence	0.26*	0.24*	0.16	0.19	0.26*	0.17	0.41**	0.60**	0.45**	0.28**	0.48**	0.20
Primitive	0.47**	0.39**	0.27*	0.40**	0.47**	0.33**	0.37**	0.37**	0.33**	0.34**	0.63**	0.30*
Egocentricity	0.22*	0.09	0.08	0.19	0.48**	0.29**	0.22*	0.06	0.03	0.27**	0.32**	0.17
Fragmentation	0.45**	0.44**	0.25*	0.41**	0.39**	0.26*	0.40**	0.39**	0.33**	0.26*	0.61**	0.31*
Lack of Structure	0.48**	0.40**	0.31**	0.41**	0.37**	0.29**	0.30**	0.41**	0.36**	0.33**	0.62**	0.27*

*$P < 0.05$.

**$P < 0.01$.

ANT indicates antisocial; AVD, avoidant; BPD, borderline; DEP, dependent; DPI, Developmental Profile Inventory; HIS, histrionic; NAR, narcissistic; OCD, obsessive-compulsive; OQ-45, Outcome Questionnaire-45; PAR, Paranoid; PDQ-4+, Personality Diagnostic Questionnaire 4+; SZD, schizoid; SZT, schizotypal.

association might be that the SIPP-118 and PDQ-4+ do not capture these mature aspects of personality functioning. The fact that discriminant validity correlations of the DPI with the OQ-45 were, in general, smaller than the convergent validity coefficients indicates that the DPI explains additional variance in personality pathology rather than mere symptom distress.

Although the DPI discriminated patients and healthy controls in a meaningful way, the PD sample scores did not differ from healthy controls with respect to the level of *Egocentricity*. This finding is consistent with earlier findings concerning the DP interview data.[26] Similar scorings on *Egocentricity* in healthy controls and PD patients might reflect the question to what extent functioning at a narcissistic level is strictly related only to maladaptive functioning. Some studies have found indications that narcissistic perceptions of the self are also related to psychological health and generate a protective value for self-esteem dysregulation.[54] Also, from a clinical point of view, narcissism has been described as a dimension that spans the broadest spectrum of severity of personality pathology, ranging from almost healthy functioning to severe psychosocial disrupting behavioral patterns.[55]

Limitations and Strengths

This study had the typical limitations of an initial, smaller scale, validation study. First, the DPI was administered in only 2 samples, both limited in size. Both samples were combined to allow CFA, which limits the generalizabilty of results and prohibits cross-validation of alternative factor models. Consequently, this study did not allow strategies to improve model fit by, for instance, discarding items or allowing related items to have correlated errors based on modification indices. Future research needs to focus on CFA in larger samples to establish construct validity at the item-level and in different populations.

Second, within the cross-sectional study, we administered the DPI only once. Future research should have a longitudinal design to investigate, for instance, the predictive value of the DPI in monitoring treatment outcome. Third, additional validation of the DPI across a wider range of clinical samples, for instance psychiatric outpatients without PD or in a forensic subsample, is needed to further provide generalizability. Finally, in line with, for instance, studies into correspondence between interview measures and self-report measures of the same model,[56] correspondence between the DPI and the DP interview-version of the measure still needs to be established. Future research should demonstrate the utility of the DPI as a screener for the DP interview. Finally, the added value of the DPI compared with measures that are designed to measure problems with personality functioning as operationalized in the Alternative DSM-5 model (ie, by self report with the LPFS-SR[19] or the LPFS Brief Form[20]; or by clinical interview, ie, STiP-5.1[13] or the recently introduced Structured Clinical Interview for the DSM-5 Alternative Model for Personality Disorders; SCID-5-AMPD[57]) needs to be investigated.

Merits of the DPI are, first, the explicit distinction of both adaptive as well as maladaptive levels of psychodynamic functioning. In the DPI, both types of functioning are complementary within a single comprehensive strength-weakness analysis. This appears to be a clinically useful property, especially in personality assessment of complex cases. In particular, when little is known about their structural organization, patients' level of personality functioning is at risk of being overestimated or underestimated, resulting in inappropriate treatment allocation.[8] Diagnostic assessment of personality pathology on a descriptive level (as in DSM-IV and DSM-5 Section II) seems insufficient for indicating psychotherapy, because it offers too little information about both strengths and impairments in personality functioning. As psychodynamic assessment procedures, such as the DPI, provide complementary information about the level of personality functioning, they seem more suitable for treatment allocation, especially indication for psychotherapy.

Second, the hierarchical, dimensional organization of the developmental levels, ranging from primitive maladaptive functioning to adaptive mature functioning, enables a clinically relevant assessment of a client's overall level of personality functioning, which is in line with the Alternative DSM-5 Model for PDs (Section III). In addition, the DPI has the advantage of measuring both pathologic and healthy aspects of personality separately.

CONCLUSION

The results of this DPI study indicate adequate reliability and validity. The results can be considered in

agreement with psychoanalytic developmental theories that assume hierarchical levels of structural personality organization and distinguish healthy personality characteristics, neurotic personality functioning, and underlying primitive (borderline or psychotic) vulnerabilities and deficits.[4,58]

The DPI is easy and efficient to administer and offers a simple method for assessing clinically relevant psychodynamic themes, supporting diagnosis and treatment planning. Because of its easy application, the DPI may also serve to measure patients' progress during treatment. The DPI, which charts not only the decrease of pathological functioning but also the increase of adaptive capabilities, may as such be helpful during the therapy process to reformulate and differentiate individual goals for the patient. This use of the DPI is supported by the good test-retest reliability, although more research into treatment effect sizes is necessary.

The results of this study indicate that the DPI is a promising self-report measure for the assessment of psychodynamic personality functioning. Because the DPI showed appropriate initial psychometric properties, a more formal test of construct validity and predictive validity in broader clinical settings is warranted.

REFERENCES

1. American Psychiatric Association (APA). Diagnostic and Statistical Manual of Mental Disorders, 4th ed. Washington, DC: APA; 1994.
2. Widiger TA, Simonson E. Alternative dimensional models of personality disorder: finding a common ground. J Pers Disord. 2005;19:110–130.
3. Widiger TA, Simonson E, Sirovatka PJ, et al. Dimensional Models of Personality Disorders Refining the Research Agenda for DSM-V. Arlington, VA: American Psychiatric Association; 2006.
4. Psychodynamic Diagnostic Manual Task Force. Psychodynamic Diagnostic Manual. Silver Spring, MD: Alliance of Psychoanalytic Organizations; 2006.
5. Livesley WJ. A framework for integrating dimensional and categorical classifications of personality disorder. J Pers Disord. 2007;21:199–224.
6. Ingenhoven TJM, Abraham RE. Making diagnosis more meaningful. The developmental profile: a psychodynamic assessment of personality. Am J Psychother. 2010;64: 215–238.
7. Morey LC, Berghuis H, Bender DS, et al. Toward a model for assessing level of personality functioning in DSM-5, Part II: empirical articulation of a core dimension of personality pathology. J Pers Assess. 2011;93:347–353.
8. Verheul R. Clinical utility of dimensional models for personality pathology. J Pers Disord. 2005;19:283–302.
9. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 5th ed. Washington, DC: American Psychiatric Association; 2013.
10. Kernberg OF, Clarkin JF. The Inventory of Personality Organization. White Plains, NY: The New York Hospital-Cornell Medical Center; 1995.
11. Ingenhoven TJM, Duivenvoorden HJ, Brogtrop J, et al. Interrater reliability for Kernberg's Structural Interview for assessing personality organization. J Pers Disord. 2009;23:528–534.
12. Abraham RE. The Developmental Profile in Clinical Practice. Assen, The Netherlands: Van Gorcum; 2005.
13. Hutsebaut J, Berghuis H, Kaasenbrood A, et al. Semi-structured Interview for Personality Functioning DSM-5 (STiP-51). Podium DSM-5 research group of the Netherlands Centre of Expertise on Personality Disorders. Utrecht: Trimbos Institute; 2015.
14. Hutsebaut J, Kamphuis JH, Feenstra DJ, et al. Assessing DSM-5 oriented levels of personality functioning: development and psychometric evaluation of the Semi-Structured Interview for Personality Functioning DSM-5 (STiP 5.1). Personal Disord. 2016;8:94–101.
15. Livesley WJ. General Assessment of Personality Disorder (GAPD). Vancouver: Department of Psychiatry, University of British Columbia; 2006.
16. Berghuis H, Kamphuis JH, Verheul R, et al. The General Assessment of Personality Disorder (GAPD) as an instrument for assessing the core features of personality disorders. Clin Psychol Psychother. 2013;20: 544–557.
17. Clarkin JF, Foelsch PA, Kernberg OF. The Inventory of Personality Organization. White Plains, NY: The Personality Disorders Institute, Department of Psychiatry, Weill College of Medicine of Cornell University; 2001.
18. Verheul R, Andrea H, Berghout CC, et al. Severity Indices of Personality Problems (SIPP-118): development, factor structure, reliability, and validity. Psychol Assess. 2008;20:23–34.
19. Morey LC. Development and initial evaluation of a self-report form of the DSM-5 Level of Personality Functioning Scale. Psychol Assess. 2017;29:1302–1308.
20. Hutsebaut J, Feenstra DJ, Kamphuis JH. Development and preliminary psychometric evaluation of a Brief Self-Report Questionnaire for the Assessment of the DSM-5 Level of Personality Functioning Scale: The LPFS Brief Form (LPFS-BF). Personal Disord. 2015;7:192–197.
21. Ingenhoven TJM, Polak MG, Van HL, et al. Developmental Profile Inventory Version 10. Amsterdam: Development Profile Foundation; 2012.
22. Abraham RE, van Dam Q. The Developmental Profile: its use in establishing indications for psychoanalytic treatment. Bull Menninger Clin. 2004;68:115–135.
23. Abraham RE, Van HL, van Foeken I, et al. The Developmental Profile. J Pers Disord. 2001;15:457–473.
24. Ingenhoven TJM, Duivenvoorden HJ, Passchier J, et al. Treatment duration and premature termination of psychotherapy in personality disorders: predictive validity of the Developmental Profile assessing psychodynamic personality diagnosis. J Psychiatr Pract. 2012;18: 172–186.
25. Ingenhoven TJM, van den Brink W, Passchier J, et al. Treatment disrupting behaviors during psychotherapy of patients with personality disorders: the predictive power of psychodynamic personality diagnosis. J Psychiatr Pract. 2011;17:21–34.
26. Polak MG, Van HL, Overeem-Seldenrijk J, et al. The Developmental Profile: validation of a theory driven instrument for personality assessment. Psychother Res. 2010;20:259–272.

27. Van HL, Ingenhoven TJM, van Foeken I, et al. The Developmental Profile: preliminary results on interrater reliability and construct validity. J Pers Disord. 2000;14: 360–365.
28. Van HL, Hendriksen M, Schoevers RA, et al. Predictive value of object relations for therapeutic alliance and outcome in psychotherapy for depression: an exploratory study. J Nerv Ment Dis. 2008;196:655–662.
29. Van HL, Dekker J, Peen J, et al. Predictive value of self-reported and observer rated defense style in depression treatment: an exploratory study. Am J Psychother. 2009;63: 25–41.
30. Van Manen JG, Kamphuis JH, Goossensen A, et al. In search for patient characteristics that may guide emperically based treatment selection for personality disorder patients—a concept mapping approach. J Pers Disord. 2012;26:481–497.
31. Spitzer RL. Psychiatric diagnosis: are clinicians still necessary? Compr Psychiatry. 1983;24:399–411.
32. Hyler SE. The Personality Diagnostic Questionnaire 4+. New York, NY: New York State Psychiatric Institute; 1994.
33. Lambert MJ, Morton JJ, Hatfield D, et al. Administration and Scoring Manual for the OQ-45. Orem, UT: American Professional Credentialing Services; 2004.
34. Arnevik E, Wilberg T, Monsen JT, et al. A cross-national validity study of the Severity Indices of Personality Problems (SIPP–118). Personal Ment Health. 2009;3:41–55.
35. Akkerhuis GW, Kupka RW, Van Groenestijn MAC, et al. PDQ-4+: Questionnaire for Personality Traits. Lisse: Swets & Zeilinger; 1996.
36. Davison S, Leese M, Taylor PJ. Examination of the screening properties of the Personality Diagnostic Questionnaire 4+ (PDQ-4+) in a prison population. J Pers Disord. 2001;15:180–194.
37. Fossati A, Maffei C, Bagnato M, et al. Brief communication: criterion validity of the Personality Diagnostic Questionnaire-4+ (PDQ-4+) in a mixed psychiatric sample. J Pers Disord. 1998;12:172–178.
38. de Jong K, Nugter MA, Polak MG, et al. The Outcome Questionnaire (OQ-45) in a Dutch population: a cross-cultural validation. J Consult Clin Psychol. 2007;14:288–301.
39. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika. 1951;16:297–334.
40. McDonald RP. Generalizability in factorable domains: "domain validity and generalizability". Educ Psychol Meas. 1978;38:75–79.
41. McDonald RP. Test Theory: A Unified Treatment. Mahwah, NJ: Lawrence Erlbaum Associates; 1999.
42. Kline P. A Handbook of Test Construction: Introduction to Psychometric Design. London: Methuen; 1986.
43. Nunnally JC, Bernstein IH. Psychometric Theory, 3rd ed. New York, NY: McGraw-Hill; 1994.
44. MacCallum RC, Widaman KF, Preacher KJ, et al. Sample size in factor analysis: the role of model error. Multivariate Behav Res. 2001;36:611–637.
45. Rosseel Y. Lavaan: an R package for structural equation modeling. J Stat Softw. 2012;48:1–36.
46. SemTools Contributors. SemTools: useful tools for structural equation modeling. R package version 0.4-14. 2016. Available at: https://CRAN.R-project.org/package=semTools. Accessed November 6, 2017.
47. Rigdon EE. CFI versus RMSEA: a comparison of two fit indexes for structural equation modeling. Struct Equ Modeling. 1996;3:369–379.
48. Marsh HW. In search of golden rules: comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. Struct Equ Modeling. 2004;11: 320–341.
49. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling. 1999;6:1–55.
50. Costello AB, Osborne JW. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. Practical Assess Res Eval. 2005;10:1–9.
51. Ferguson E, Cox T. Exploratory factor analysis: a user's guide. Int J Select Assess. 1993;1:84–94.
52. Landis RJ, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33: 159–174.
53. Cohen J. A power primer. Psychol Bull. 1992;112: 155–159.
54. Sedikides C, Rudich EA, Gregg AP, et al. Are normal narcissists psychologically healthy? Self-esteem matters. J Pers Soc Psychol. 2004;87:400–416.
55. Caligor E, Levy KN, Yeomans FE. Narcissitic personallity disorder: diagnostic and clinical challenges. Am J Psychiatry. 2015;172:415–422.
56. Helle AC, Trull TJ, Widiger TA, et al. Utilizing interview and self-report assessment of the Five-Factor Model to examine convergence with the Alternative Model for Personality Disorders. Personal Disord. 2017;8:247–254.
57. Bender DS, Skodol AE, First MB, et al. Structured Clinical Interview for the DSM-5® Alternative Model for Personality Disorders (SCID-5-AMPD) Module I Level of Personality Functioning Scale. Washington, DC: APPI; 2018.
58. Kernberg OF. Severe Personality Disorders Psychotherapeutic Strategies. New Haven, CT and London: Yale University Press; 1984.