Towards a new perspective on personality disorder classification

The alternative model for personality disorders in clinical practice

Weekers, L.C.

Publication date
2024

Document Version
Final published version

Laura Cornelia Weekers

Towards a New Perspective on Personality Disorder Classification

The Alternative Model for Personality Disorders in Clinical Practice

ISBN: 978-94-6483-446-8

Provided by thesis specialist Ridderprint, ridderprint.nl
Printing: Ridderprint
Layout and design: Timo Wolf Kamp, persoonlijkproefschrift.nl

Towards a New Perspective on Personality Disorder Classification:
The Alternative Model for Personality Disorders in Clinical Practice

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. P.P.C.C. Verbeek
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op dinsdag 4 juni 2024, te 13:00

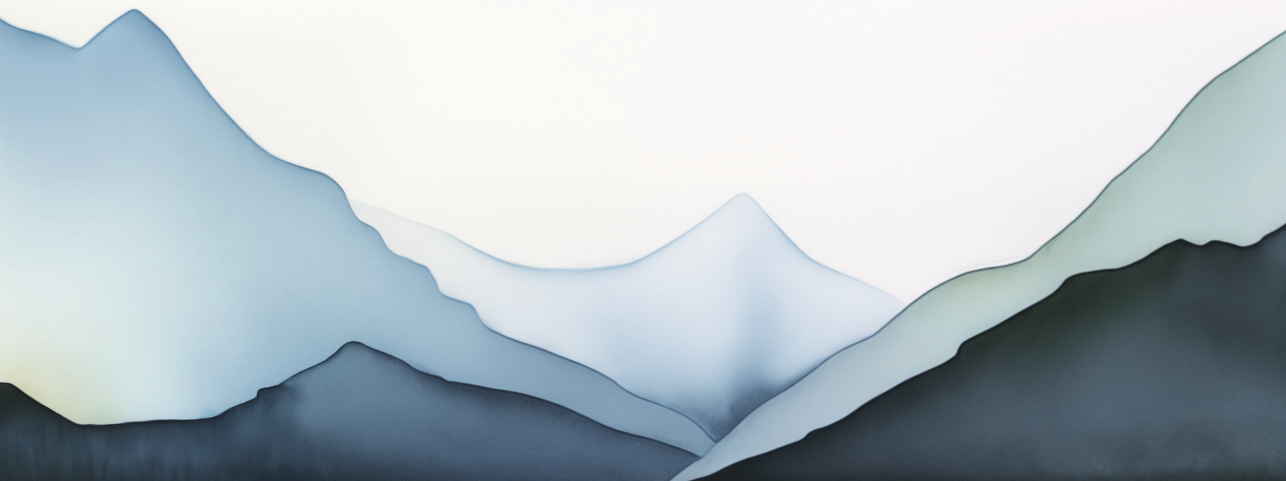
door Laura Cornelia Weekers
geboren te Hellevoetsluis

Promotiecommissie

Promotores:	prof. dr. J.H. Kamphuis	Universiteit van Amsterdam
	prof. dr. J. Hutsebaut	Tilburg University
Overige leden:	prof. dr. P.T. van der Heijden	Radboud Universiteit
	prof. dr. L. Claes	KU Leuven
	prof. dr. M. Kindt	Universiteit van Amsterdam
	prof. dr. R.W.H.J. Wiers	Universiteit van Amsterdam
	dr. A.A.P. van Emmerik	Universiteit van Amsterdam

Faculteit der Maatschappij- en Gedragswetenschappen

Contents



Chapter 1

General Introduction

The prevailing categorical classification system for personality disorders (PDs) in Section II of DSM-5 (American Psychiatric Association, 2013) has been extensively criticized for its lack of reliability, validity, and clinical utility. Studies have shown limited diagnostic reliability in terms of interrater agreement (Heumann & Morey, 1990; Samuel, 2015), limited validity as reflected by extensive heterogeneity within categories (Krueger et al., 2014; Skodol, 2014), (arguably) arbitrary diagnostic thresholds with types varying widely in terms of underlying severity (Balsis et al., 2011), temporal instability (Gunderson et al.,2011; Shea et al., 2002), high co-occurrence between types of PDs (Grant et al., 2005; Zimmerman et al., 2005), poor coverage of personality pathology by types with PD not otherwise specified being the most prevalent PD classification (Eaton et al., 2011; Krueger et al., 2014; Verheul & Widiger, 2004; Verheul et al., 2007), and limited clinical utility in terms of treatment selection and treatment planning (Natoli, 2019; Widiger & Samuel, 2005). These findings question the validity and utility of the categorical classification system.

To remedy many of these shortcomings, DSM-5 introduced the Alternative Model for Personality Disorders (AMPD; American Psychiatric Association, 2013). The research base of the AMPD model has been growing extensively over the past decade (e.g., Zimmerman et al., 2019). Research on validity and clinical utility of the model however, using ecologically valid designs, is still scarce. This thesis aims to address this gap in the research base and focusses on the validity and utility of the AMPD model in clinical practice. The research described in this thesis was written over a seven-year period. The context in which this project began will be discussed below.

The Alternative Model for Personality Disorders

The Alternative Model for Personality Disorders was introduced in Section III ('Emerging Measures and Models') and provides a hybrid ('categorical-dimensional') model consisting of three elements: 1) a dimensional assessment of the level of personality functioning (LPFS, Criterion A), (2) a dimensional assessment of pathological personality

traits (Criterion B) and (3) six specific personality disorders defined by disorder specific patterns of level of personality functioning and specific pathological personality traits. The PD diagnosis can be further specified by identifying elevated personality traits and level of severity.

Table 1. *Criterion A (APA, 2013)*

Self-functioning ***Identity***: Experience of oneself as unique, with clear boundaries between self and others; stability of self-esteem and accuracy of self-appraisal; capacity for, and ability to regulate a range of emotional experiences ***Self-direction:*** Pursuit of coherent and meaningful short-term and life goals; utilization of constructive and prosocial internal standards of behavior; ability to self-reflect productively
Interpersonal functioning ***Empathy:*** Comprehension and appreciation of others' experiences and motivations; tolerance of differing perspectives; understanding the effects of one's own behavior on others ***Intimacy:*** Depth and duration of connection with others; desire and capacity for closeness; mutuality of regard reflected in interpersonal behavior.

Criterion A requires moderate or more severe impairments in self- and interpersonal functioning. To assess these impairments, the Level of Personality Functioning Scale was introduced (LPFS; Bender et al., 2011). The LPFS assumes that PDs share 'essential commonalities' (Morey et al., 2011); impairments in capacities important for adaptive self- and interpersonal functioning. Elements of self-functioning are identity and self-direction, while empathy and intimacy constitute interpersonal functioning. These elements are further refined in twelve facets (see Table 1). Identity for example pertains to the ability to experience oneself as unique, with clear boundaries between self and others, self-esteem, and emotion-regulation. Functioning is rated on a continuum from Level 0 (no impairments), Level 1 (mild impairments), Level 2 (moderate impairments), Level 3 (severe impairments), to Level 4 (extreme impairments). Criterion B is used to describe *stylistic* elements of PD, comprising typical (maladaptive) ways of thinking, feeling, and acting (i.e., personality traits; see Table 2). A hierarchical trait model was developed with five higher order domains: Negative

affectivity, Detachment, Antagonism, Disinhibition, and Psychoticism (Krueger et al., 2012). These domains are further specified in 25 trait facets. For example, Negative affectivity contains the facets 'emotional lability', 'anxiousness', 'separation anxiety', 'submissiveness', 'hostility', 'depressivity', 'suspiciousness', and 'perseveration'.

Table 2. *Criterion B (APA, 2013)*

Negative affectivity: Frequent and intense experiences of high levels of a wide range of negative emotions (e.g., anxiety, depression, guilt/shame, worry, anger) and their behavioral (e.g., self-harm) and interpersonal (e.g., dependency) manifestations.
Detachment: Avoidance of socio-emotional experience, including both withdrawal from interpersonal interactions (ranging from casual daily interactions to friendships to intimate relationships) and restricted affective experiences and expression, particularly limited hedonic capacity.
Antagonism: Behaviors that put the individual at odds with other people, including an exaggerated sense of self-importance and concomitant expectations of special treatment, as well as callous antipathy toward others, encompassing both an unawareness of others' needs and feelings and a readiness to use others in the service of self-enhancement.
Disinhibition: Orientation toward immediate gratification, leading to impulsive behavior driven by current thoughts, feelings, and external stimuli, without regard for past learning or consideration of future consequences.
Psychoticism: Exhibiting a wide range of culturally incongruent odd, eccentric, or unusual behaviors and cognitions, including both process (e.g., perception, dissociation) and content (e.g., beliefs).

For a PD diagnosis, at least moderate impairments in personality functioning (Criterion A) and one pathological personality trait (Criterion B) must be present. Six specific PDs are retained in the AMPD model: Borderline, Antisocial, Avoidant, Obsessive-compulsive, Narcissistic, and Schizotypal PD. A trait-specified PD can be classified when patients do not meet criteria for a specific PD but have moderate or more severe impairments in personality functioning and at least one pathological personality trait.

The publication of the AMPD model introduced a major paradigm shift in the operationalization of PDs. The field is shifting from a descriptive model of specific symptoms related to patterns of experiencing and behavior towards an explanatory model describing

underlying mental processes and dispositions. Allport (1937) stated almost a century ago "personality is something and personality does something". PD classification prior to the AMPD focused mainly on what personality *is* in terms of behaviors and less on what personality *does* (Livesley, 1998). The AMPD model is a shift towards the latter, as personality pathology is reframed as the inability to achieve certain universal life tasks: e.g., developing a stable sense of self, establishing intimate relationships with others, cooperating effectively. This is radically different from the traditional Section II PD model in several ways. The Section II PD model focusses on observable symptoms such as social avoidance or clinging behavior, while the AMPD model focusses on underlying mental processes such as an inability to establish satisfying intimate relationships and a disposition to experience separation anxiety in relationships, which can be accompanied by certain behaviors (such as clinging or avoidance), but this is not a prerequisite for classification. Furthermore, the Section II PD model adopts a categorical system, in which a certain number of symptoms are required for the patient to meet the diagnostic threshold, arbitrarily distinguishing 'pathology' from 'no pathology'. The AMPD model also retains a categorical component (at least moderate impairments in personality functioning), but by introducing the LPFS a dimensional severity component was added describing the whole spectrum of personality functioning, from healthy to extremely impaired. Moreover, the AMPD model was the first in DSM history to offer a definition of healthy personality functioning (the levels '0' of the LPFS). Of note, whereas the AMPD model is a radical change for DSM, self- and interpersonal functioning have been central in most theories concerning (mal)adaptive personality (Livesley, 1998; Waugh et al., 2017), as well as in empirical studies for decades (Zimmerman et al., 2022). Since the publication of DSM-III however, PDs were described as distinct categories with specific symptoms (Natoli, 2019). Symptom focus aided interrater reliability, but probably at the cost of validity and clinical utility.

Utility and validity

Validity and clinical utility are complex multifaceted constructs with some definitions of utility also containing aspects of validity. Diagnostic validity can be broadly divided into the following aspects: face validity, descriptive or discriminant validity, predictive validity, incremental validity, and construct or convergent validity (First et al, 2004; Zimmermann, 2022). Face validity pertains to whether the diagnostic criteria appear to accurately describe the disorder. Descriptive or discriminant validity refers to whether the description of criteria is unique for a specific diagnosis and differs sufficiently from other diagnoses (First et al., 2004; Zimmermann et al., 2022). Predictive validity is the degree with which a diagnosis can predict future clinical course and outcomes, while incremental validity is concerned with whether a (new) diagnosis/classification system provides additional information for predicting clinical outcomes. Lastly, construct or convergent validity is related to whether a diagnosis has meaningful associations with expected external validators, i.e., the same or similar constructs.

A broad definition of clinical utility was provided by First and colleagues (2004). They describe clinical utility of a classification system in terms of five diagnostic functions (First et al., 2004, p.947):

1. Conceptualizing diagnostic entities
2. Communicating clinical information to relevant others
3. Ease of use of the diagnostic categories and criteria
4. Choosing effective interventions to improve clinical outcomes
5. Predicting future clinical management needs

Several authors suggest that the conceptualization of diagnostic entities and prediction of future clinical needs are better considered as aspects of (construct and predictive) validity (Mullins-Sweatt & Widiger, 2009; Zimmermann et al., 2022). Core elements of clinical utility can thus be described as: ease of use, communication, and treatment planning. The latter is also referred to as treatment utility,

and often seen as the 'holy grail' of utility (Nelson-Gray, 2003; First et al., 2004).

Although efforts were made to clarify the distinction between validity and utility, it should be noted that they are inherently related to one another. A classification system loses its utility when it has no validity. Improving validity often has a positive impact on utility, for example a higher predictive validity (predicting clinical course) can aid treatment selection. However, validity and utility can also have a negative impact on one another, for example when a diagnosis is simplified and easier to use at the cost of construct validity (First et al., 2004).

Rationale for the current thesis

The publication of the AMPD model and radical paradigm shift implied huge challenges for the clinical field, one of the reasons why the model was placed in Section III (Emerging measures and models) instead of replacing the traditional PD model in Section II of DSM-5. There were several important gaps in the research base.

First, there were no assessment instruments to reliably assess the LPFS, hindering the utility of the AMPD model. With the publication of DSM-5 the Personality Inventory DSM-5 (PID-5; Krueger et al., 2012), a self and informant-report questionnaire to assess Criterion B became available. The LPFS (Criterion A) was described in DSM-5 but upon publication no measures were available for assessing impairments in functioning in a structured way. Initially the DSM-5 workgroup predicted the LPFS would be relatively easy to assess in a general clinical interview, however, several early studies demonstrated low interrater reliability when using interview schedules not specifically developed to assess the LPFS (Few et al., 2013; Young & Beazley, 2023; Zimmerman et al., 2014). There was a need for new assessment instruments to reliably assess the LPFS. The Semi-structured interview for personality functioning DSM-5 (STiP 5.1; Hustebaut et al., 2017) was one of the first interview schedules developed to assess the LPFS in a structured way. Psychometric properties of the STiP 5.1 were promising

in an adult sample, however replication of these findings in different populations (e.g. adolescents) was necessary. Moreover, the Level of Personality Functioning Brief Form (LPFS-BF; Hutsebaut et al., 2016) was the first brief self-report questionnaire developed to provide a global impression of personality functioning. Preliminary findings were promising, but the questionnaire had several limitations which needed to be addressed. One of the aims at the start of this thesis was thus to further develop and evaluate these instruments for assessing the LPFS and test the utility of the LPFS for identifying personality problems in youngsters.

Second, there was the controversial issue of clinical utility. Although developed to improve clinical utility of PD classification, early critics were concerned that the model would be cumbersome to use and difficult to implement in clinical practice (Clarkin & Huprich, 2011). Guidelines for implementation were missing and needed. Another utility related concern was the fear of large discontinuity with the previous model in terms of identifying PD types (Clarkin & Huprich, 2011; Frances, 2012). A translational gap could have major implications for optimal treatment selection. Comparing the Section II PD- to the AMPD model in terms of case-identification of PDs was thus another important subject.

Lastly, there was a need for head-to-head comparisons of the AMPD and Section II PD model in terms of validity and utility. Radically changing a classification system is only justified when improvement, or at least equivalence, in validity and/or utility is demonstrated. It was hypothesized that the AMPD model better captures the underlying vulnerability inherent to PDs, however this notion had to be tested in empirical studies. Another aim of the current thesis was thus to compare both models in terms of their ability to predict long-term functioning. Furthermore, although several consumer surveys had been conducted on clinician rated utility of both models (Bornstein & Natoli, 2019), a more robust test of clinical utility was still lacking. Also, clinical utility research was solely focused on clinicians, patients themselves were never included in clinical utility research. This thesis

describes a randomized controlled trial comparing the clinical utility of the AMPD- and Section II PD models as rated by both clinicians and patients.

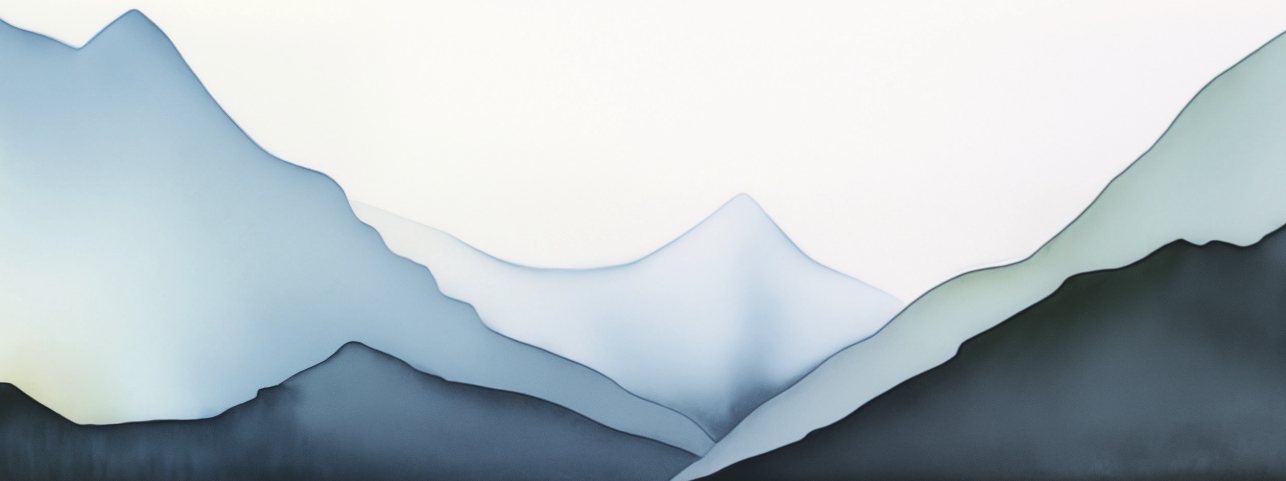
Aims and outline

Summarizing, the overall aim of this thesis was to assess the validity and clinical utility of the AMPD model and facilitate implementation of the model in clinical practice. This led us to the following research aims:

1. Develop specific instruments for assessing Criterion A and evaluate the impact of using the AMPD model in terms of case-identification of PD diagnoses;
2. Describe an AMPD assessment procedure which can be implemented in clinical practice and define what constitutes clinical utility of PD assessment;
3. Compare the AMPD model to the traditional Section II PD model in terms of predictive validity and clinical utility.

The first part of this thesis describes the evaluation of several instruments for assessing Criterion A and the impact of using the model on case-identification of PDs. **Chapter 2** describes the development and psychometric evaluation of the LPFS-BF 2.0, a brief self-report questionnaire for assessing Criterion A. In **Chapter 3** the Semi-structured interview for personality functioning DSM-5 (STiP 5.1) is evaluated for assessing Criterion A in adolescents. **Chapter 4** describes the impact of using the AMPD model in terms of case-identification of PDs. The second part of this thesis focusses on implementation and clinical utility of the model. **Chapter 5** presents a detailed case to demonstrate the use of the complete AMPD model in clinical practice. **Chapter 6** presents the development of client and clinician definitions of clinical utility of assessment. The third part of this thesis focusses on a comparison of the AMPD model to the traditional Section II PD model. **Chapter 7** compares the models in terms of predictive validity

In **Chapter 8** a randomized controlled trial is described, comparing the clinical utility of the models as rated by patients and clinicians. In the final part of this thesis, the general discussion **(Chapter 9)**, the main findings are summarized and discussed.



Chapter 2

The Level of Personality Functioning Scale – Brief Form 2.0 (LPFS-BF 2.0): Update of a brief instrument for assessing level of personality functioning

Laura C. Weekers, Joost Hutsebaut, Jan H. Kamphuis
Personality and Mental Health (2019), 13, 3-14

Abstract

Section III of The *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; DSM-5; American Psychiatric Association, 2013) introduced the alternative model for personality disorders that includes assessing levels of personality functioning. Here, we describe the development, preliminary psychometric evaluation, and sensitivity to change of a revised brief self-report questionnaire, the Level of Personality Functioning Scale-Brief Form 2.0 (LPFS-BF 2.0). Patients (N = 201) referred to a specialized center for the assessment and treatment of personality disorders completed the LPFS-BF 2.0, the Brief Symptom Inventory, the Severity Indices of Personality Problems Short Form, and were administered the Structured Clinical Interview for DSM-IV Axis I and Axis II Disorders. Internal structure and aspects of construct validity were examined. A subsample of 39 patients also completed the questionnaires after 3 months of inpatient treatment. Confirmatory factor analyses demonstrated better fit for a two-factor solution (interpretable as Self- and Interpersonal Functioning) than for a unidimensional model, though acceptable model fit was evident only after two post-hoc modifications. The LPFS-BF 2.0 demonstrated satisfactory internal consistency and promising construct validity. Sensitivity to change after 3 months of treatment was high. The LPFS-BF 2.0 constitutes a short, user-friendly instrument that provides a quick impression of the severity of personality pathology.

Introduction

The Level of Personality Functioning Scale (LPFS) was introduced in The Alternative Model for Personality Disorders in DSM-5 (American Psychiatric Association [APA], 2013) to provide a measure for the assessment of impairments in personality functioning. The model builds upon the assumption that all types of personality disorders are characterized by 'essential commonalities' with regard to moderate or more severe limitations in self and interpersonal functioning (Morey et al., 2011; Hutsebaut et al., 2017). These 'commonalities' are thought to be reflected by 12 facets, including impairments in identity (experience of oneself as unique, stability of self-esteem, capacity for and ability to regulate a range of emotional experience), self-direction (pursuit of coherent and meaningful goals, constructive and prosocial internal standards of behavior, self-reflection), empathy (comprehension and appreciation of others' experiences and motivations, tolerance of differing perspectives, understanding the effects of one's own behavior on others), and intimacy (depth an duration of connection with others, desire and capacity for closeness, mutuality of regard). The LPFS identifies 5 levels of functioning for each of these 12 facets, offering a severity index for personality pathology. The addition of a severity dimension is seen as a major addition to the traditional assessment of maladaptive personality traits. Severity of personality pathology is a strong predictor of current and future functioning (Hopwood et al., 2011), and likely has greater impact on treatment planning and course of treatment than the particular type of personality problems (Hopwood, 2018; Hopwood et al., 2011; Morey et al., 2015; Tyrer, 2005). Although the LPFS is described in the DSM-5 as a unidimensional construct, studies to date yielded inconsistent factor structures. Morey (2017) found a single factor solution and thus argued that the Criterion A is a unidimensional construct. Zimmerman and colleagues (2015), however, concluded that the LPFS was best conceptualized as a two-dimensional construct. They found two distinct factors:

Self- and Interpersonal Functioning. This is in line with a study by Berghuis and colleagues (2013), which corroborated the two-factors of the General Assessment of Personality Disorders (GAPD): Self- and Interpersonal pathology. Bastiaansen and colleagues (2013) used the Severity Indices of Personality Problems 118 (SIPP-118) to assess the LPFS and concluded that the LPFS consists of four-factors, i.e., Self-Control, Identity Integration, Relational Functioning and Responsibility. Previous research on the structure of the SIPP-118 by Verheul and colleagues (2008) yielded a five-factor solution. In sum, the results to date are inconclusive with regard to the structure of Criterion A of the AMPD (i.e., levels of personality functioning). In addition to the LPFS, the Alternative Model for Personality Disorders included 25 pathological personality traits, organized by 5 higher order domains (Negative Affectivity, Detachment, Antagonism, Disinhibition, and Psychoticism) for which the Personality Inventory for DSM–5 was proposed as assessment instrument (Krueger et al., 2012).

The AMPD was primarily designed to meet the shortcomings regarding validity and clinical utility of the prevailing model. However, soon after publication, concerns were raised concerning the presumed complexity of the model . Indeed, findings with regard to the application of the LPFS revealed some mixed results when using clinical interview data or SCID-data (Few et al., 2013; Thylstrup et al., 2016). On the other hand, other studies demonstrated the model lends itself well for instruction, such that graduate students and inexperienced raters were able to apply the model with adequate interrater reliability (Dereboy et al., 2018; Garcia et al., 2018; Zimmermann et al., 2014). An important way to improve clinical utility and ease of use, is to develop assessment instruments for assessing the LPFS and pathological personality traits. Since its publication, several instruments for assessing the LPFS have been developed independently by different research groups, including two interview schedules (Hutsebaut et al., 2017; Thylstrup et al., 2016) and (at least) three self-report questionnaires. Huprich and colleagues (2017) developed the DSM-5 Levels of Personality Functioning Questionnaire

(DLOPFQ), a 132-item questionnaire assessing the LPFS in both social and work/school domains. Initial results were promising, with high internal consistency rates and conceptually relevant correlations with maladaptive personality traits and overall wellbeing. Morey (2017) developed the Level of Personality Functioning Scale – self report, an 80-item self-report scale. The LPFS-sr includes items for each marker of severity as proposed by the LPFS, leading up to 80 items to represent 60 descriptions of severity. The LPFS-sr demonstrated high internal consistency, high test-retest reliability, high intercorrelations between each of its dimensions and high correlations with related instruments (Hopwood et al., 2018; Morey, 2017). Our group developed the Level of Personality Functioning – Brief Form (LPFS-BF; Hutsebaut et al., 2016). This instrument was initially developed as a quick screening tool related to the LPFS. Our primary aim was to formulate one item for each facet of the LPFS, yielding a global estimate of impairment related to personality functioning. The LPFS-BF thus became a very brief instrument, including only 12 items to be rated 'yes' or 'no'. Therefore, both the LPFS-BF and LPFS-sr may have different areas of application, with the LPFS-BF offering a 'quick and dirty' assessment of general impairment in personality functioning, while the LPFS-sr might enable a more precise and detailed assessment of different domains of personality functioning (Morey, 2017).

Although the LPFS-BF was initially developed to only serve as a website screening tool for patients to self-assess whether their problems might be related to personality dysfunction, the instrument showed acceptable psychometric properties. It yielded a clear two-factor solution, resembling Self and Interpersonal domains, and the internal consistencies in a sample of patients with personality pathology were borderline acceptable, with coefficient alphas of .69 for the total score, and .57 and .65 for the subscales respectively (Hutsebaut et al., 2016). With regard to construct validity, the LPFS-BF scores were associated as expected with related measures of personality pathology. On the other hand, analyses also demonstrated that some items of the original scale did not perform well, specifically

item 6 (I am often very strict with myself, referring to impairments in constructive and prosocial internal standards of behavior as an aspect of self-direction) and item 11 (There is almost no one who is really close to me, referring to impairments in desire and capacity for closeness as an aspect of intimacy). The item-total correlation of these questions was low and deletion of these items resulted in better internal consistency. With the newly formulated item 11 we tried to capture the subjective sense of a lack of safety in close relationships which is characteristic of more severe disturbance in the closeness facet. The reformulated item now reads as 'I often feel very vulnerable when relations become more personal'. We reformulated item 6 to capture a more severe level of self-direction: 'I often make unrealistic demands on myself'. Furthermore, (only) one of the initial items (item 4) was reversed (*I have clear aims in my life and succeed in achieving these*, referring to 'goals' as an aspect of self-direction). However, as the absence of health might not necessarily equal the presence of pathology and vice versa, we changed the reversed item. The updated LPFS-BF 2.0 therefore consists of 9 of the original items and 3 reformulated items.

In addition, to improve psychometric functioning, we opted for a response scale instead of a binary yes/no response format. This modification related to our aim of expanding the use of the LPFS-BF 2.0 as a screening tool to a tool for assessing changes in personality functioning during treatment. Assessing (lack of) progress during treatment is increasingly included in treatments of mental disorders in order to inform treatment decisions, for example reformulating treatment goals or terminating treatment (Lamber, 2007; Lambert et al., 2005). In the Netherlands, Routine Outcome Monitoring (ROM) was introduced nationwide in 2011 and typically consists of systematic periodic data collection on the mental health and level of functioning of patients as an indicator of treatment outcome (Buwalda et al., 2011; Nugter & Buwalda, 2012). Although using ROM during treatment to inform treatment decisions is considered clinically useful by its advocates, several prominent clinical researchers have raised

concern about indiscriminate use of ROM for benchmarking (using ROM data to compare treatment results), potential bias, confounds, and the need for disorder-specific instruments to more accurately assess the complexity of what constitutes treatment outcome (Mulder & Kortrijk, 2012; Van Os et al., 2012). Moreover, implementation of disorder-specific instruments in treatment for personality disorders is hindered by lack of data on sensitivity to change for most personality questionnaires, and many conceptually relevant questionnaires are too lengthy for multiple assessments over treatment. By including a response scale – similar to the Personality Inventory for DSM-5 (PID-5) response scale – we intended to increase variation in responses and therefore facilitate sensitivity of the instrument to identify relevant changes in personality functioning during treatment.

In sum, this study investigated aspects of reliability and construct validity of the updated version of the LPFS-BF (Hutsebaut et al., 2016), the LPFS-BF 2.0. We expected the internal structure of the LPFS-BF 2.0 to reflect two intercorrelated, internally consistent factors corresponding to Self- and Interpersonal functioning domains. Futhermore, we expected conceptually meaningful associations with related measures of personality functioning, the SIPP-SF and the DSM-IV-TR personality disorders. With respect to ROM purposes, we tested associations with a widely used routine oucome measuring questionnaire, the Brief Symptom Inventory (BSI; Derogatis, 1975), and compared their respective sensitivities to change in the context of a residential treatment program for PD.

Method

Participants

Two subsamples of patients were used in the analysis. All participants were treatment seeking adults who were referred to *de Viersprong*, a specialized mental health care center for the assessment and treatment of adolescents and adults with personality disorders. The

first sample of 201 participants completed the LPFS-BF 2.0 as part of the standard admission procedure. All intakes took place between April 2016 and February 2017. About two thirds of the total sample (*n* = 131; 65.2%) were female. Patients' age ranged from 18 to 62 years old, with a mean age of 36.2 (*SD*=11.0). Clinical characteristics of the participants are presented in Table 1; for 18 participants data on clinical characteristics were missing. Most patients met criteria for at least one personality disorder (90.7%), with borderline and personality disorder not otherwise specified (PD-NOS) being the most prevalent personality disorders. The second sample of 47 participants were administered the LPFS-BF 2.0 at the start of their 3-month residential treatment program, based on a Transactional Analysis treatment model (Berne, 1996). The comprehensive treatment program specifically targeted patients with a cluster C personality disorder and includes psychotherapy, psychomotor- and art therapy, sociotherapy, and milieu therapy. Questionnaires were collected between September 2016 and November 2017. Clinical characteristics of the second sample are presented in Table 1; data were missing for one participant. Thirty-nine of the 47 participants also completed the LPFS-BF 2.0 at the end of treatment. These data were used in the subsequent (treatment responsivity) analyses.

Table 1 *Diagnostic Characteristics of Sample 1 & 2*

DSM-IV-TR Diagnosis	*Sample 1* (*N* = 183) *N (%)*	*Sample 2* (*N* = 46) *N (%)*
Personality disorders		
Avoidant PD	40 (21.9)	29 (63)
Dependent PD	5 (2.7)	1 (2.2)
Obsessive-compulsive PD	24 (13.1)	9 (19.6)
Paranoid PD	3 (1.6)	0 (0)
Histrionic PD	1 (0.5)	0 (0)
Narcissistic PD	9 (4.9)	3 (6.5)
Borderline PD	63 (34.4)	7 (15.2)
Antisocial PD	5 (2.7)	0 (0)
PD-NOS	81 (44.3)	21 (45.7)
Any PD	166 (90.7)	44 (95.7)
Clinical disorders		
Mood disorder	97 (64.2)	24 (52.2)
Anxiety disorder	65 (36.3)	14 (30.4)
Substance use disorder	19 (11.8)	2 (4.3)
Psychotic disorder	1 (0.5)	0 (0)
Somatoform disorder	19 (10.4)	4 (8.7)
Eating disorder	16 (9)	3 (6.5)
Any Axis-I disorder	142 (86.1)	34 (79.1)

Note: PD= personality disorder; NOS = not otherwise specified; The sum of the number of patients across the different diagnostic groups is higher than the total number of patients because of comorbidity.

Measures

Level of Personality Functioning Scale-Brief Form 2.0.

The LPFS-BF 2.0 is a brief self-report questionnaire which assesses the LPFS as described in Section III of the DSM-5 (APA, 2013). The LPSF consists of 12 items, clustered into two higher order domains Self- and Interpersonal functioning. Participants are asked to rate the 12 items on a 4-point Likert scale from 1 (completely untrue) to 4 (completely true). Table 2 shows the distribution of responses of all items in the current sample.

Table 2 *Distribution of LPFS-BF 2.0 responses (N =201)*

		Mean	*SD*
1.	I often do not know who I really am	3.04	1.01
2.	I often think very negatively about myself	3.49	0.77
3.	My emotions change without me having a grip on them	3.27	0.87
4.	I have no sense of where I want to go in my life	3.17	0.92
5.	I often do not understand my own thoughts and feelings	3.14	0.90
6.	I often make unrealistic demands on myself	3.10	0.87
7.	I often have difficulty understanding the thoughts and feelings of others	2.17	1.03
8.	I often find it hard to stand it when others have a different opinion	2.46	1.00
9.	I often do not fully understand why my behavior has a certain effect on others	2.33	0.97
10.	My relationships and friendships never last long	2.29	1.07
11.	I often feel very vulnerable when relations become more personal	3.20	0.95
12.	I often do not succeed in cooperating with others in a mutually satisfactory way	2.14	0.99

Structured Clinical Interview for DSM-IV Axis I disorders.
The SCID-I (First et al., 1997; Groenestijn et al., 1999) is a semi-structured interview designed to assess the DSM-IV Axis I disorders. The SCID-I has demonstrated good interrater reliability in a diversity of samples, especially when interviewers had received a formal training; overall kappa was .85 (Ventura et al., 1998).

Structured Clinical Interview for DSM-IV Axis II personality disorders.
The SCID-II (First et al., 1997; Weertman et al., 1996) was used to diagnose DSM-IV Axis II personality disorders. The Axis II personality disorder criteria were largely kept unchanged in DSM-5 allowing the SCID-II to assess also DSM-5 personality disorders. Criteria were scored when the clinician deemed sufficient evidence present that the targeted behaviors were present, as well as pathological, pervasive, and persistent. Personality disorder not otherwise specified (PD-NOS) was classified when 5 criteria from personality disorders were present

(Verheul et al., 2007). The SCID-II has good interrater and test-retest reliability in personality disorder samples (Maffei et al., 1997; Weertman et al., 2003) with sum ICC's reported as high as .90 for avoidant and .95 for borderline personality disorder in a Dutch sample (Lobbestael et al., 2011).

Brief Symptom Inventory.
The BSI (Derogatis, 1975; De Beurs, 2006) was used to assess symptom severity. It consists of 53 items covering nine symptom dimensions (i.e., Somatization, Obsession-Compulsion, Interpersonal Sensitivity, Depression, Anxiety, Hostility, Phobic anxiety, Paranoid ideation and Psychoticism). The present study only utilized the BSI total score, which provides an index of the intensity of distress by psychological symptoms during the past week. Respondents rate each item on a 5-point scale ranging from 0 (not at all) to 4 (extremely). Cronbach's α in the present sample was .95.

Severity Indices of Personality Problems.
The SIPP-SF (Feenstra et al., 2011; Verheul et al., 2008) is a dimensional self-report measure designed to assess core components of (mal-)adaptive personality functioning. The 60 item SIPP-SF asks respondents to think back to the past three months and indicate the extent to which they agree with the presented statements. The response categories range from 1-4 and are described as fully disagree, partly disagree, partly agree, and fully agree. The measure comprises five higher-order domains labelled: (a) Self-control, (b) Identity Integration, (c) Relational Capacities, (d) Responsibility, and (e) Social Concordance. High scores indicate better adaptive functioning. The comprising SIPP-SF subscales have generally yielded adequate to strong internal consistencies in personality disorder samples, with Cronbach's α ranging from .62 to .89 (Feenstra et al., 2011; Verheul et al., 2008). In the current sample α scores ranged from .83 to .89.

Results

Internal Structure

To test the hypothesized two-factor model of the LPFS, and compare this fit to a unidimensional rendering of personality dysfunctioning, we conducted Confirmatory Factor Analyses (CFA) using M*plus* 7 (Muthén & Muthén, 2010). Model fit was evaluated by using absolute fit indices including the Root Mean Square Error of Approximation (RMSEA) and Standardized Root Mean Residual (SRMR), and relative fit indices including the Confirmatory Fit Index (CFI) and the Tucker-Lewis Index (TLI). We followed common guidelines for the interpretation of fit, with RMSEA and SRMR values of .05–.08 suggesting acceptable fit, and CFI and TLI values of .90–.95, respectively (Brown, 2014; Little, 2013). The Chi-square statistic is also reported but this statistic is generally considered less useful for the evaluation of model fit as it is overly sensitive to sample size.

Table 3 reports the fit indices of the alternative models. First, we tested a 1-factor model in line with previous research suggesting the LPFS is a unidimensional construct (Morey, 2017). All fit indices indicated a poor fit to the data. Next, we tested the hypothesized 2-factor model of the LPFS-BF 2.0 (Self- and Interpersonal functioning). This improved model fit considerably, though all fit indices remained below acceptable levels. Closer inspection of fit indices led to subsequent respecifications of the model, particularly with respect to the two-factor solution. Specifically, the modification indices suggested that item 11 ('I often feel very vulnerable when relations become more personal') was highly correlated to factor 1 (Self functioning), and that specification of a crossloading of item 11 on factor 1 would improve fit. Moreover, allowing the error terms of items 10 and 11 to correlate would also enhance model fit. We tested these modifications in subsequent models 3 and 4. Fit indices for model 3 were generally below acceptable thresholds, whereas for model 4 absolute fit (as measured by RMSEA and SRMR) was acceptable, with relative fit

indices slightly below (TLI) or above (CFI) customary thesholds. The post-hoc modifications made conceptual sense, as item 11 mentions feelings of vulnerability that (also) map onto deficits in Self functioning (model 3), and both item 10 and item 11 have a unique feature in introducing the context (and key word) of "relationship", beyond the specification of experienced difficulties in core tasks of personality functioning (model 4). Model 4 is shown in Figure 1. Of note, the LPFS-BF 2.0 showed robustness in that analyses with and without item 11 yielded highly similar results. The internal consistency estimates for the LPFS-BF 2.0 were high, with α=.82 for the total scale and α=.79 and α=.71 for the Self- and Interpersonal Functioning Scales. Correlation between the Self- and Interpersonal Functioning Scales was moderate (*r* = .44).

Table 3 *Confirmatory Factor Analyses: Fit Indices for Alternative Model Specifications*

	χ^2	df	RMSEA	90% CI	SRMR	TLI	CFI
Model 1 (1 factor; Unidimensional model)	229.802	54	.127	.111; .144	.096	.646	.710
Model 2 (2 factors; Self and Interpersonal Functioning)	145.294	53	.093	.075; .111	.083	.811	.848
Model 3 (2 factors; Self and Interpersonal; crossloadings 11 on factor 1)	120.198	52	.081	.062; .100	.066	.857	.888
Model 4 (2 factors; crossloadings 11 on factor 1, correlated errors 10 – 11)	106.282	51	.073	.054; .093	.061	.882	.909

Note. RMSEA = Root Mean Square Error of loadings; CI = Confidence Interval; SRMR = Standardized Root Mean square Residual; TLI = Tucker Lewis Index; CFI = Comparative Fit Index.

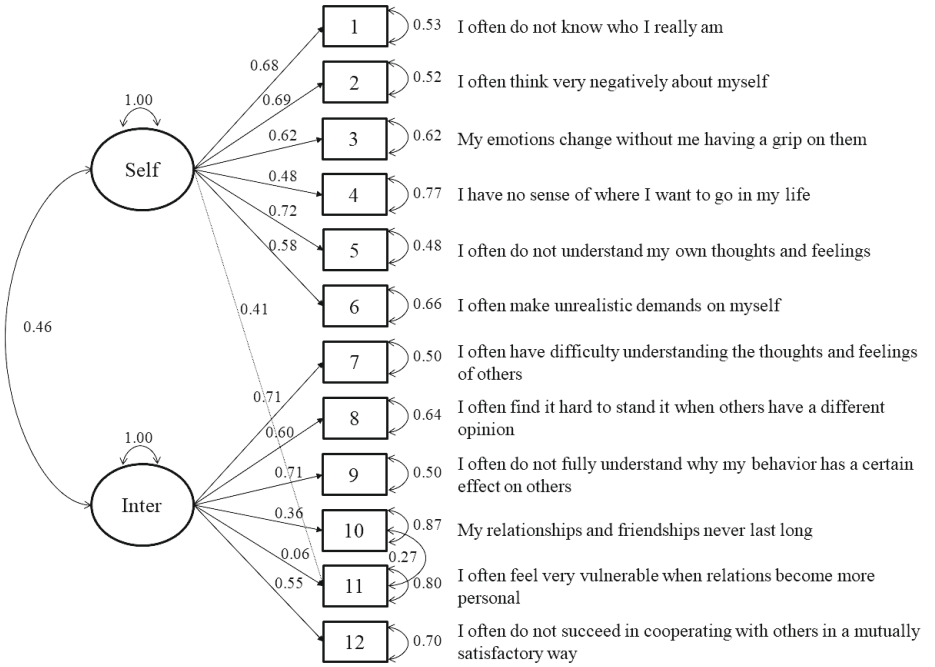


Figure 1. *LPFS-BF 2.0 final model after Confirmatory Factor Analyses.*
Note. Self = Self functioning domain; Inter = Interpersonal functioning domain.

Construct validity

Small to moderate associations were observed between the LPFS-BF 2.0 and severity of personality disorder, as measured by the number of personality disorder diagnoses (r = .33, r = .27 and r = .28 for the Total-, Self- and Interpersonal scales, respectively). In addition, the number of personality disorder criteria were significantly associated with the LPFS-BF 2.0 (r = .38, r = .33 and r = .30 for the Total-, Self- and Interpersonal scales, respectively). We also assessed whether the LPFS-BF 2.0 differentiated between patients with and without a borderline personality disorder, as several studies indicate borderline PD may be considered a measure of general severity (Bender & Skodol, 2007; Sharp et al., 2015). In our sample 63 patients met criteria for a borderline PD and 138 patients did not meet criteria for BPD (see also Table 1 for the distribution of PD diagnoses in our sample). An independent samples t-test showed a significant difference on the

LPFS-BF 2.0 between patients with a borderline PD (M = 37, SD = 5.72) and without a borderline PD (M = 32.38, SD = 5.72; t = -5.07, p < .001).

The LPFS-BF 2.0 showed moderate correlations with the BSI and SIPP-SF domains (see Table 4). All correlations were significant at p < .01. The BSI and SIPP Identity Integration domain were more strongly related to the LPFS-BF 2.0 Self-Functioning domain than the LPFS-BF 2.0 Interpersonal Functioning domain (z = 2.82, p = .005 and z = 5.74, p < .001 respectively). In addition the SIPP Social Concordance domain had a stronger relationship with the LPFS-BF 2.0 Interpersonal Functioning domain than the Self-Functioning domain (z = 6.69, p < .001). No differences were found for other SIPP domains. To test whether the LPFS-BF 2.0 and the SIPP are exclusively correlated due to shared general psychopathology variance, we assessed their relationship while controlling for the BSI by calculating partial correlations. The LPFS-BF 2.0 total score remained significantly correlated to all SIPP-SF domains. After controlling for BSI scores, correlations of LPFS-BF 2.0 Self-functioning score and the SIPP self control and social concordance domains were no longer significant. For the LPFS-BF 2.0 Interpersonal functioning score only the correlation with the identity integration domain was no longer significant.

Table 4 *Pearson Correlations with Self-report Measures of Personality Problems and Symptom Severity (N=182-187)*[a]

	SIPP-SF Self control	SIPP-SF Identity Integration	SIPP-SF Responsibility	SIPP-SF Relational capacities	SIPP-SF Social concordance	BSI total score
LPFS-BF 2.0 Total	-.50	-.50	-.37	-.49	-.52	.56
LPFS-BF 2.0 Self	-.38	-.62	-.29	-.38	-.24	.57
LPFS-BF 2.0 Interpersonal	-.48	-.25	-.36	-.48	-.66	.39
Diff (p)	.10	<.001**	.31	.15	<.001**	.005**

Note: SIPP-SF = Severity Indices of Personality Problems Short Form; BSI= Brief Symptom Inventory; Diff (p)= p-value of difference between Self functioning and Interpersonal functioning domains. [a] N varies due to missing values.
**p < .01.

Sensitivity to change

Sensitivity to change is the ability of an instrument to detect changes when these occur. Three methods have been described to assess sensitivity to change (Deyo et al., 1990): (1) effect sizes (*M2 – M1 / SD1*, *M1*= mean at time 1, *M2*= mean at time 2, *SD1* = standard deviation at time 1; Kazis et al., 1989), (2) standardized response mean (*M2 – M1 / SDdiff*, *SDdiff*= standard deviation of score changes; Liang et al., 1990) and (3) responsiveness index (*M2 – M1 / SDstable*, *SDstable*=standard deviation in unchanged subjects; Guyatt et al., 1987). There currently is no consensus about the best measure of sensitivity to change. We calculated both effect sizes (Cohen's *d*) and the standardized response mean to assess sensitivity to change of the LPFS-BF 2.0. Since results were very similar we chose to only report Cohen's *d*. Table 5 shows a summary of the main findings. Mean time between start and end of treatment was 92.13 days (*SD*=14.55). The LPFS-BF 2.0 shows high sensitivity to change, yielding an effect size of *d*=1.05 at the end of the 3-month inpatient treatment. Effect size of the LPFS-BF 2.0 was comparable to or higher than other measures commonly used for Routine Outcome Monitoring (BSI and SIPP-SF). The Self-functioning domain of the LPFS-BF 2.0 appeared to be especially sensitive to change, yielding an effect size of *d* = 1.22. Rank order stability measured by pre- post correlations is also reported in Table 5. Due to lack of power, no meaningful comparison of these estimates across instruments is possible in the present sample, but moderate rank order stability across respondents and instruments can be observed.

Table 5 *Sensitivity to change of the LPFS-BF 2.0 and related constructs (N=36-39[a])*

	Start of treatment M(SD)	End of treatment M(SD)	Correlation pre-post	Change score M (SD)	*p-value*	Effect size (*d*)
LPFS-BF 2.0 Total	30.54 (5.83)	24.23 (6.32)	.49	6.31 (6.15)	<.001	1.05
LPFS-BF 2.0 Self	17.54 (3.79)	12.73 (4.19)	.47	4.81 (4.14)	<.001	1.22
LPFS-BF 2.0 Interpersonal	13.05 (3.09)	11.46 (3.26)	.51	1.59 (3.13)	.003	0.51
SIPP-SF Self control[b]	35.67 (6.60)	38.61 (7.05)	.53	-2.94 (6.65)	.012	0.44
SIPP-SF Identity integration	25.89 (9.0)	34.36 (8.57)	.72	-8.47 (6.59)	<.001	0.98
SIPP-SF Responsibility	35.33 (7.49)	37.81 (5.92)	.70	-2.47 (5.41)	.010	0.37
SIPP-SF Relational capacities	28.64 (8.25)	33.72 (8.37)	.68	-5.08 (6.66)	<.001	0.62
SIPP-SF Social concordance	36.94 (6.46)	38.53 (6.58)	.66	-1.58 (5.34)	.084	0.25
BSI	1.40 (0.66)	0.79 (0.58)	.47	0.61 (0.64)	<.001	1.0

Note: SIPP-SF = Severity Indices of Personality Problems Short Form; BSI= Brief Symptom Inventory. [a] *N* varies due to missing values. [b] SIPP-SF scores are T-scores, comparing the scores to the normal population, with higher scores reflecting more adaptive functioning (T<30 very low, T= 30-40 low, T= 40-60 average, T= 60-70 high, T>70 very high).

Discussion

In this study, we tested the factor structure, reliability, construct validity and sensitivity to change of the LPFS-BF 2.0 in two samples of personality disordered patients. In line with our previous study (Hutsebaut et al., 2016), the structure of the LPFS-BF 2.0 total scale grossly adhered to two meaningful subscales: Self- and Interpersonal

Functioning. Distribution of the items over the subscales was mostly in line with what was expected, supporting the content validity of the LPFS-BF 2.0. Moreover, internal consistencies of the LPFS-BF 2.0 were satisfactory for both the total scale and the Self- and Interpersonal Functioning subscales. However, one item (item 11) hypothesized to load on the Interpersonal Functioning domain, loaded on the Self-Functioning domain (*'I often feel very vulnerable when relationships become more personal').* In retrospect this item, especially its first part, might also tap (deficits in) emotion regulation, an aspect of Self functioning. Future translations and adaptations may consider modifying this item to more accurately reflect its interpersonal facet origin. On the other hand, it is noteworthy that Zimmerman and colleagues (2015) found a similar deviation from the original theoretical model in an observer-report questionnaire. In their study, the Depth and Duration of Connections facet (most equivalent to item 10 of the LPFS-BF 2.0) was more indicative of Self- than Interpersonal functioning, and the Desire and Capacity for Closeness (counterpart of item 11 of the LPFS-BF 2.0) was equally related to Self- and Interpersonal Functioning.

Conceptually meaningful associations were observed between the LPFS-BF 2.0 and other measures of severity of personality disorders. As expected, the SIPP identity integration domain was more strongly related to the LPFS-BF 2.0 Self functioning domain, whereas the SIPP social concordance domain had the strongest relationship with the Interpersonal functioning domain of the LPFS-BF 2.0. Associations between number of personality disorders and number of personality disorder criteria were moderate. Patients with a borderline PD showed higher impairment scores than patients without a borderline PD. In line with previous research, borderline personality disorder appears to be a general marker of severity of personality pathology (Sharp et al., 2015). Conceptually, borderline personality disorder and the level of personality functioning scale show considerable overlap; borderline personality disorder is often conceptualized as a disorder of self- and interpersonal dysfunction (Bender & Skodol, 2007). The Self- and

Interpersonal Functioning subscales showed positive associations with similar constructs as measured by the SIPP-SF, with the identity integration subscale of the SIPP-SF showing a stronger relationship with the Self functioning domain and the Social Concordance subscale showing a stronger relationship with the Interpersonal Functioning domain. These findings support the construct validity of the scale.

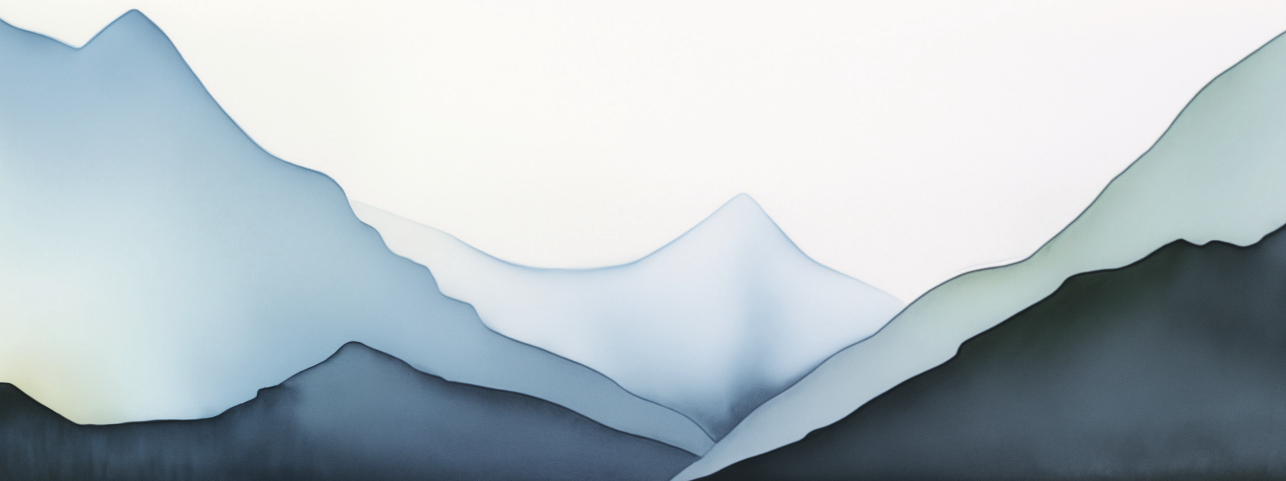
The LPFS-BF 2.0 showed high sensitivity to change, reflected by a high effect size after three months of inpatient treatment. The LPFS-BF 2.0 was as sensitive (total score) or more sensitive (self functioning domain) to change than the BSI and more sensitive to change than most SIPP-SF domains, providing preliminary evidence that, at least from a perspective of sensitivity to change, the LPFS-BF 2.0 may serve as a Routine Outcome Monitoring instrument in patients with personality disorders. That said, it warrants mentioning that many other conditions need to be met before such use can become good practice (see for example van Os et al., 2012). Notably, the LPFS-BF 2.0 showed sensitivity to change commensurate with the BSI (measuring symptom distress). As personality syndromes are generally presumed to be more stable than symptom syndromes, this finding warrants further study. It may be explained by the efficacy or the treatment program focussing on personality problems, but it may also point to shared variance with symptom distress. Intercorrelations between the measures were moderate, with the Interpersonal facet of the LPFS being more distinct. Future studies should also include the sustainability of the changes after treatment, by including follow-up assessments of level of personality functioning, and assessing the presence of PD diagnoses after treatment has been completed.

Some limitations should be kept in mind, most notably the restricted range of personality disorder types within this sample, with virtually no Cluster A PD present, and very few antisocial PD, dependent PD and histrionic PD, along with a predominance of borderline PD, PD-NOS and avoidant PD diagnoses. To mitigate this concern, this composition is consistent with most reported research in non forensic mixed samples of PD (Berghuis et al., 2012). Secondly, although there

were multiple instances of the SIPP-SF domains showing discriminant relationships with the LPFS-BF 2.0 Self- and Interpersonal Functioning domains, this was not the case for three of the five domains. Several factors may account for this lack of discriminative associations. First, the concepts operationalized by both the LPFS-BF 2.0 and the SIPP are clinically "rich" concepts, which tends to complicate the inherent trade-off between coverage and clarity of factor structure. Second, it is also possible that the brevity of the LPFS-BF 2.0 limits its ability to discriminate between Interpersonal and Self Functioning. The longer LPFS-sr, for example, showed higher correlations with the SIPP domains and other questionnaires measuring personality functioning (Morey, 2017). Also, compared to the original LPFS-BF, we deleted all reversed items to reflect the notion that absence of health does not necessarily equal the presence of pathology. However, this arguably renders the questionnaire more vulnerable to response tendencies, like acquiescence bias. Of course, the absence of a (virtual) gold standard for assessing the level of personality functioning remains a limitation that plagues all research in this domain. For future research, we recommend comparing the psychometric performance of the present short LPFS-BF 2.0 to more full assessment measures of the LPFS (e.g. the LPFS-sr; Morey, 2017). Also future research could assess the sensitivity and specificity of the LPFS-BF 2.0 for distinguishing patients with and without a PD in a more heterogenous sample. Taken together, we suggest this study demonstrates the potential value of the LPFS-BF 2.0 as a brief instrument that may also serve to assess individual change in personality functioning during treatment as a complement to assessing symptom reduction.

Acknowledgements

We would like to thank Carlijn Wibbelink of the University of Amsterdam, for her help with conducting the Confirmatory Factor Analyses in *M*plus.



Chapter 3

Assessing Criterion A in adolescents using the Semistructured Interview for Personality Functioning DSM – 5

Laura C. Weekers, Susanne C.E. Verhoeff, Jan H. Kamphuis, Joost Hutsebaut

Personality Disorders: Theory, Research, and Treatment, (2021), 12, 312-319

Abstract

Accumulating evidence supports the reliability and validity of the diagnosis of personality disorders (PDs) in adolescents, but whether the current DSM-5 criteria are optimal to capture and help detect emerging PDs in this age group remains controversial. The Level of Personality Functioning Scale (LPFS) may provide a more developmentally sensitive way to identify impaired personality features in young people. This study investigates the feasibility of the LPFS in adolescents by examining the psychometric properties of the Semistructured Interview for DSM-5 Personality functioning (STiP-5.1) in a clinical sample of referred adolescents (N = 84) and in a community sample (N = 12). Additionally, referred adolescents completed self-report questionnaires pertaining to symptom severity, personality functioning and personality traits. In general, good inter-rater reliability and internal consistency were observed, and the associations with external variables largely followed theoretical prediction. Interestingly, and in contrast to data on adults, we found no significant associations between the LPFS-scores on the one hand and traditional DSM-5 PD diagnoses in the clinical sample on the other (except for borderline PD criteria). In discussing these findings, we argue that the assessment of personality functioning may be better suited for detecting personality pathology in adolescence than the traditional Section II criteria.

Introduction

A growing body of research has provided evidence that PDs are common in adolescents (e.g. Feenstra et al., 2011; Grilo et al., 1998; Johnson et al., 2000; Westen et al., 2003). Collectively supporting the validity of the diagnosis in this age group, research has shown that adolescents with putative PDs are at a greater risk of having a broad range of problems than adolescents without PDs, including problems at school (Westen et al, 2003), behavioral problems (Johnson et al., 2005), interpersonal difficulties and stress (Daley et al., 2006), substance abuse (Serman et al., 2002), suicide attempts (Braun-Scharm, 1996; Westen et al., 2003), emergency admissions (Kasen et al., 2007), and deviant sexual behavior (Lavan & Johnsons, 2002). These findings warrant for early detection, setting the stage for the recent development of early intervention programs aiming to tackle PD problems (Chanen & McCutcheon, 2008; Hutsebaut et al., 2019).

Discussion continues whether the current DSM-5 criteria are optimal to assess personality pathology in the adolescent population. Indeed, many manifestations of PDs in DSM-5 seem to refer to adult roles and symptoms (Videler et al., 2019). Examples of such adult-oriented criteria include several criteria for dependent PD; e.g. "needing others to assume responsibility for most major areas of life" (Criterion 2), or "seeking another relationship as a source of care and support when a close relationship ends" (Criterion 7). Similarly, Criterion 1 of borderline PD (BPD) referring to "frantic efforts to avoid real or imagined abandonment" may typically apply to a late-adolescent or adultlike expression of fear of abandonment. There is also empirical evidence to suggest that current PD criteria are not completely age-neutral, such that symptoms in the areas of affect dysregulation and impulse dyscontrol seem to manifest early in the course of the disorder, while symptoms of interpersonal disturbance only seem to be expressed later on (Debast et al., 2017; Sharp et al., 2015). These conjectures and observations call into question the validity of the

current diagnostic criteria to detect PD expressions across the lifespan. Assessment of personality pathology could therefore benefit from more developmentally sensitive descriptors of personality dysfunction.

The fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association, 2013) features in its Section III an Alternative Model for Personality Disorders (AMPD), which represents an interesting alternative for assessing personality pathology in adolescents. Central to this newly proposed classification of PDs is an assessment of the level of impairment in personality functioning (Criterion A). Impairments constituting personality pathology are assumed to manifest in self- and/ or interpersonal functioning (Bender et al., 2011). In order to assess these impairments, DSM-5 has introduced the Level of Personality Functioning Scale (LPFS). This scale uses 12 capacities of self- and interpersonal functioning to differentiate between five levels of severity of personality pathology, ranging from little or no impairment (Level 0) to extreme impairment (Level 4). The model reflects the viewpoint that these psychological capacities, much like maladaptive traits (De Clercq et al., 2014), develop over the lifespan and express themselves differently according to the developmental phase, paving the way for a lifespan perspective on PDs (Hutsebaut et al., 2019). As such, personality pathology is defined in terms of impairments in personality functioning, rather than in terms of their (age-related) behavioral manifestations (as in the DSM-5 Section II model). The AMPD model has the potential to improve the assessment of personality pathology in adolescents, as it relates severity of personality pathology directly to personality processes, independent of behavioral manifestations, social and vocational outcomes, or experienced burden of disease. Finally, by adopting a dimensional conceptualization, the AMPD implicitly recognizes that normal and pathological development are not qualitatively distinct, and in so doing also posits a gradient model of healthy personality development (Hutsebaut et al., 2017). Dimensionality may also contribute to early detection when the PD has not yet fully developed. On the other hand, one may also argue that

many of the concepts of the LPFS, including intimacy and empathy, may refer to concepts of personality development that are still "under construction". Such concepts may make it difficult to distinguish normative shortcomings in areas of personality development from expressions of "true" personality dysfunction, which is predictive for actual and future social and mental problems.

Emerging evidence shows that aspects of personality functioning may differentiate well between "normal" and "clinical" adolescents. For example, one study found that maladaptive aspects related to the LPFS, including identity integration and relational capacities as assessed by the Severity Indices of Personality Problems (SIPP-118), were more strongly related to clinical status than to age and these dimensions appeared to improve during treatment, supporting the notion that they capture (personality) pathology rather than developmental issues (Feenstra et al., 2014). More closely related to the AMPD, Goth and colleagues (2018) developed the Level of Personality Functioning Questionnaire (LoPF-Q) as an adaptation of their AIDA (Assessment of Identity Development in Adolescence) questionnaire. For both instruments they showed good capacity to differentiate clinical adolescents from community adolescents (d = 0.7 -2.2). To our knowledge, to date only Goth and colleagues (2018) used specifically tailored AMPD-instruments to study Criterion A in young persons.

This study uses the Semi-Structured Interview for DSM-5 Personality functioning (STiP-5.1) to assess the LPFS in adolescents and tested its ability to distinguish normative developmental phenomena from psychopathology. The STiP-5.1 was specifically designed as a multi-item assessment of each of the 12 capacities of the LPFS (Hutsebaut et al., 2014). It has shown promising results in adults (Hutsebaut et al., 2017; Zettl et al., 2019). We investigated aspects of the reliability of the STiP-5.1 to assess severity of personality pathology in adolescents, expecting good inter-rater reliability and internal consistency. Furthermore, we investigated aspects of construct validity by studying the capacity of the STiP-5.1 to differentiate between clinical and

community adolescents and by calculating the associations with theoretically relevant measures of personality pathology.

Method

Participants

Both a clinical and a community sample were recruited. Participants in the clinical sample (N = 84) were treatment seeking adolescents, referred to a mental health care center specialized in the assessment and treatment of adolescents and adults with personality-, conduct-, or family problems. Their age ranged from 12 to 17 (M = 15.60, SD = 1.39) and 89.3% were female. Participants lived with both parents (45.2%), with one of their parents (28.6%), in a newly formed family (10.7%), with foster parents (8.3%), or in an institution (6.0%). Information on educational level was available for only half of the sample. Of these participants 11.9% had attained a low educational level, 40.4% an intermediate-, and 23.8% a high educational level.

In the community sample (N = 12), participants' age ranged from 13 to 17 (M = 15.08, SD = 1.16), and 75% were female. Almost all participants lived with both parents (91.7%). All had intermediate (8.3%) or higher educational (91.7%) levels. No participants were in treatment or had ever received individual or group psychotherapy. Comparisons on demographic variables (age, sex, educational level) showed no other significant differences between the two samples.

Procedure

In addition to the standard admission procedure, which included administration of semi-structured interviews for the assessment of DSM-IV-TR Axis I and Axis II disorders and of selected self-report questionnaires (SIPP-SF, BSI, LPFS-BF 2.0 and PID-5 BF; see measures), all adolescents in the clinical sample were administered the STiP-5.1 interview. The STiP-5.1 was administered after the adolescents and their parents had received an initial consultation with a clinician. The

STiP-5.1 was integrated in the standard admission procedure and administered by a psychologist who was only given the name and age of the participant. The interviewer asked permission to videotape the interview and obtained informed consent from the adolescent and his/ her parents to use the recording for scientific purposes, including re-scoring of the interview by an independent rater. The second rater, who was equally uninformed concerning the adolescent's personal and clinical background, scored the LPFS independently based upon the videotaped interview.

Based upon previous effect sizes (Hutsebaut et al., 2017), a desired level of power of $\beta = .80$, and using a false positive rate of $\alpha = .05$, we included 12 youngsters in the community sample to study the STiP-5.1's ability to differentiate between clinical and community youngsters. The participants from the community sample were recruited through a call among schools, sports clubs, and relatives of personnel working at the treatment center. We asked mentors/ teachers of different classes (1st – 6th grade) and different levels of education to inform their students about the study and to ask for volunteers. In order to match the community sample as much as possible we asked them to recruit adolescents from specific ages and educational levels for our study.

Participants from the community sample were contacted by the interviewer and were administered the STiP 5.1. They were also asked permission to videotape the interview and the interviewer obtained informed consent from the adolescent and his/her parents. Additional sociodemographic information (age, level of education, living situation) was collected. No additional diagnostic interviews or self-report questionnaires were administered. All participants were interviewed in the same or a similar consultation room to make second raters as blind as possible to clinical status, preventing bias.

Measures

Semi-Structured Interview for Personality Functioning DSM-5 (STiP-5.1). The STiP-5.1 (Hutsebaut et al., 2014) is an interview schedule assessing the level of personality functioning as operationalized by the LPFS in Section III of the DSM-5. Its format is sufficiently user-friendly that after only a brief training, clinicians without specialized experience are able to competently administer it (Hutsebaut et al., 2017). The diagram of the interview, which is organized around the capacities of the LPFS, is divided into three columns. The left column features the criteria for the different levels of each of the capacities. In the outer right column, the aspects of information that should be collected in order to rate the different LPF levels are described. Specific questions that should be posed to the patient are displayed in the middle column. Sixty descriptors of severity are encompassed in the LPFS, divided into 12 facets (capacities) with each 5 levels. A "funnel" strategy is applied during the interview, which allows the interviewer to narrow down to the level of impairment through the questioning sequence, instead of having to check each of the sixty descriptors separately (Hutsebaut et al., 2016). The interview consists of 28 open questions, with optional clarifying questions. A broad open question is used as the start of each section of the interview. Auxiliary questions may be used, contingent on the prior answer to the starting question, to subsequently focus on the remaining levels. Reframing the respondents' information in terms that correspond with the exact description of the level in the LPFS may be used as an additional strategy to check the assumed level of impairment. Ratings of each capacity should be performed during the interview, interviewers are encouraged to give one score per capacity ranging from 0 (*little or no impairment*) to 4 (*extreme impairment*). The average interview duration is about 45 minutes. Internal consistency of the STiP-5.1 was high in an adult sample, with a Cronbach's alpha of .97 for the total scale, and .94 for both the Self-functioning and Interpersonal functioning subdomain. Interrater reliability was good, with intraclass correlation coefficients (*ICCs*) ranging from .81 to .92 in

the total sample, and *ICCs* ranging from .58 to .81 in the clinical sample (Hutsebaut et al., 2017).

Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I)
The SCID-I (First et al., 1997; translated by Groenestijn et al., 1999) is a semistructured interview to measure *DSM–IV* Axis I disorders. The SCID-I has demonstrated good interrater reliability in various samples, especially when interviewers had received a formal training; overall κ = .85 (Ventura et al., 1998).

Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II)
The SCID II (First et al., 1997, translated by Weertman et al., 1996), which is essentially identical to the current SCID-P (DSM-5), was used to diagnose Axis II PDs. Criteria were scored if they were pathological, pervasive, and persistent. PD not otherwise specified was classified when five PD criteria were present (Verheul et al., 2007). The SCID-II has good interrater and test–retest reliability in PD samples (see, e.g. Maffei et al., 1997; Weertman et al., 2003) with sum *ICCs* of .90 for avoidant and .95 for borderline PDs (Lobbestael et al., 2011).

Severity Indices of Personality Problems – Short Form (SIPP-SF)
The SIPP-SF (Feenstra et al., 2011; Verheul et al. 2008) is a dimensional self-report measure assessing the generic and changeable components of personality functioning. It consists of 60 items, all rated on a 4-point Likert scale ranging from 1 (*fully disagree*) to 4 (*fully agree*). Respondents are asked to answer to what extent they agree with the statements, referring to the last 3 months. The SIPP-SF comprises five higher-order domains: Self-Control, Identity Integration, Responsibility, Relational Functioning, and Social Concordance. High scores (sum scores) on the facets indicate more adaptive personality functioning, whereas lower scores suggest more maladaptive functioning. The SIPP-SF is the shortened version of the SIPP-118, which has good psychometric features in both adults and adolescents. Internal

consistency in the current sample was high, with Cronbach's α's of .87 (Self-control), .93 (Identity Integration), .83 (Responsibility), .85 (Relational Capacities), and .85 (Social Concordance), respectively.

Brief Symptom Inventory (BSI)

The BSI (Derogatis, 1975; translated by de Beurs, 2011) is used to assess symptom severity. It consists of 53 items covering nine symptom dimensions, but the present study uses the total score (sum score) that provides an index of the intensity of distress by psychological symptoms during the past week. Respondents rank each item on a 5-point scale ranging from 0 (*not at all*) to 4 (*extremely*). Internal consistency in the current sample was high, with Cronbach's α = .96 for the total score.

Level of Personality Functioning Scale–Brief Form (LPFS-BF 2.0)

The LPFS-BF 2.0 (Bach & Hutsebaut, 2018; Weekers et al., 2019) is a brief self-report questionnaire for assessing the LPFS as described in Section III of the DSM–5 (APA, 2013). It consists of 12 items corresponding to the 12 capacities of the LPFS. Participants are asked to rate the 12 items on a 4-point Likert scale from 1 (*completely untrue*) to 4 (*completely true*). Both a total score and subdomain scores (Self-functioning and Interpersonal functioning) can be calculated. Internal consistency in the current sample, as measured by Cronbach's α, was .64 for the LPFS-BF total scale, and .73 and .58 for the Self and Interpersonal subscales respectively.

Personality Inventory for the DSM-5 Brief Form (PID-5 BF)

The PID-5-BF (APA, 2013b; Dutch version: van der Heijden et al., 2014) describes 25 trait facets organized in five higher-order domains: Negative Affectivity (vs. Emotional Stability), Detachment (vs. Extraversion), Antagonism (vs. Agreeableness), Disinhibition (vs Conscientiousness), and Psychoticism (vs. Lucidity). The PID-5-BF measures the DSM-5 trait domains using a total of 25 items (five per domain), computed following the APA guidelines. Items are measured

on 4-point Likert scales. The PID-5 BF has been validated in a sample of adolescents (Koster et al., 2019). Cronbach α's in the current sample ranged from .64 (Detachment) to .73 (Antagonism).

Table 1 *Prevalence of DSM diagnoses in the Clinical Sample (N = 64-83)*

	N	*%*
Syndrome disorders		
Anxiety Disorders	35	45.5
Mood Disorders	49	75.4
Somatization Disorders	7	8.8
Eating Disorders	20	25.3
Substance use Disorders	13	17.6
Conduct Disorder	5	5.2
Oppositional-Defiant Disorder	2	2.5
Attention-Deficit/Hyperactivity Disorder	6	7.4
Autism Spectrum Disorder	3	3.6
Any Axis I diagnosis	69	92
Personality disorders		
Avoidant PD	17	20.7
Obsessive-compulsive PD	3	3.6
Borderline PD	23	27.4
PD not otherwise specified	22	26.2
Any PD	53	63.1

Results

Clinical characteristics of the sample

Table 1 provides an overview of diagnostical information. The majority of participants met criteria for at least one axis I disorder (92%), with mood disorders (75.4%) and anxiety disorders (45.5%) being the most prevalent. Fifty-three (63.1%) of the adolescents from the clinical

sample had a PD whereas the remaining 31 adolescents did not meet criteria for a PD diagnosis (25 of whom met at least 1 PD criterium). BPD (27.4%) and PD not otherwise specified (26.2%) were most prevalent.

Reliability

Interrater reliability was assessed using a one-way random, absolute agreement, single-measures *ICC* (McGraw & Wong, 1996). Twenty-six interviews were scored by a second rater, (16 and 10 interviews of the clinical sample and community sample, respectively). Internal consistency of the STiP 5.1 was high, with α = .96 for the Total score and α= .94 and α = .92 for the Self- and Interpersonal functioning subdomains, respectively. Interrater reliability was high for the total sample, with *ICCs* ranging from .88 to .99 (see Table 2). For the clinical sample, interrater reliability was acceptable to good, with one exception for "experience of oneself as unique" (*ICC* = .47). *ICCs* for the remaining Self-functioning capacities ranged from .57 to .96 and *ICCs* for the Interpersonal functioning subdomain ranged from .73 to .97.

Table 2 *Inter-rater Reliability: Intraclass correlation coefficients of STiP 5.1*

Scale	Clinical (*N* = 16)	Total (*N* = 26)
STiP-5.1 total score	.69	.95
Domain Self-functioning	.57	.95
Identity	.65	.96
Experience of oneself as unique	.47	.92
Self-esteem	.96	.99
Emotions	.80	.98
Self-direction	.79	.93
Goals	.76	.94
Norms	.76	.91
Self-reflection	.76	.88

Table 2 *Inter-rater Reliability: Intraclass correlation coefficients of STiP 5.1* (continued)

Scale	Clinical (*N* = 16)	Total (*N* = 26)
Domain Interpersonal functioning	.75	.96
Empathy	.92	.98
Understanding others	.92	.95
Perspectives	.73	.91
Impact	.93	.98
Intimacy	.85	.97
Connection	.90	.99
Closeness	.79	.95
Mutuality	.97	.99

Note. STiP-5.1 = Semistructured Interview for Personality Functioning DSM–5; DSM–5 = Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition.

Construct validity

Table 3 shows associations of the STiP 5.1 with SCID-I and SCID-II indices. No significant correlations were observed between the STiP 5.1 on the one hand, and the number of PDs, number of PD criteria, or number of axis I disorders on the other. However, the number of BPD criteria was significantly associated with the STiP 5.1 total score as well as its subscales.

As expected, self-report measures of personality problems generally showed moderate positive associations with the STiP 5.1, with higher levels of self-reported personality problems being associated with more severe levels of personality functioning. More specifically, the STiP 5.1 was positively associated with the LPFS-BF 2.0 and most of the SIPP-SF and PID-5-BF domains (see Table 4 for correlation coefficients). The SIPP-SF identity integration domain, the PID-5 Negative Affectivity and Antagonism domain were not significantly associated with the STiP 5.1. Symptom severity as measured by the BSI was significantly related to level of personality functioning as measured by the STiP 5.1; particularly the Self-functioning subdomain but not to the Interpersonal functioning subdomain. For means and

standard deviations of the STiP 5.1 and the self-report questionnaires, see supplemental material.

Independent samples *t*-tests showed the community sample had significantly healthier STiP 5.1 scores (M = 0, SD = 0) than the clinical sample (M = 2.61, SD = 0.60; t (83) = -39.72 , p < .001, d = 4.68), with a very large effect-size. Subsequently, in the clinical sample there was no significant difference on STiP 5.1 scores between adolescents with a PD diagnosis (M = 2.66, SD = 0.59) and adolescents without a PD diagnosis (M = 2.52, SD = .63; t (82) = -1.06, p = .292).

Table 3 *Correlations of STiP-5.1 Scores (N = 77-84[a]) with Axis I and PD disorders*

	STiP-5.1 Total	Self	Interpersonal
STiP-5.1 Self-functioning	.78***		
STiP-5.1 Interpersonal functioning	.84***	.66***	
SCID-II Number of PDs	.19	.14	.11
SCID-II Number of PD criteria	.19	.17	.15
SCID-II Number of avoidant PD criteria	.06	.08	.03
SCID-II Number of borderline PD criteria	.38**	.29**	.36**
SCID-I Number of Axis I diagnoses	.16	.01	.17

Note: *** Correlation is significant at the .001 level; ** Correlation is significant at the .01 level; [a] *n* varies due to missing values.

Table 4 *Pearson Correlations of STiP-5.1 with Self-report Measures of Personality problems and Symptom severity in the clinical sample (N = 59-64)*

	STiP-5.1 Total score	Self functioning	Interpersonal functioning
LPFS-BF 2.0 Total score	.49***	.46***	.43***
LPFS-BF 2.0 Self	.21	.29*	.14
LPFS-BF 2.0 Interpersonal	.57***	.43***	.56***
SIPP-SF Self-control	-.46***	-.46***	-.39**
SIPP-SF Identity integration	-.19	-.20	-.13
SIPP-SF Responsibility	-.32*	-.26*	-.35**

Table 4 *Pearson Correlations of STiP-5.1 with Self-report Measures of Personality problems and Symptom severity in the clinical sample (N = 59-64)* (continued)

	STiP-5.1 Total score	Self functioning	Interpersonal functioning
SIPP-SF Relational capacities	-.33**	-.16	-.32*
SIPP-SF Social concordance	-.53***	-.46***	-.44***
PID-5-BF Total score	.44***	.44***	.46***
PID-5-BF Negative affectivity	.19	.28*	.19
PID-5-BF Detachment	.26*	.12	.30*
PID-5-BF Disinhibition	.45***	.50***	.46***
PID-5-BF Antagonism	.21	.17	.25*
PID-5-BF Psychoticism	.27*	.29*	.26*
BSI total score	.25*	.31**	.17

Note. LPFS-BF 2.0: Level of Personality Functioning Scale – Brief Form 2.0; SIPP-SF: Severity Indices for Personality Problems – Short Form; PID-5-BF: Personality Inventory for DSM-5 – Brief Form; *** Correlation is significant at the .001 level; ** Correlation is significant at the .01 level; * Correlation is significant at the .05 level.

Discussion

This study investigated the potential utility of the LPFS in adolescents by exploring reliability and validity of the STiP-5.1 in this age group. We found moderate to excellent interrater reliability in the clinical sample and excellent interrater reliability in the total sample, supporting the reliability of assessment of personality functioning using the STiP-5.1. An exception was observed for "unique sense of self" with a comparable lower *ICC* in the clinical sample. Construct validity was supported by the instrument's ability to differentiate clinically referred from community youngsters, and by a theoretically meaningful pattern of associations with related constructs of personality pathology. More specifically, we found significant correlations between the STiP-5.1 scores and the majority of self-report measures of personality pathology, including total and scale scores of the LPFS-BF-2.0, SIPP-SF and PID-5-BF. Higher scores on these personality functioning

and -trait measures (indicating greater dysfunction) covaried with higher levels of impairments in personality functioning (STiP 5.1). No associations were found between Negative affectivity and Antagonism (PID-5-BF) and the STiP 5.1 scores in the current sample. Restriction of range (associated with high levels of Negative affectivity and low levels of Antagonism) may account for this null finding. Furthermore, the STiP 5.1 clearly distinguished healthy adolescents from adolescents in the clinical sample. Impairments in personality functioning are indeed distinguishable from normal adolescent struggles.

Of note, as opposed to findings observed in an adult sample in our clinic (Hutsebaut et al., 2017), the STiP 5.1 was not associated with (Section II informed) traditional PDs. Moreover, the STiP 5.1 did not differentiate between adolescents with and without a (full) Section-II PD diagnosis in the clinical sample. An exception was observed for features of BPD, such that adolescents who displayed more features of BPD, were also rated as more disturbed on level of personality functioning using the STiP-5.1.

Arguably, these results suggest that the STiP 5.1 does not adequately capture the severity of personality pathology in younger populations or even that the LPFS does not capture the common core of PDs in adolescents. However, we deem this explanation less plausible, as the STiP 5.1 did show theoretically consistent associations with validated self-report questionnaires assessing personality pathology. As an alternative explanation, we would argue that the Section II and III assessment approaches capture different aspects of personality pathology. Whereas in adults those different aspects generally converged strongly (Hutsebaut et al., 2017), meaning that impairments in personality functioning co-occur with classic diagnostic criteria of PD diagnoses, in adolescents these aspects were only loosely connected, if at all. Given the highly impaired and severe LPFS-scores in this study we propose that severe impairments as assessed by the STiP-5.1 do not necessarily express or manifest themselves in traditional DSM-5 criteria of PDs in youngsters. This discrepancy may be related to the reliance on formal diagnostic criteria whose behavioral manifestations

are more prevalent in adults than in adolescents and that in fact may be more representative of adult personality impairments in personality functioning (e.g. avoiding social situations, being dependent on another adult, impulsive drug and alcohol use; Videler et al., 2019).

An implication of this hypothesis would be that an approach towards the assessment of personality functioning, e.g. through the STiP-5.1, may be better suited to detect severity of personality pathology at a young age and may be more informative for planning treatment than an exclusive DSM-5 Section-II based approach. While the SCID-5-P relies heavily on behavioral manifestations of PD, the STiP 5.1 allows the clinician to assess core aspects of personality functioning and adjust questions and severity ratings to the developmental level of the adolescent; for example by taking into account that interpersonal impairments as described by the LPFS-criteria, may express themselves differently at 14 or 28 years of age. Further studies should investigate the predictive value of both Section II and III criteria for early identification of youngsters at risk for developing full, chronic, and severe personality pathology.

Features of BPD were moderately associated with impaired personality functioning as assessed by the STiP-5.1, and as such constituted an exception to the non-relation between STiP 5.1 assessed LPFS and Section II PDs. These findings may support the notion that features of BPD represent a rather generic marker for severity of personality pathology (Sharp et al., 2015). Moreover, these findings provide suggestive evidence that the core components of PDs in general are first expressed in features of BPD (Chanen & McCutcheon, 2013). Such an hypothesis aligns well with the notion of clinical staging, with the core vulnerability in personality development expressing itself primarily in affective and impulsive dysregulations, captured by some of the criteria of BPD (Hutsebaut et al., 2019). Other section-II PD-criteria may only emerge later in the course of the disorder, either because the criteria are not age-neutrally formulated or because these problems only arise later in life. An implication may be that BPD features are more sensitive than other PD criteria to detect

personality impairment earlier in life, but these conjectures are in need of empirical testing.

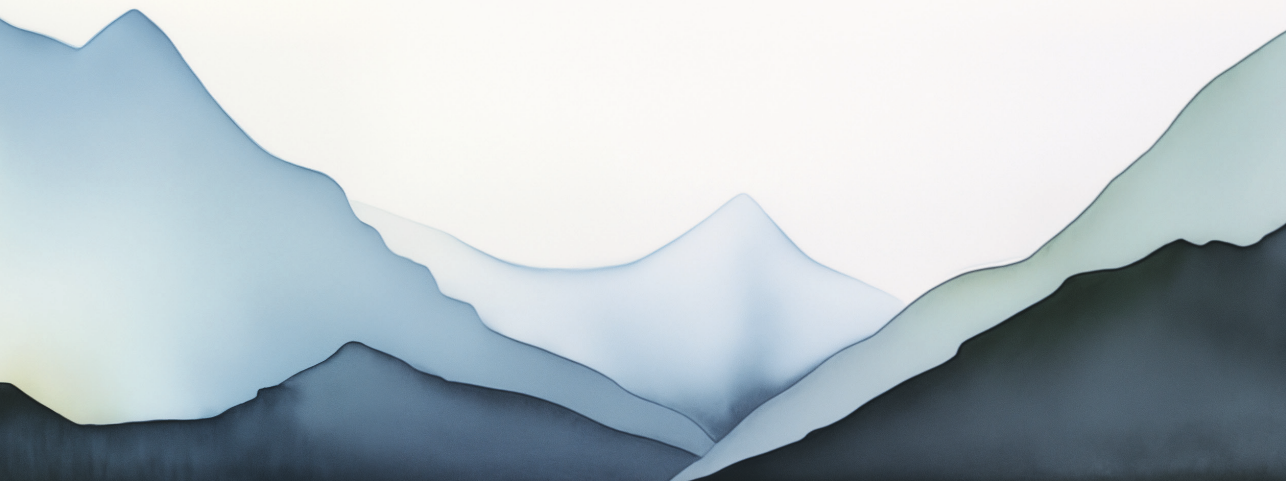
Of note, interrater reliability differed across subdomains and capacities. Most capacities seem to be easy to score reliably, e.g. self-esteem or mutuality of regard. One exception seems to be "unique sense of self", with a remarkably lower *ICC* (.47), probably explaining the somewhat reduced *ICC* for the Self-functioning subdomain too. The emergence of a unique sense of self throughout adolescence may be affected by normal developmental struggles, apparently troubling clinicians whether to consider the answers of the young person reflective of true personality pathology or (relatively) within the normal developmental range. This might also explain null findings for associations between STiP severity scores and self-reported identity integration (SIPP-SF). It seems that different clinicians may use the LPFS-criteria somewhat differently in order to assess the level of personality pathology given this dilemma. As an implication, it may be that the LPFS-criteria might be enhanced by adding developmentally specific criteria that could assist clinicians in making these decisions.

A number of strengths and limitations of this study deserve mention. It is the first interview-based study of the level of personality functioning in adolescents. Interviews were integrated in the regular intake procedure, supporting ecological validity of the study and demonstrating usefulness of the STiP-5.1 in regular clinical practice. However, several limitations should be mentioned. First, due to the design of the study, the interviewers were not blind to the clinical status of the adolescents. This may be a potential source of bias. Second, as in most clinical samples, some types of personality pathology (i.e. borderline / avoidant PDs) are more prevalent than others (i.e. Cluster A and antisocial PDs). Although this clinical sample may be representative for a severe and complex personality disordered clinical sample, it does not cover the whole range of personality pathology and, for example, it remains questionable whether the STiP-5.1 may be useful in samples of antisocial youngsters too. Indeed, in a study by Bach and Hutsebaut (2018), incarcerated adults reported healthier levels

of personality functioning than outpatients, calling into question the validity of the LPFS self-report instrument in antisocial samples. Third, this study mainly draws on self-report and does not include informant-based assessments of personality pathology. A multi-informant approach for assessing personality pathology is recommended, particularly for youngsters who may have some introspective and motivational limitations (Shiner & Allen, 2013; Weekers et al., 2020). Moreover, we recommend future studies include developmentally sensitive, age-specific measures such as the LoPF-Q (Goth et al., 2018) to assess construct validity of the STiP 5.1. Furthermore, for the purpose of this study – focusing on the psychometric qualities of the STiP-5.1 interview schedule – except for clinical status, interviewers were kept uninformed of any other information (e.g. living situation, treatment history, reasons for seeking help), in order not to bias them towards certain levels of severity. However, in clinical practice, it would be recommended to include all sources and types of information in order to make a comprehensive and valid assessment of someone's personality pathology (Weekers et al., 2019). Also, although our a priori power analysis deemed the sample size of the community sample large enough to detect a large effect (as was also observed in adults; Hutsebaut et al., 2017), we acknowledge that this comparison group was small and invariably in good mental health, which limits its representativeness and generalizability. Moreover, the SCID-P was unfortunately not administered to the community sample, which precludes testing to what extent Section II criteria would also have discriminant ability with respect to the clinical and community samples. Finally, possibly due to the tertiary, specialist nature of the setting, we included mostly high-end severity cases of adolescent personality pathology. Future studies may include a more heterogeneous sample of adolescents.

In sum, this study provided (preliminary) support for the reliability and validity of the STiP 5.1 in adolescents. Additionally, it lends support to the use of the LPFS to detect personality pathology at an early stage and accordingly may provide a framework for detecting

young people at risk, even before their vulnerability to personality pathology is expressed in classic symptoms of PD. We recommend further research be aimed at cross-validating the added value of the LPFS to detect adolescent personality pathology in a developmentally sensitive way.



Chapter 4

Changes in the classification of personality disorders: Comparing the DSM–5 Section II personality disorder model to the Alternative Model for Personality Disorders using structured clinical interviews

Laura C. Weekers, Joost Hutsebaut, Johannes Zimmermann, Jan H. Kamphuis

Personality Disorders: Theory, Research, and Treatment (2021), 13, 527-535

Abstract

The current study examined the continuity of personality disorder (PD) diagnoses from Section II to Section III (alternative model for personality disorders [AMPD]) when using structured interviews. We investigated the continuity both in terms of stability of prevalence rates and in terms of convergent validity. A clinical sample of 189 participants were concurrently administered both Section II PD and AMPD interviews for diagnosing PD by 2 independent interviewers. Stability of prevalence between the models for specific PD diagnoses was generally supported. A higher prevalence of trait-specified PD in the AMPD model resulted in higher prevalence of PD in general when using the AMPD model compared with the Section II PD model. Correlations between matching criterion counts according to both models were generally high. Convergence between the Section II PD and AMPD model categorical diagnoses was adequate for the most frequently diagnosed and studied PDs (i.e., avoidant-, borderline-, and antisocial PD), but lower than previously found, likely due to the stringent test–retest design used in this study. Convergence between the models for narcissistic and obsessive–compulsive PD was low and could not be estimated for schizotypal PD. Future studies should investigate which of both models may prove to be most valid in terms of predicting current and future impairments.

Introduction

The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM–5), introduced a new, alternative model in Section III "emerging measures and models" (American Psychiatric Association, 2013). The alternative model for personality disorders (AMPD) was developed to enhance clinical utility and address a number of limitations of the current Section II model, including arbitrary diagnostic thresholds (Balsis et al., 2011), extensive co-occurrence of personality disorders (PDs; Grant et al., 2005), heterogeneity among patients receiving the same diagnosis (Krueger et al., 2014), and low validity (Waugh et al., 2017). Presenting a hybrid dimensional–categorical classification system, the AMPD aims to improve validity and clinical utility of the diagnosis of PDs (Bach et al., 2015) and bring the assessment of PDs more in line with evidence-based models of PD (Waugh et al., 2017). The AMPD defines PDs in terms of impairments in self and interpersonal functioning (Criterion A), and the presence of maladaptive personality traits (Criterion B; American Psychiatric Association, 2013). Although both criteria are dimensional, the AMPD still enables categorical PD diagnoses. In particular, out of the 10 Section II types, the AMPD retained six specific types and newly defined them as specific combinations of Criteria A and B descriptors. In addition, the AMPD included a "trait-specified" PD for combinations that do not meet criteria of a specific PD. The current study was designed to assess the continuity of PD diagnoses from Section II to Section III when using the specifically developed structured interviews associated with each model. We investigated the continuity both in terms of stability of prevalence rates and in terms of convergent validity.

Although the AMPD was initially proposed to replace the prevailing classification system, the APA Board of Trustees refused to include the AMPD in Section II (Diagnostic Criteria and Codes) as the leading classification system (American Psychiatric Association, 2013). The AMPD was argued to represent too radical a change for the existing

clinical practice and to be lacking in empirical evidence. Members of the field feared that the new assessment system would lead to a large discontinuity with past studies, causing existing research on the treatment of PDs to be of only limited value (Clarkin & Huprich, 2011; Frances, 2012). Since DSM–5 was published, however, a growing number of studies have mitigated these concerns. Several studies suggest adequate to high convergence between borderline PD (Bach & Sellbom, 2016; Evans & Simms, 2018; Sellbom et al., 2014), obsessive–compulsive PD (Liggett & Sellbom, 2018; Liggett et al., 2017), antisocial PD (Wygant et al., 2016), narcissistic PD (Fossati et al., 2017; Wright et al., 2013), avoidant PD (Sellbom et al., 2017), and schizotypal PD (Somma et al., 2019) and their corresponding Section III maladaptive traits (Watters et al., 2019; Zimmermann et al., 2019). Furthermore, Morey et al. (2013) demonstrated that a moderate or higher score on Criterion A had a specificity of 72.2% and a sensitivity of 84.6% for identifying patients that had at least one PD according to the Section II PD model. Furthermore, Morey and Skodol (2013) reported an average convergent validity of $\kappa = .54$ between PD diagnoses from both models, which is similar to the consensus from the transition of the DSM-III to DSM-III-Revised. In a more recent study, Morey (2019) demonstrated a high level of convergence ($\kappa = .51$) of borderline PD diagnosis between Section II and III when using case vignettes. Furthermore, convergence between Section II and Section III was demonstrated for obsessive–compulsive PD (Liggett et al., 2017) and avoidant PD (Sellbom et al., 2017) in community samples, when using self-report.

The changes in (specific) PD diagnoses might therefore be less radical then initially thought. However, these studies have several limitations. First, most studies have focused on associations between Criterion B traits and Section II PDs and only a few studies investigated associations between Criterion A impairment and Section II PDs. Studies comparing the full AMPD diagnoses with the Section II PD diagnoses are even more scarce. Second, most studies that did compare the full models focused on convergence (i.e., correlations)

between both models (Liggett & Sellbom, 2018; Liggett et al., 2017; Sellbom et al., 2017) but not on stability of the prevalence of respective PD diagnoses. This is problematic, as even when Section II PD and AMPD traits might be highly correlated, diagnostic thresholds may differ, resulting in shifts in prevalence rates of specific diagnoses. An exception is the study by Morey (2019), but the reported estimates were based on case vignettes and restricted to borderline PD. Comparison of the models in terms of stability of prevalence of PD diagnoses in clinical practice using structured interview data is still lacking.

Whereas Section II PDs can be reliably assessed by experts by means of the Structured Clinical Interview for DSM–5 Personality Disorders (SCID-5-P; First et al., 2016), the AMPD assessment was initially based upon unstructured clinical judgment (Criterion A) and a self-report measure for the maladaptive traits (Personality Inventory DSM–5; Krueger et al., 2012). Initial statements from the DSM–5 workgroup suggested (unstructured) assessment of Criterion A would be a relatively easy task (Skodol et al., 2011). However, several studies showed modest to low interrater reliability when severity scores were based upon interviews not specifically designed to assess Criterion A (Few et al., 2013; Zimmermann et al., 2014). Shortly after its publication in DSM–5, several specifically designed instruments were developed for the assessment of Criterion A. Hutsebaut and colleagues (2017) developed the Semistructured Interview for Personality Functioning DSM–5 (STiP 5.1). Recently, the Structured Clinical Interview for DSM–5 Alternative Model for Personality Disorders (SCID-5-AMPD; First et al., 2018) was developed, which comes in three modules: Criterion A (Module I), Criterion B (Module II), and type-specific criteria (Module III). The availability of these instruments should further support more reliable and valid classification.

The present study aims to investigate the continuity of PD diagnoses between Section II and AMPD diagnoses of PD using expert ratings based upon (semi-) structured clinical interviews. Our results will provide an estimate of the degree of change involved in the transition of these classification systems, and hence of the degree to which

previous findings based upon the traditional categories might be extended to the newly defined PD categories. First, we will explore to what extent Section II versus AMPD lead to similar prevalence rates of PD diagnoses. We focus on the six PD diagnoses that are present in both Section II and the AMPD, except for schizotypal PD (that was absent in our sample), and on other-specified/trait-specified PD. Second, we will explore whether Section II versus AMPD lead to convergent PD diagnoses in the sense that the differences between patients in diagnoses are stable across systems. As an exploratory part of this question, we will also focus on discriminant validity by inspecting the full correlation matrix (including correlations of divergent PD diagnoses) and inspect matches and mismatches for the most prevalent PD diagnoses. The present study utilizes newly available instruments to classify PDs according to both models. To our knowledge, this is the first study in which the prevalence of PDs according to the AMPD is systematically examined in a clinical sample using widely used standardized interview instruments.

Method

Procedure

Participants were treatment-seeking adults who were referred to De Viersprong, a mental health-care facility specialized in the assessment and treatment of personality disorders. A first interviewer administered the Section II interviews, which include (a) the Structured Clinical Interview for DSM–5 Syndrome Disorders (Arntz et al., 2017; official Dutch translation of the Structured Clinical Interview for DSM–5 Disorders–Clinician Version) and (b) the SCID-5-P (Arntz et al., 2017; official Dutch translation). In addition, a second, independent, interviewer administered the AMPD interviews (i.e., the STiP 5.1 and the Structured Clinical Interview for DSM–5 Alternative Model for Personality Disorders, Module II, SCID-AMPD-II) as an integral part of the admission procedure. A total of 20 different interviewers administered

the Section II interviews, and eight interviewers administered the AMPD interviews. Both the SCID-5-P interviewers and the STiP 5.1/SCID-AMPD-II interviewers were trained in administering the interviews and received regular supervision. Interviewers (Structured Clinical Interview for DSM–5 Syndrome Disorders and SCID-5-P resp. STiP-5.1 and SCID-AMPD-II) were blind for information from the other interviewer. All interviewers received standardized background information before conducting the interviews that included: demographic information, reasons for seeking help, educational level and work history and (if applicable) risk assessment (suicidality, self-harm, aggression).

Data were collected between November 2018 and November 2020. Because of COVID-19, interviews were conducted online from March 2020 until November 2020. Interviews were included in the routine assessment procedure. A total of 233 AMPD interviews were conducted. Patients were informed about the goals of the study, and 215 patients signed informed consent to use data for research purposes. Eighteen participants refused participation in the study. Of the 215 participants who signed informed consent, three were excluded because they had an autism spectrum disorder as primary diagnosis, and an additional four participants were excluded because later assessment revealed an IQ below 80. Nineteen participants had partly incomplete interview data (for example, some data were lost at the beginning of COVID-19 when transitioning from face to face to video calling, some patients dropped out of the intake procedure, some patients were not administered the Section II interviews because they had recently been administered a SCID-5-P interview in another mental healthcare facility) and were excluded using listwise deletion. A final sample of 189 participants were included in the analysis.

Sample Characteristics

Of the 189 participants, 138 (73%) were female. Their age ranged from 18 to 66, with a mean age of 34.48 (*SD* = 11.81). Most patients had at least one syndrome disorder (74.4%), with mood disorders being the

most prevalent (54.8%), followed by anxiety disorders (23.6%) and substance use disorders (18.6%).

Measures

Structured Clinical Interview for DSM–5 Personality Disorders

The SCID-5-P is a structured interview designed to assess DSM–5 PDs. Each PD criterion is scored using a 0 (absent), 1 (subclinical), or 2 (present) rating (Arntz et al., 2017; First et al., 2016). The SCID-5-P is a slightly revised version of the Structured Clinical Interview for DSM–IV Axis II PDs (SCID-II; First et al., 1997), the psychometric properties of the SCID-5-P are expected to be comparable with the SCID-II. The interrater and test–retest reliability of the SCID-II is high in samples of PDs (Maffei et al., 1997; Weertman et al., 1996) with intraclass correlation coefficients of .90 for avoidant and .95 for borderline PD in a Dutch sample (Lobbestael et al., 2011). All interviewers were trained and supervised in the administration of the interview.

Semistructured Interview for Personality Functioning DSM–5

The STiP-5.1 is a semistructured interview for the multi-item assessment of the level of personality functioning (Criterion A of the AMPD) and consists of 28 open questions, with optional clarifying questions (Hutsebaut et al., 2017). Internal consistency of the STiP 5.1 is high, with Cronbach's α of .97 for the total scale and .94 for both the self-functioning and interpersonal-functioning domains. Interrater reliability is moderate to high, with intraclass correlation coefficients ranging from .58 to .80 in a clinical sample (Hutsebaut et al., 2017). The STiP 5.1 has shown theoretically consistent associations with other instruments assessing personality pathology, supporting construct validity (Hutsebaut et al., 2017). Replication studies have been conducted in different languages, age groups, and in a diversity of patient samples (Heissler et al., 2021; Hutsebaut et al., 2021; Weekers et al., 2020; Zettl et

al., 2019). Taken together, the STiP-5.1 interview has demonstrated good psychometric properties across different samples. Interviewers were trained and supervised in the administration of the interview.

Structured Clinical Interview for DSM–5 Alternative Model for Personality Disorders, Module II

The SCID-5-AMPD-II is a semistructured interview for assessing the AMPD maladaptive personality traits (Criterion B; First et al., 2018). The SCID-5-AMPD-II assesses 25 traits, which clinicians assess on a dimensional scale ranging from 0 (very little or not at all descriptive), 1 (mildly descriptive), 2 (moderately descriptive), to 3 (very descriptive). No studies have been published on the validity or reliability of this interview yet.

Assignment of diagnoses

DSM–5 Section II Personality Disorders Diagnoses

The SCID-5-P was used to assess Section II PD criteria and assign specific PD diagnoses to patients. Criteria were scored if they were pathological, pervasive, and persistent. For example, for the borderline PD diagnosis at least five of the nine criteria had to be scored as "present." Other-specified PD was diagnosed when patients did not meet criteria for a specific PD diagnosis and five or more criteria across all PDs were present (Verheul et al., 2007). Internal consistency of the Section II criterion counts was acceptable to good with Cronbach's α's of .64 for obsessive–compulsive PD, .72 for borderline PD, .74 for avoidant PD, .77 for narcissistic PD, and .91 for antisocial PD. We also estimated unidimensional confirmatory factor analyses for each of five PDs. The results suggested good model fit and acceptable to good reliability based on McDonald's ω (see online supplemental materials).

DSM–5 Alternative Model for Personality Disorders Diagnoses

DSM–5 AMPD criteria were used to assign specific diagnoses to patients. Interviewers rated each PD-specific criterion as described

in DSM–5 based on information from the STiP 5.1 and SCID-AMPD-II interviews. For example, for the borderline PD diagnosis to be scored as "present" patients had to meet the general PD criteria (C-G) and have (a) moderate or more severe impairments in personality functioning, manifested in at least a "moderate impairment" score on at least two of the four borderline PD-specific Criterion A criteria (identity, self-direction, empathy, intimacy), and (b) a moderate or very descriptive score on at least four of the borderline PD-specific pathological personality traits (emotional lability, anxiousness, separation anxiety, depressivity, impulsivity, risk taking, hostility) with at least one of the following traits: impulsivity, risk taking, or hostility. Trait-specified PD was classified when patients did not meet criteria for any specific PD diagnosis but did meet the general PD criteria (C-G) and had (a) moderate or more severe impairments in personality functioning on at least two Criterion A domains (as measured by the STiP 5.1) and (b) one or more pathological personality trait (scored as moderate or very descriptive).

Internal consistency of the AMPD criterion counts was generally good with Cronbach's α's of .72 for antisocial PD, .75 for avoidant PD, .77 for borderline PD, and .86 for narcissistic PD. Internal consistency of the obsessive–compulsive PD diagnosis was poor with $\alpha = .54$. Criterion counts for AMPD diagnoses were the sum of the number of Criterion A and B criteria for each diagnosis. Again, we also estimated unidimensional confirmatory factor analyses for each of five PDs. The results suggested that model fit was at least acceptable for avoidant, narcissistic, and antisocial PD, whereas it was not satisfying for obsessive–compulsive and borderline PD. Reliability based on McDonald's ω was very similar to Cronbach's α, with poor results for obsessive–compulsive PD (see online supplemental materials).

Statistical analysis

To assess stability of prevalence of PD diagnoses between the classification systems, McNemar tests were conducted for pairs of Section II PD and AMPD diagnoses. Furthermore, Fleiss kappa was used

to assess convergence between Section II PD and AMPD diagnoses. To account for small prevalence rates of some PD diagnoses, exact McNemar's and Fisher's tests were conducted to derive *p* values (Fay, 2010). Lastly, as an exploratory part of our research question, Pearson correlation coefficients were used to assess associations between Section II PD and AMPD criterion counts. We assessed associations between convergent Section II PD and AMPD diagnoses (heteromethod–monotrait correlations), associations within Section II PD diagnoses between divergent PD diagnoses and within AMPD diagnoses between divergent PD diagnoses (monomethod–heterotrait correlations) and associations between divergent PD diagnoses between both models (heteromethod–heterotrait correlations; Campbell & Fiske, 1959). To assess average correlations, we used Fisher's *r* to *z* transformations.

Results

Stability of Prevalence of PD Diagnoses

Table 1 shows prevalence rates of PD diagnoses of both models. With regard to Section II PD, other-specified PD was the most frequently diagnosed PD (32.3%), followed by avoidant PD (21.2%) and borderline PD (18.5%). Similarly, for diagnoses based on the AMPD, trait-specified PD was the most prevalent (52.4%), followed by borderline PD (23.8%) and avoidant PD (20.6%). McNemar tests showed a significant difference between Section II PD and AMPD prevalence rate of PD: More patients were classified as having a PD in the AMPD model compared with the Section II PD model (see Table 1). McNemar tests showed no significant differences between Section II PD and AMPD prevalence rates for avoidant PD, borderline PD, narcissistic PD, antisocial PD, and obsessive–compulsive PD. A significant difference between the models emerged for other/trait-specified PD.

Table 1 *Prevalence of DSM-5 Section II and AMPD PD Diagnoses (N =189)*

Diagnosis[a]	Section II PD *n (% - 95% CI)*	AMPD *n (% - 95% CI)*	McNemar's test *p*	κ *(95% CI)*	Fisher's test *p*
Avoidant PD	40 (21.2 – *15.3-27.2*)	39 (20.6 – *15.3-26.4*)	1.0	.41 *(.25-.57)*	<.001
Obsessive-compulsive PD	9 (4.8 – *2.1-8.0*)	10 (5.3 – *2.6-8.9*)	1.0	.06 *(-.14-.25)*	.394
Narcissistic PD	5 (2.6 – 0.5-*5.3*)	10 (5.3 – *2.6-8.7*)	.267	.10 *(-.14-.34)*	.240
Borderline PD	35 (18.7 – *13.2-24.3*)	45 (23.8 – *18.1-30.3*)	.143	.40 *(.24-.56)*	<.001
Antisocial PD	5 (2.5 – *0.5-5.3*)	1 (0.5 – *0-1.7*)	.125	.33 *(-.16-.81)*	.026
Other/trait specified PD	61 (32.3 – *25.9-38.9*)	99 (52.4 – *45.3-59.5*)	<.001	.04 (*-.09-.17)*	.538
Any PD	132 (69.8 – *63.3-76.5*)	181 (95.8 – *92.9-98.4*)	<.001	.09 *(-.02-0.19)*	.055

Note. DSM–5 = Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition; AMPD = alternative model for personality disorders; PD = personality disorder; 95% CI = 95% confidence interval; [a]Section II dependent-, paranoid-, schizoid-, schizotypal- and histrionic PD were absent in the sample and therefore not mentioned in the table.

Convergence Between Section II PD and AMPD

Agreement between Section II PD and the AMPD model for the presence of any PD was low with κ = .09. For specific PD diagnoses, kappa ranged from nonsignificant (κ = .06) for obsessive– compulsive PD to good agreement (κ = .41) for avoidant PD (see Table 1).

Table 2 shows correlations between all Section II PD and AMPD criterion counts. Correlations between convergent Section II and AMPD criterion counts (heteromethod–monotrait correlations) were moderate to high, with *r*'s ranging from .37 (for obsessive–compulsive PD) to .59 (for avoidant PD) and an average correlation of *r* = .49. The average correlation between divergent Section II PD criterion counts was *r* = .04 and the average correlation between divergent AMPD criterion counts was *r* = .15, averaging to a correlation of *r* = .05 for the monomethod–

heterotrait correlation. Lastly the average heteromethod–heterotrait correlation was *r* = .004.

To exploratively conjecture about the source of the discrepancies between the Section II PD and AMPD model, we analyzed the source of the mismatches from both perspectives for the most prevalent PDs: avoidant and borderline PD.

For avoidant PD, 152 out of 189 cases (80.4%) were in agreement. Of the 19 cases in which Section II avoidant PD was present but AMPD avoidant PD was not, 68.4% (*n* = 13) received an AMPD trait-specified PD diagnosis, 15.8% (*n* = 3) received an AMPD borderline PD diagnosis, 5.3% (*n* = 1) an AMPD obsessive–compulsive PD diagnosis, and 10.5% (*n* = 2) did not meet criteria for an AMPD PD diagnosis. Conversely, of the 18 cases in which AMPD avoidant PD was present, but Section II avoidant PD was not, 38.9% (*n* = 7) received a Section II other-specified PD diagnosis, 22.2% (*n* = 4) received a Section II borderline PD diagnosis, 5.6% (*n* = 1) a Section II antisocial PD diagnosis, and 61.1% (*n* = 7) did not meet criteria for PD diagnosis according to Section II.

For borderline PD, 151 out of 189 cases (79.9%) were in agreement. Of the 14 cases in which Section II borderline PD was present but AMPD borderline PD was not, 64.3% (*n* = 9) met criteria for AMPD trait-specified PD, 14.3% (*n* = 2) met criteria for AMPD avoidant PD, 14.3% (*n* = 2) met criteria for AMPD narcissistic PD, and 7.1% (*n* = 1) met criteria for AMPD obsessive–compulsive PD. Conversely, of the 24 cases in which AMPD borderline PD was present but Section II borderline PD was not, 58.3% (*n* = 14) met criteria for Section II other-specified PD, 16.7% (*n* = 4) met criteria for Section II avoidant PD, 12.5% (*n* = 3) met criteria for Section II antisocial PD, 4.2% (*n* = 1) met criteria for Section II narcissistic PD, 4.2% (*n* = 1) met criteria for Section II obsessive–compulsive PD, and 33.3% (*n* = 8) did not meet criteria for any Section II PD.

Table 2 *Pearson Correlations Between Section II PD and AMPD Criterion Counts (N = 189)*

Variable	1	2	3	4	5	6	7	8	9
Section II PD criterion counts									
1.AvPD	1								
2.OCPD	-.04	1							
3.NPD	-.32**	-.07	1						
4.BPD	-.14*	-.11	.04	1					
5.APD	-.19**	-.05	.29**	.20**	1				
AMPD criterion counts									
6.AvPD	.59**	.08	-.21**	-.03	-.12	1			
7.OCPD	-.08	.37**	.07	-.15*	..05	.28**	1		
8.NPD	-.28**	.10	.53**	.09	.21**	-.18*	.14	1	
9.BPD	-.04	-.16*	.04	.56**	.18*	.17*	-.06	.16	1
10.APD	-.30**	-.19**	.47**	.34**	.48**	-.23**	-.05	.59**	.51**

Note. PD = personality disorder; AvPD = Avoidant PD, OCPD = obsessive-compulsive PD, NPD = narcissistic PD, BPD = borderline PD, APD = antisocial PD; * $p < .05$, ** $p < .01$.

Discussion

This study investigated the continuity of PD diagnoses between Section II and AMPD diagnoses of PD when using (semi-) structured clinical interviews. Our study demonstrated stability of prevalence rates between the models for all specific PD diagnoses recognized in the AMPD (except schizotypal PD, which was absent in our sample): avoidant PD, borderline PD, narcissistic PD, antisocial PD, and obsessive–compulsive PD. Trait-specified PD had a significantly higher prevalence rate than other-specified PD, resulting in a higher prevalence rate of PD in the AMPD model than in the Section II model. Convergence of PD diagnoses between the models was, according to the current standards (Kraemer, 2014), good for avoidant and borderline PD, questionable for antisocial PD, and nonsignificant for obsessive–compulsive and narcissistic PD. Associations between

Section II PD and AMPD criterion counts were generally high with an average correlation of r = .49.

Previous studies on the convergence between Section II PD diagnoses and AMPD maladaptive traits, showing high levels of convergence, have mostly used self-report measures comparing continuous variables (e.g., correlations). Convergence is likely to decrease when examining convergence between categorical decisions as in this study (Markon et al., 2011). Indeed, correlations between Section II PD and AMPD (continuous) criterion counts were generally high in our sample, and markedly higher than correlations between divergent diagnoses supporting discriminant validity. It is worth noting that several AMPD PD types were probably not unidimensional. This was especially true for borderline and obsessive–compulsive PD types, confirming previous findings with the Section II model that the categorical types are not necessarily homogeneous entities (Sharp et al., 2015; Widiger & Trull, 2007). AMPD PD types are indeed, except for narcissistic PD, comprising combinations of trait facets that are derived from different trait domains, which theoretically contradicts the assumption of a unidimensional scale. This heterogeneity may have affected the correlations of criterion counts.

Compared with previous research that was also based on clinical judgments of PD criteria and categorical decisions, we found a slightly lower convergence between both classification models. Morey and Skodol (2013) reported an average kappa of .54 between the AMPD and Section II PD models across PD diagnoses. Furthermore, Morey (2019) found a kappa of .51 for the borderline PD diagnosis. These differences can, at least in part, be explained by the methods used in both studies. In the Morey and Skodol (2013) study, the same clinician was asked to apply both Section II PD and AMPD criteria to a patient, eliminating both interrater variability and variability within patients in the expression of disorder-related characteristics. Furthermore, data from this study were used to derive the criteria for the AMPD diagnoses by maximizing the convergence with Section II, which likely biased convergence due to overfitting. In the more recent study

by Morey (2019), independent raters were asked to assess either Section II PD or AMPD criteria of borderline PD. However, clinicians were presented the exact same information: case vignettes including borderline PD as a prominent aspect of the clinical picture. Higher convergence is to be expected when the same clinical information is presented in a standardized way, because information variance in how patients express their characteristics is again eliminated (Chmielewski et al., 2015; Kraemer, 2014). Chmielewski and colleagues showed that diagnostic reliability dropped from a mean kappa of .80 when clinicians received the same information (audio-taped) to a mean kappa of .47 when using a test–retest design. We believe the test–retest design employed in this study provides a more ecologically valid and stringent test of convergence between the models.

Our findings show some aspects of continuity and discontinuity from the Section II to the Section III classification system. We found slightly lower convergence compared with DSM–5 field trials, where Section II borderline PD had an average kappa of .54 (Regier et al., 2013). However, lower kappa values are expected when using different interview schedules and different models to assess PD as in this study. Some discontinuity is to be expected and perhaps desired given the AMPD was developed to improve PD classification. Generally speaking, there seems to be fair to adequate continuity between both classification models for specific types of PDs, at least when focusing on frequently diagnosed types of PDs, like the avoidant, borderline, and to a lesser extent antisocial PD. Interestingly, these three specific types are also the most frequently studied types in treatment studies (Gibbon et al., 2020; Keefe et al., 2020; Storebø et al., 2020). Our results suggest that study results based upon previous classifications of these types may therefore largely generalize to the newly defined types. This may mitigate concerns that the evidence base for treatment of specific PD types may become futile when changing to the new system (Clarkin & Huprich, 2011; Frances, 2012). Still, a number of patients did not receive concordant diagnoses. In

avoidant and borderline PD mismatches were mostly classified as trait-specified or other-specified PD in the other model.

Indeed, an important change is the increased rate of trait-specified PD (AMPD) as compared with other-specified PD in the Section II model, which contributes to an overall increased prevalence of PD diagnoses when using the AMPD. The threshold for diagnosing PDs seems therefore lower when using the new classification system. This may be due to the guidelines that have been followed when using the interview measures. For the other-specified PD to be present, a minimum of five specific criteria across different types was required beyond the fulfillment of the general PD criteria. Criteria for the trait-specified PD were already met when subjects displayed moderate impairments on two or more elements of personality dysfunction (Criterion A), whereas the criterion of simultaneously displaying one or more pathological personality traits (Criterion B) seemed to be met in any of these cases and therefore did not impose an additional threshold for the diagnosis. In fact, based upon these findings, one could question the incremental value of Criterion B in differentiating PD from no-PD from the perspective of the AMPD.

The finding that the AMPD seems to result in more categorical PD diagnoses—as much as 95.8% in our selected sample as compared with 69.8% according to the Section II model—can be interpreted in two ways. It may imply that the AMPD is overly sensitive and "falsely" detects PDs or it may mean that the AMPD is more sensitive than the Section II model and succeeds in detecting "true" PD patients that remain undiagnosed within the Section II model. More studies are needed to compare both models in terms of construct and predictive validity with regard to essential outcomes, like social and occupational functioning, treatment outcomes and long-term disability, to determine which of both models may be capturing the core of PDs more veridically.

Convergence between the AMPD and Section II PD models was low for narcissistic and obsessive–compulsive PD, and the number of cases were also rather low according to both models, complicating

interpretations. The low convergence for narcissistic PD may be explained by differences between both models in the ways they represent narcissism. The Section II model has been criticized for capturing only the grandiose or "overt" types of narcissism, whereas the AMPD allows a narcissistic PD diagnosis for vulnerable or "covert" presentations of narcissism as well (Levy, 2012; Pincus et al., 2015; Skodol et al., 2014). This may explain why the AMPD identified twice as much narcissistic PD as the Section II model (although this difference was not statistically significant). This hypothesis is in need of further investigation. Also, obsessive–compulsive PD appears to be captured differently by both models, at least in our sample. This is in line with findings from Liggett and colleagues (2017), who found Rigid Perfectionism was the only trait that uniquely predicted Section II obsessive–compulsive PD. Subsequently, several studies did not support the inclusion of Restricted Affectivity and/or Intimacy Avoidance in the AMPD model (Anderson et al., 2014; Bastiaens et al., 2016; Fossati et al., 2013; Liggett et al., 2017). Of note, AMPD obsessive–compulsive PD showed low internal consistency and seemed to exhibit a more complex factor structure in our sample. Taken together, our data suggest that the AMPD obsessive–compulsive PD criteria should be modified if the aim would be to ensure continuity and to prevent radical changes from Section II.

Finally, we want to highlight that a more fundamental issue underlying this research concerns the empirical and clinical utility of categorical diagnoses (i.e., types of PD). Although the AMPD enables typological classifications of personality pathology, the 11th revision of the World Health Organization (WHO) International Classification of Diseases has deleted all types (excepting for a borderline specifier). Several authors have indeed disputed the usefulness of PD types (Livesley, 2012; Widiger, 2013). Our study design did not address this fundamental issue and our data analytic strategy does not allow for a discussion on the utility of the categorical aspect of the AMPD. Future studies should address this issue and involve aspects of incremental validity of types (compared with a purely dimensional approach) as

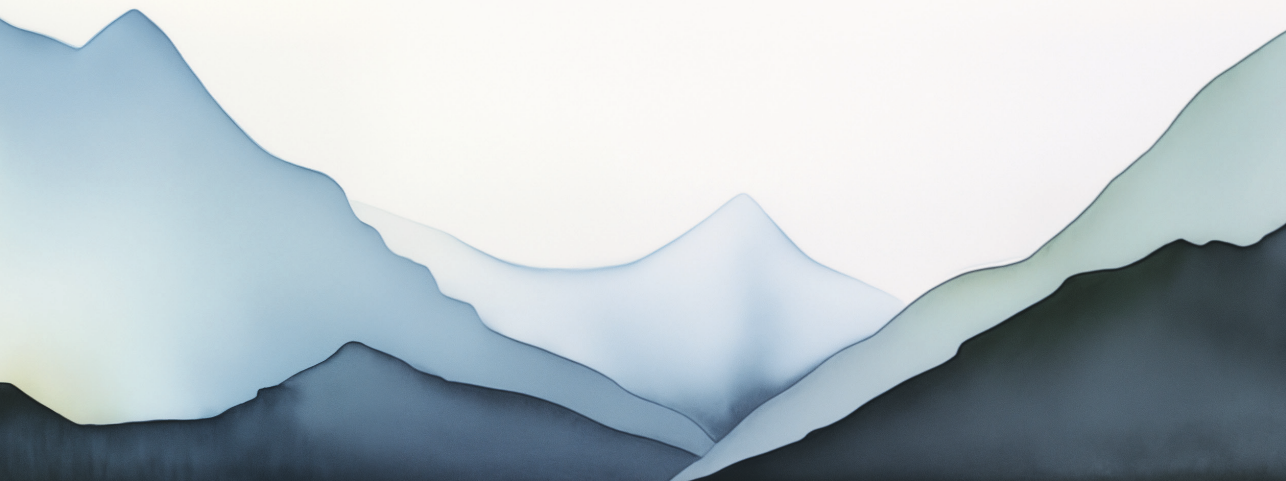
well as aspects of clinical utility related to the use of types. A related issue is how these types should then be defined. The AMPD aimed to provide continuity with Section II PDs and additionally dropped specific types that could be represented by one single trait domain (like the paranoid PD). Another option would be to define empirically more homogeneous types, based upon extreme positions on specific maladaptive traits. For example, the multifactorial AMPD criteria for obsessive–compulsive PD could be redefined based upon the trait dimension of (Extreme) Conscientiousness or Anankastia. However, one may then question the additional value of types given their overlap with single trait dimensions. Still another option would be to define specific types as emergent interpersonal syndromes that consist of components that may be uncorrelated (i.e., heterogeneous) but show interaction effects on interpersonal outcomes (Lilienfeld et al., 2019).

Our study has several strengths and limitations. It is, to our knowledge, the first study in which the Section II PD and AMPD models are compared using (semi-) structured clinical interviews specifically designed for each model, thereby complying with general recommendations to use structured interview designs for assessing PDs (Landelijke Stuurgroep Multidisciplinaire Richtlijnontwikkeling in de GGZ, 2008). We are also not aware of any previous study that has assessed AMPD traits using an interview schedule. Interviews were implemented in the regular admission procedure, supporting ecological validity. Both PD diagnoses were independently assessed by well-trained and supervised professionals. On the other hand, several limitations should also be considered when interpreting our findings. First, despite our relatively large sample, there was a low prevalence rate for some specific PD types, like antisocial, narcissistic, and obsessive–compulsive PD. To account for the low prevalence rates, we used exact *p*-values and confidence intervals. However, conclusions for these diagnoses should be interpreted with caution. Second, we did not use the SCID-AMPD Module III interview, specifically designed to assess the AMPD diagnostic criteria, which might have lowered the convergence between the models. We chose the STiP-5.1 instead

of the SCID-AMPD Module I because more validation studies have been published on it so far. Also, no studies on the Dutch version of the SCID-AMPD Module I have been published. Third, due to the specialized nature of the setting, our sample consisted of a rather homogeneous group of patients with PDs. The relatively low stability of the general prevalence may partly be explained by the specific nature of the sample, with all subjects being specifically referred for personality problems and therefore displaying at least some traits of PD. Stability rates may increase within a more heterogeneous sample that would be more representative for all patients seeking help for mental problems. Hence, future studies should focus on convergence between the models in a more heterogeneous clinical sample, as well as community samples. Finally, although the choice of specifically designed standardized interview measures has been mentioned as a strength, there is also a limitation in the assessment procedure used in this study. Previous studies have shown modest reliability of interview measures (Zimmerman, 1994) and divergence between diagnostic interviews assessing the same (Section II) categorical diagnoses (Pilkonis et al., 1995). Lack of convergence may therefore be also reflecting in part measurement error or interview-specific variance. Including multiple measures and raters for both the Section II as Section III models may provide a better test of the true (lack of) convergence of PD types between both models.

In sum, this study showed stability of prevalence rates for specific PD diagnoses between the Section II PD and the AMPD model. The AMPD model, however, appears to diagnose more PD in general, reflected by a high number of trait-specified PD. Studies are needed to compare both models in terms of construct and predictive validity to determine which of the models captures PDs more accurately. Convergence between the Section II PD and AMPD model was adequate for the most frequently diagnosed and studied PDs, but lower than previously found, likely due to the more stringent test–retest design used in this study. Future studies should investigate prevalence and convergence of PD diagnoses in other clinical and community samples and should

identify which of both models may prove to be most valid in terms of predicting current and future impairments.



Chapter 5

Scripting the DSM-5 Alternative Model for Personality Disorders assessment procedure: A clinically feasible multi-informant multimethod approach

Laura C. Weekers, Joost Hutsebaut, Bo Bach, Jan H. Kamphuis
Personality and Mental Health (2020), 14, 304-318

Abstract

Published case studies on the DSM-5 (section III) Alternative Model for Personality Disorders (AMPD) generally utilized unstandardized assessment procedures or mono-method approaches. We present a case from clinical practice to illustrate a standardized, clinically feasible procedure for assessing personality pathology according to the full AMPD model, using a multi-method approach. We aim to present a procedure that can guide and inspire clinicians that are going to work with dimensional models as presented in DSM-5 and ICD-11. Specifically, we show how questionnaire and interview data from multiple sources (i.e., patient and family) can be combined. The clinical case also illustrates how Criterion A (i.e., functioning) and B (i.e., traits) are interrelated, suggesting that the joint assessment of both Criterion A and B is necessary for a comprehensive and clinically relevant case formulation. It also highlights how multi-method information can enhance diagnostic formulations. Finally, we show how the AMPD model can serve treatment planning and provide suggestions for how patient feedback might be delivered.

Introduction

In this paper, we will present a standardized, clinical approach to assessing personality pathology using the AMPD model. Using a multi-informant multi-method (MI-MM) approach, we aim to demonstrate how different instruments (questionnaires and interviews) from different sources (patient and family) can be combined in a semi-structured procedure. Furthermore, we aim to demonstrate how this information may be profitably shared with the patient and how it may inform treatment planning. To contextualize this procedure, we offer an elaborate case-presentation to illustrate each step. Before detailing the procedure, we will provide a quick review of AMPD research findings that guided our choices in designing our AMPD assessment procedure.

The personality disorder field is currently shifting from categorical models of personality disorders (PD) towards dimensional models. ICD-11 (World Health Organization, 2019) recently introduced a dimensional model in their chapter on PDs, DSM-5 however introduced a dimensional model as an alternative approach to the assessment of personality pathology in DSM-5 section III (i.e., AMPD; American Psychiatric Association [APA], 2013). Because the AMPD model has already been extensively evaluated and used for clinical purposes in a number of years, the present article focuses on this approach, while underscoring that the same utility is expected to apply to the ICD-11 classification as well (Bach & First, 2018). The DSM-5 AMPD comprises a profile of impairments in self- and interpersonal functioning along with a constellation of pathological traits (APA, 2013; Hopwood et al., 2011). Assessment follows a stepwise procedure, enabling subsequent diagnostic refinement. Clinicians start with assessing impairments in self- and interpersonal functioning (Criterion A), using the Level of Personality Functioning Scale (LPFS), followed by an assessment of 25 maladaptive trait facets that are organized in the five broad domains of Negative Affectivity, Detachment, Antagonism, Disinhibition, and

Psychoticism (Criterion B). By checking stipulated type-specific criteria, clinicians can determine whether the profile of Criterion A impairments and Criterion B trait facets matches one of six specific types of PDs, provided that patients meet criteria C-G of the general diagnostic requirements. Additionally, a trait-specified PD diagnosis is provided for patients suffering from at least moderate impairments in personality functioning, but whose presentation is not matching one of the specific types (corresponding to "other specified" in DSM-5 Section II). Finally, the clinician may refine this global assessment by specifying the different severity scores and relevant trait facets, allowing dimensional specifiers beyond the categorical diagnosis. The AMPD model encompasses both strengths and impairments in functioning, along with resilient versus pathological features. The profile of personality functioning and traits may thus yield a balanced picture of the patient's psychological infrastructure, interpersonal dynamics, and clinical prognosis, and may accordingly be especially informative for treatment planning (Berghuis et al., 2014; Rodriguez-Seijas et al., 2019).

The first and foremost step in the AMPD model is the assessment of personality related impairment, i.e. Criterion A, without which no PD can be present. Research into the reliability of the level of personality functioning ratings has yielded mixed results. Using a case-vignette methodology in which brief case information was selected and narratively organized by the research team, Garcia and colleagues (2018) observed promising reliability of LPFS ratings. However, reliability ratings were lower when students or clinicians had to self-select the information from the clinical interviews to infer LPFS ratings (Few et al., 2013; Zimmermann et al., 2015). To assess Criterion A, several interview and self-report instruments have been specifically developed (Bach & Hutsebaut, 2018; First et al., 2018; Huprich et al., 2018; Hutsebaut et al., 2016; Hutsebaut et al., 2017; Morey, 2017; Thylstrup et al., 2016; Weekers et al., 2019). Studies using these specific interview instruments showed superior interrater reliability compared to non-specific clinical interviews with respect to the assessment of Criterion A (Hutsebaut

et al., 2017; Christensen et al., 2018). Furthermore, there is supportive evidence for internal consistency and construct validity for self-report questionnaires assessing personality functioning (Bach & Hutsebaut, 2018; Huprich et al., 2018; Hutsebaut et al., 2016; Morey, 2017; Weekers et al., 2019). However, no studies have investigated the convergent validity of self-report versus clinical interview ratings.

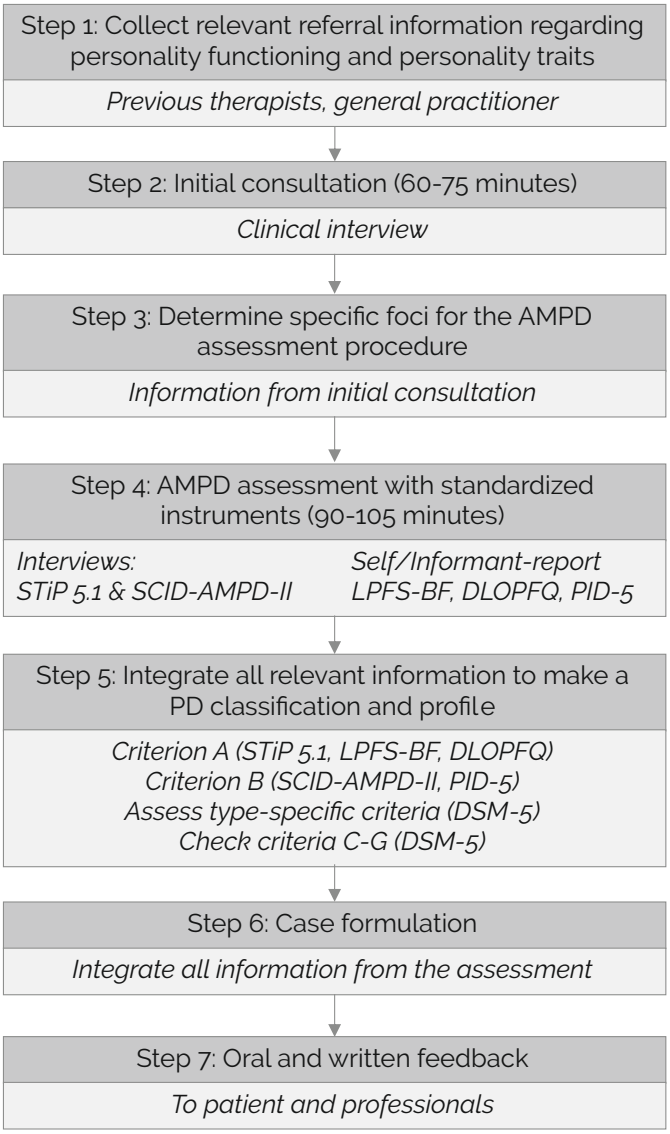
Subsequent to the assessment of the LPFS (Criterion A), the specific expression of personality dysfunction is delineated in terms of stylistic traits (i.e. Criterion B). The majority of Criterion B research draws upon a self-report instrument: the Personality Inventory for DSM-5 (PID-5; Krueger et al., 2012). The PID-5 shows a stable factor structure across different samples and cultures, with good internal consistency at domain and facet levels (Watters & Bagby, 2018). Furthermore, Bach and colleagues (2017) demonstrated that findings from non-clinical data were generalizable to clinical populations, thus supporting the results of many non-clinical studies. Although the Structured Clinical Interview for DSM-5 Alternative Model of Personality Disorders Module II (SCID-AMPD; First et al., 2018) provides an interview-based alternative, no studies to date have examined its reliability. Previous studies showed a wide variation in reliability scores when traits were assessed based upon clinical vignettes (Garcia et al., 2018). As was noted for Criterion A, no convergence studies have been conducted between questionnaire-based assessment of traits versus interview-based ratings.

Clinical application of the AMPD not only requires assessment of Criteria A and B, but also a clinical integration of the comprising elements in a way that represents the nature of a patient's problems and informs subsequent treatment. Such profitable integration relates to the issue of clinical utility, which has received rather scant attention in the AMPD research to date, but may be what matters most to clinicians. Early critics of the AMPD model have questioned the ease of use of the model in clinical practice (Clarkin & Huprich, 2011; Shedler et al., 2010). Morey and colleagues (2014) assessed the clinical utility of the AMPD model by asking clinicians to diagnose their own patients

using the DSM-IV-TR categorical diagnosis and the AMPD model. Clinicians reported that the AMPD model was as useful or more useful than the categorical system, especially with respect to communication with patients, treatment formulation, comprehensiveness, and global descriptive utility. A similar survey on the ICD-11 PD classification using the same approach concluded that mental health professionals (i.e., psychologists, psychiatrist, and nurses) generally preferred the ICD-11 dimensional approach over the ICD-10 categorical approach, particularly in respect to utility for treatment formulation (Hansen et al., 2019). Furthermore, some authors have illustrated the clinical value of the AMPD model by describing case studies (Bach et al., 2015; Bach & Bernstein, 2019; Morey & Stagner, 2012; Pincus et al., 2016; Schmeck et al., 2013; Skodol et al., 2015; Waugh et al., 2017; Waugh et al., 2019). Although these demonstrations of the clinical utility of the AMPD model are informative and inspiring, they also have some limitations. First, most case studies lack the standardized use of specifically tailored assessment instruments for both Criterion A and B (Morey & Stagner, 2012; Schmeck et al., 2013; Skodol et al., 2015; Waugh et al., 2017; Waugh et al., 2019). As noted, the reliability of the Criterion A assessment appears to benefit from the use of specifically designed instruments. Second, most case studies relied on the patient as the exclusive source of information, especially with regard to Criterion B (Bach et al., 2015; Pincus et al., 2016; Waugh et al., 2019). However, patients with (severe) personality disorders frequently have difficulty reflecting on their internal experiences, and may offer an incomplete and/or biased picture of their functioning (Carnovale et al., 2019). In similar vein, the wholesale reliance on a self-report inventory (e.g., PID-5) to assess traits could be questioned (Huprich et al., 2011). Finally, most case studies (Pincus et al., 2016; Schmeck et al., 2013; Skodol et al., 2015) were limited in their description of how the AMPD information may be used to inform treatment or how feedback could be provided to the patient. To address these issues, we developed a standardized clinical approach to assessing personality pathology using the AMPD model. Given the conceptual overlap between the

AMPD and the PD chapter in ICD-11, this clinical approach may also be informative for clinicians that are going to use ICD-11.

Figure 1 Stepwise approach to AMPD assessment.



A Multi-Method Multi-Informant AMPD Assessment Procedure

We will now describe the successive steps of the AMPD assessment procedure, and illustrate each step by describing the case of a 39-year old man, henceforth called "Adam", referred for help by his general practitioner for psychiatric assessment and evaluation for treatment. Adam gave full written consent to utilize his clinical records for the current case illustration.

Step 1: Collect relevant referral information regarding personality functioning and personality traits in social, occupational and relationship domains

At referral, some information pertinent to the patient's personality functioning may be immediately available (e.g., based upon earlier treatment or reasons for referral). Additional information may be collected from previous therapists or the general practitioner (GP). Relevant topics include current social network, stability of intimate and family relationships, and course of academic and professional career.

Adam described a long history of prematurely terminated studies, discontinued jobs, along with enduring sleeping problems, stress complaints, and low mood. The GP surmised that this pattern of problems might be rooted in his personality. At the time of the intake, Adam held no paid job but did some voluntary work in a home for the elderly. According to the GP, Adam spent a lot of time gaming and seemed to have a very limited social network.

Step 2: Conduct a clinical intake interview

The initial consultation will usually consist of an open-ended clinical interview. The therapist explains the full assessment procedure and invites the patient to talk about his or her own reasons for seeking help. Furthermore, s/he invites information of the patient's developmental background, family of origin, current and past relational and professional context, previous treatment history, and medication use.

It warrants mentioning that Adam presented at the clinical intake interview accompanied by his mother, though not upon our request.

His mother explained that Adam might be inclined to "mislead" the clinician by providing "too positive a view of himself and his problems". Adam in fact agreed and added he experienced difficulty in providing an overview of his current situation. Both Adam and his mother seemed highly stressed and fearful. We observed that Adam frequently behaved rather submissively; being overly polite, highly apologetic, or frequently praising the clinician. Moreover, Adam tended to intellectualize when talking about his problems, using difficult and often rather vague language, without reference to specific, concrete examples of his personal issues. His mother would interrupt and then fill in for Adam during the consultation.

With regard to his family of origin, Adam described his father as a "verbally abusive man" who would often target him for severe scolding. His parents divorced when he was 16 years old. Adam and his mother drew a picture of persistent social and emotional problems over the course of his childhood. First, starting in early adolescence, Adam had recurring depressive episodes. Moreover, Adam had always experienced difficulties in connecting with his peers. He was repeatedly and severely bullied in primary school, and changed schools several times to escape this. His parents put him in a protective school for children with special needs. After completing primary school, he was advised to follow a lower level technical secondary school, although later testing revealed intelligence in the superior range.

After leaving secondary school, Adam initiated several studies, including Engineering, Philosophy, Law, Psychology, and most recently Social Work Studies. However, he failed to complete any of these studies: Adam would begin enthusiastically, but after some time got increasingly stressed and then prematurely terminated his studies. Over the past 20 years, he accrued a study loan of 40.000 euros. During this period, he held many temporary jobs which usually ended up in conflicts with superiors. Adam then either withdrew and simply stopped showing up, or got into a verbal fight and was subsequently fired. He explained his behavior as being a response to experienced injustice from superiors.

Adam explained that he typically got very upset when feeling pressured to do something. He preferred to be alone and not feel demands from others, but also recognized that complete withdrawal created increased feelings of loneliness and depression. Adam disclosed that he felt "like a failure" with no future perspective, and he had become increasingly desperate. In fact, he reported feeling anxious and depressed most of the time, and that he sought relief from the negative emotions by withdrawing socially. Instead, he had been seeking refuge in online gaming. At times, the stress was also causing him physical complaints including severe headaches. He had also developed a pattern of compulsions, like counting, to avoid feeling miserable.

Adam reported a history of suicidal ideation but never to the point of planning or making a suicide attempt. He had never engaged in deliberate self-injury, but his mother reported extended periods of neglected self-care. Adam seemed ambivalent about seeking help. On the one hand, he would often minimize his problems (for example saying he was "just lazy"). On the other hand, he had grown increasingly demoralized and despondent, and indicated he did not see his way out of this situation.

Step 3: Integrate referral and intake information to determine specific foci for the AMPD assessment procedure

Prior to conducting the Criteria A and B interviews, the clinician integrates all available information to appraise which areas of personality functioning or trait facets may be especially relevant for subsequent exploration. Although all elements of personality functioning and trait facets will be explored in the assessment, noting specific areas of interest based on the referral and intake information helps the clinician to develop some tentative hypotheses regarding the level of personality functioning and trait elevations that can help focus the assessment.

Based on the collected information, several foci of attention could be identified. Regarding the LPFS domains, there was clear evidence of

severe problems in self-direction, as reflected by Adam's longstanding inability to complete studies and hold jobs. Adam's history was also suggestive of impaired self-esteem, and his social isolation and inability to collaborate in a professional context as expressed in repeated conflicts with superiors pointed to interpersonal impairment. With respect to personality traits, there were several indications of easily triggered antagonism, as exemplified by his recurring conflicts at work. Conversely, his general overly friendly and compliant demeanor suggested submissiveness. His chronic stress may point to increased dysfunctional negative emotions. Finally, several indications of detachment were evident, e.g. his extensive social withdrawal and him seeking refuge in isolated activities (e.g., gaming). In sum, Adam's history was consistent with a wide range of impairments in personality functioning along with several pronounced maladaptive traits, to be explored further in specific AMDP assessment.

Step 4: Administer standardized measures of personality functioning and traits, involving different sources of information

Our procedure includes the collection of self- and informant-report data along with clinician-ratings based on structured clinical interviews. Here, we will briefly describe the instruments used in the assessment.

The Level of Personality Functioning – Brief Form 2.0 (LPFS-BF 2.0). The LPFS-BF 2.0 is a 12-item self-report questionnaire (Weekers et al., 2019) with a 4-point Likert scale for assessing Criterion A of the AMPD. Internal consistency estimates for the LPFS-BF 2.0 were high in a sample of patients with PD, with α=.82 for the total scale and α=.79 and α=.71 for the Self- and Interpersonal Functioning Scales (Weekers et al., 2019). We developed an informant version of the LPFS-BF 2.0 as an adaptation from the original LPFS-BF 2.0 for the current study.

DSM-5 Levels of Personality Functioning Scale (DLOPFQ). The DLOPFQ is a 66-item self-report questionnaire (Huprich et al., 2018) for assessing the level of personality functioning (Criterion A) of the AMPD. Items are rated on a scale from 1 (strongly disagree) to 6 (strongly agree). The questionnaire yields scores for the four elements of the

LPFS (Identity, Self-direction, Empathy & Intimacy). Internal consistency of the scales was high in a sample of in- and outpatient psychiatric patients with α's ranging from .72 to .94 (Huprich et al., 2018).

Personality Inventory DSM-5 (PID-5). The PID-5 is a 220-item self-report questionnaire for assessing the Criterion B pathological traits of the AMPD (Krueger et al., 2012). Items are rated on a scale from 0 (very false or often false) to 3 (very true or often true). The questionnaire consists of 25 facets (maladaptive personality traits), constituting 5 higher order domains (Negative Affectivity, Detachment, Antagonism, Disinhibition & Psychoticism). Internal consistency of the facets was high in a community sample with Cronbach's α's ranging from .72 to .96 (Krueger et al., 2012). We used both the self-report as well as the informant version of the PID-5.

Semi-Structured Clinical Interview for Personality Functioning DSM-5 (STiP-5.1). The STiP 5.1 (Hutsebaut et al., 2017) is a semi-structured interview for assessing the 12 capacities of the LPFS; clinicians rate each capacity from level 0 (little or no impairment), level 1 (some impairment), level 2 (moderate impairment), level 3 (severe impairment) to level 4 (extreme impairment). In a previous study reporting on both a clinical and community sample (Hutsebaut et al., 2017), internal consistency of the STiP-5.1 was high with a Cronbach's alpha of .97 for the total scale, and .94 for both the Self-functioning and Interpersonal functioning domain. Interrater reliability was good, with ICCs ranging from .81 to .92 in the total sample, and ICC's ranging from .58 to .81 in the clinical sample (Hutsebaut et al., 2017). Administration time of the instrument requires between 45 and 60 minutes to administer and yields clinician-rated element-, domain- and total impairment scores for all 12 capacities.

Structured Clinical Interview for DSM-5 Personality Disorders – Alternative Model of Personality Disorders, Module II (SCID-AMPD-II; specifically translated to Dutch for this study by the authors). The SCID-AMPD Module II (First et al., 2018) is a semi-structured interview assessing pathological personality traits. The clinician is to evaluate the degree to which each trait facet is descriptive of the patient: 0 (not

descriptive), 1 (mildly descriptive), 2 (moderately descriptive), or 3 (very descriptive). To our knowledge, no information on the psychometric properties of the interview are available yet. We omitted the "general overview" questions of the SCID-AMPD Module II because this information (demographic variables, education and work history, current and previous psychiatric complaints) was presumably already covered in the intake interview. Administration of SCID-AMPD Module II requires another 45 minutes; total administration of the interview schedules thus ranges from 90-105 minutes.

Table 1 and 2 display the scores of both Adam and his mother. Scores on the LPFS-BF 2.0 range from 1 to 4, with higher scores reflecting more severe personality dysfunction. The DLOPFQ scores range from 14 to 114, again with higher scores reflecting more severe personality dysfunction. PID-5 scores range from 0 (not at all descriptive) to 3 (very descriptive). We calculated T-scores (in parenthesis in Tables 1 and 2) to compare Adam's scores to a normative clinical sample of treatment seeking adults (descriptions of the normative samples: LPFS-BF 2.0 Weekers et al., 2019; DLOPFQ Huprich et al., 2018; PID-5 Wright et al., 2015). The LPFS-BF 2.0 and DLOPFQ suggested average impairment in all domains of personality functioning as compared to clinical samples, with above average impairment in intimacy as assessed by the DLOPFQ. When comparing informant (i.e., his mother's) report with Adam's self-report, the following picture emerged. They largely agreed on the severity of his personality dysfunction (severe) and were fairly consistent in pinpointing elevated problematic personality traits (e.g., Negative Affectivity, Detachment), as well as domains that were relatively unproblematic (Disinhibition and Psychoticism). However, Adam endorsed more antagonistic traits, especially Grandiosity, than his mother recognized. On the other hand, his mother considered Adam more anxious and suspicious than Adam reported.

After completing the questionnaires, Adam was administered the (semi-) structured standardized interviews for the systematic assessment of Criteria A and B. Standardized instruments were selected because of their superior psychometric qualities as

compared to regular clinical interviews. In addition to the LPFS and the assessment of maladaptive traits, the DSM-5 AMPD describes Criterion A and B criteria for specific types of PD. We used the information as collected in these interview procedures to assess the specific Criterion A and B criteria for the different PD types (i.e. Avoidant-, Obsessive-Compulsive-, Narcissistic-, Borderline-, Antisocial- and Schizotypal PD), enabling us to omit the Module III of the SCID-AMPD (which assesses these type-specific criteria).

Adam's ratings based upon STiP-5.1 (Criterion A) and SCID-AMPD Module II (Criterion B) are displayed in Tables 1 and 2. In general, the level of severity as based on the STiP-5.1 corresponded to 'Severe impairment', fully consistent with both Adam's and his mother's questionnaire-based ratings of his level of personality functioning. Self-esteem was especially impaired (extreme impairment). Based on the SCID-AMPD Module II, several trait domains were elevated. Most descriptive were the Grandiosity and Attention Seeking trait facets. Interview ratings and self/informant ratings were generally fairly consistent, with a few noteworthy discrepancies in the Antagonism domain. For instance, Attention Seeking and Grandiosity were rated rather higher by the clinician (i.e., interview data) than both Adam and his mother had endorsed on the PID-5. Likewise, the clinician rated Intimacy avoidance higher than Adam, and Irresponsibility higher than both Adam and his mother had. We will return to the clinical utility of discussing such patterns of convergence and divergence in the next steps.

Table 1 *Criterion A results (N=1)*

Measure	Scale	Clinician Raw score	Self-report Raw score (T-score)	Informant Raw score (T-score)
LPFS-BF 2.0	Self-functioning		3.33 (52.1)	3.00 (46.8)
	Interpersonal functioning		2.60 (52.7)	3.00 (59.0)
	Total LPFS score		3.00 (53.4)	3.00 (53.4)
DLOPFQ	Identity		48 (53.4)	-
	Self-Direction		49 (52.3)	-
	Empathy		53 (52.6)	-
	Intimacy		76 (60.2)	-
STiP 5.1	TOTAL SEVERITY SCORE	3		
	Identity	3		
	Uniqtue Self	3		
	Self-esteem	4		
	Emotions	2		
	Self-direction	3		
	Goals	3		
	Values	2		
	Self-reflection	3		
	SELF FUNCTIONING	3		
	Empathy	3		
	Understanding others	2		
	Perspectives	3		
	Impact	3		
	Intimacy	3		
	Connectedness	3		
	Closeness	3		
	Reciprocity	3		
	INTERPERSONAL FIE	3		

Note: LPFS-BF 2.0 = Level of Personality Functioning Scale Brief Form 2.0; DLOPFQ = DSM-5 Level of Personality Functioning Questionnaire; STiP 5.1 = Semi-structured interview for Personality functioning DSM-5.

Table 2 *Criterion B PID-5 and SCID-AMPD module II results (N=1)*

Scale	Clinician Raw score	Self-report Raw score (T-score**)	Informant Raw score (T-score**)
Negative Affectivity		1.31 (47.5)	1.74 (54.6)
Emotional Lability	2	1.86 (54.7)	2.00 (56.6)
Anxiousness	3	1.78 (50.9)	2.50 (61.1)
Separation Insecurity	1	0.29 (37.2)	0.71 (43.3)
*Submissiveness**	2	1.50 (50.8)	1.25 (53.1)
*Hostility**	2	1.50 (53.0)	1.60 (54.5)
*Perseveration**	3	2.44 (69.7)	1.89 (60.8)
Detachment		1.44 (54.3)	1.87 (62.5)
Withdrawal	2	1.90 (57.6)	2.20 (62.4)
Intimacy Avoidance	2	0.17 (41.7)	1.17 (56.2)
Anhedonia	2	2.25 (61.4)	2.25 (61.4)
*Depressivity**	3	-	2.14 (64.1)
*Restricted Affectivity**	2	1.57 (60.0)	0.71 (45.4)
*Suspiciousness**	2	1.43 (52.5)	2.83 (74.8)
Antagonism		1.31 (60.8)	0.33 (41.2)
Manipulativeness	1	1.00 (52.0)	0.40 (42.9)
Deceitfulness	1	1.10 (58.0)	0.10 (39.8)
Grandiosity	3	1.83 (67.5)	0.50 (45.7)
*Attention Seeking**	3	1.00 (50.9)	0.25 (39.5)
*Callousness**	1	0.21 (43.5)	0.21 (43.5)
Disinhibition		1.04 (50.5)	1.21 (53.7)
Irresponsibility	2	0.57 (48.5)	0.57 (48.5)
Impulsivity	0	1.00 (49.3)	1.17 (51.7)
Distractability	2	1.56 (53.6)	1.89 (58.4)
*Risk taking**	0	0.50 (36.9)	1.14 (47.9)
*Rigid Perfectionism**	2	1.10 (45.0)	1.80 (56.3)
Psychoticism		0.84 (49.8)	0.93 (51.4)
Unusual Beliefs / Experiences	0	0.38 (45.3)	0.25 (43.17)
Perceptual Dysregulation	1	0.83 (52.3)	0.83 (52.3)
Eccentricity	2	1.31 (51.7)	1.69 (56.3)

Note: PID-5 = Personality Inventory for DSM-5; SCID-AMPD = Structured Clinical Interview for the DSM-5 Alternative Model for Personality Disorders; *These facets are not included in the PID-5 domain score calculation. For each PID-5 domain, only the three primary facets are included in its aggregate score; **T-scores were computed relative to a clinical reference sample.

Step 5: Develop the AMPD classification, and determine the profile of personality impairments and traits

Next, the clinician uses all available information to follow the different scoring and classification steps of the AMPD model. Both convergence and divergence between clinician- self- and informant report should be considered. Ultimately, the assessment is a clinician-based procedure, assigning the clinician the responsibility to weigh different sources of information and make clinical judgments based on all available information. Areas of convergence and divergence may be especially informative for structuring feedback to patients (see step 6 and 7).

The clinician first determines the severity at element- (identity, self-direction, empathy, intimacy), domain- (self- and interpersonal functioning), and general (personality functioning) level. Although conceptually the LPFS is considered a single dimension, in our experience some differentiation may be seen with regard to specific elements and aspects, thus highlighting areas of strengths or increased vulnerability. Second, the clinician makes a profile of elevated personality trait facets. Again, it may be helpful to highlight not only (extreme) maladaptive traits, but also to note relatively intact functioning. Third, in keeping with traditional clinical practice, the clinician systematically assesses the type-specific criteria using the DSM-5 criteria for the six types (Watters et al., 2018). Integrating all of Adam's scores (in this MI-MM procedure), the clinician concluded that Adam's level of personality functioning was best captured by severe impairment (i.e. level 3), with an extremely impaired self-esteem aspect (i.e., level 4). Taking all measures into account, there was robust evidence for elevations in the domains of Antagonism (especially Grandiosity and Attention Seeking), Negative Affectivity (especially Anxiousness, Perseveration, and Submissiveness), and Detachment (especially Intimacy Avoidance, Depressivity and Anhedonia). Conversely, on virtually all measures, Adam scored relatively low on most facets of Disinhibition and Psychoticism, suggesting that impulse control and reality testing were intact.

With regard to type-specific criteria (APA, 2013), the clinician concluded that Adam met Criteria A and B for Narcissistic and Avoidant PD. His self-esteem alternated between grandiose/ inflated and deflated, and he was extremely vulnerable to experiencing criticism or slights from others (NPD, A1 Identity; APA, 2013). Personal standards were unrealistically high in order to view himself as exceptional, but he often withdrew because of fear of failure (NPD, A2 Self-direction; APA, 2013). He exhibited a pervasive inability to appraise his impact on others, leading to interpersonal problems and conflicts (NPD, A3 Empathy; APA, 2013). He was overly sensitive to criticism and rejection and quick to infer that others perceived him in a very negative way (APD, A3 Empathy; APA, 2013). Although he was sensitive to reactions of others, this appeared to be motivated by the desire to avoid criticism and negative feelings; he did not appear to be motivated by a genuine interest in the feelings and experiences of others. Mutuality was limited by either a submissive stance to avoid feeling ridiculed (APD, A4 Intimacy; APA, 2013) or an overly controlling and superior stance to protect his self-esteem (NPD, A4 Intimacy; APA, 2013). Adam reported that he considered himself destined for "something special" and often felt slighted or misunderstood by others leading to condescension toward others (NPD, B1 Grandiosity; APA, 2013). He was inclined to withdraw socially as a way of protecting against criticism or negative feedback (APD, B2 Withdrawal; APA, 2013). Feelings of nervousness, tension, and a fear of being shamed were prevalent (APD, B1 Anxiousness; APA, 2013). Although not stereotypically attention seeking, Adam was strongly motivated to gain the admiration of others, reflected by his high need for achievement and grandiose fantasies (NPD, B2 Attention Seeking; APA, 2013). Finally, Adam also endorsed significant Anhedonia (APD, B3; APA, 2013), which is also consistent with his escape into online gaming, and not engaging in real life experiences.

Fourth, the clinician checks whether the general criteria C-G are met. In Adam's case this was clear: his impairment was inflexible and pervasive, relatively stable across time, and not better explained by

another mental disorder, nor attributable to the effects of a substance or medical condition nor normal for his developmental stage and sociocultural environment.

Finally, the clinician summarizes all information and makes a classification, using additional specifiers. In Adam's case: Narcissistic and Avoidant PD with Submissiveness, Perseveration, Hostility, Suspiciousness, Distractibility and Rigid Perfectionism.

Step 6: Develop a case-formulation on the dynamic interaction of maladaptive personality traits and impaired personality functioning

Arguably, the depth and clinical utility of the AMPD model resides not as much in the specific diagnostic notation it provides, but is especially evident in the information the AMPD yields for the construction of a comprehensive case-formulation: a narrative clinical integration of all information, detailing the specific interplay between traits and level of personality functioning.

Based on all the information collected, the following case-formulation was made.

Adam was a 39-year old man, referred by his GP for assessment and treatment evaluation. He presented with several persistent social and emotional problems. For his entire adult life, he had been unable to successfully complete an education or hold a job, which led to longstanding feelings of depression and anxiety (demoralization). To avoid feelings of failure, helplessness and hopelessness, he had adopted a socially withdrawn lifestyle, primarily seeking refuge in online gaming. The present AMPD assessment suggested to us that these problems were rooted in severe self-esteem issues. Indeed, Adam held an extremely vulnerable self-concept, alternating between grandiose self-expectations and severe self-defeating tendencies. On the one hand, he stated a deep conviction of being destined for "something special" and endorsed high standards. On the other hand, the anticipated failure triggered strong negative feelings in him that he was unable to confront, which led to flight in phantasy and extensive social withdrawal. Interpersonally, he was extremely sensitive to

rejection and slights (especially with superiors), and therefore was heavily invested in pleasing others, meeting their expectations as best he could, by taking on a submissive and overly friendly stance. However, this relational position had built up frustration and anger, because of unmet needs for recognition and admiration. He did not appear to understand his impact upon other persons and felt like he "was getting a raw deal from others". This realization triggered strong aversive feelings in him, leading him to either withdraw or to have emotional outbursts that interfered with cooperating with others. His understanding of his pattern of interpersonal involvement was quite limited, which left him confused and highly arousable.

Step 7: Provide oral and written feedback to patient and professionals
The final step of this procedure is to share the case formulation and diagnostic information with the patient and with colleagues involved in the follow-up care. Focus should be on the interplay between traits and impaired functioning, and to collaboratively build a narrative description that will help the patient make sense of his personality functioning. In our experience elements of Therapeutic Assessment (Finn, 2007; Kamphuis & Finn, 2018) are compatible with the AMPD model and can be used to structure the feedback session. Patients are more inclined to accept and integrate assessment information when the assessor starts with information that matches or is close to their self-concept (Finn, 2007; Kamphuis & Finn, 2018). Both the convergences and discrepancies across self-, informant-, and clinician-rated instruments can inform us on the (expected) optimal sequence in which to present the results from the AMPD assessment. If the case-formulation allows for it, it is best practice to start with issues on which self-report, informant-report and clinical ratings converge.

In Adam's case, we started with his self-reported reason for referral and history of presenting problems. His primary concerns were his inability to complete studies or hold jobs, and the associated feelings of failure, chronic stress and demoralization. We discussed how his inability to attain his goals was linked to his vulnerable self-esteem

and suggested to him how his withdrawal and emotional avoidance served to protect him from being emotionally overwhelmed by failure. Next, we linked this withdrawal to his feelings of self-loathing and how it also contributed to his anhedonia. We then introduced a finding that was a bit more discrepant from Adam's self-concept: underneath his feelings of self-loathing, he also seemed to harbor very high (grandiose) expectations for himself, which seemed to feed his fear of failure. A more tentative, not-knowing stance would be appropriate for discussing the findings that are most difficult to integrate for the client (in Adam's case how his pleasing and submissive stance was a way to control others, his 'blind spot' for the impact of his behavior on others). Empathy and ample validation are important inputs for fostering acceptance of these highly personal (and in part novel and discrepant) findings. For example, we helped Adam to an initial understanding of how aspects of his developmental history (most notably his father's verbal abuse, and the severe bullying in primary school) had rendered him extra vulnerable to impaired self-esteem, and we validated how he had tried to solve these emotional issues as best he could by adopting high internal standards, and by pleasing and controlling others; but also how this strategy had left him demoralized and depleted.

Next, specific areas of attention for treatment were discussed. We explained that treatment might help Adam confront his fear of failure and enhance his ability to tolerate the associated emotions. With the support of therapy, Adam might process his emotional injuries instead of his current coping strategy of extensive social withdrawal.

Discussion

This case study illustrated how different methods and different sources of information collection may serve to yield a comprehensive picture of the nature and degree of the patient's personality pathology. In some domains, Adam's self-reported personality problems aligned

with informant reports and clinical ratings based on structured interviews. These are the topics for which feedback is most readily integrated. However, notable areas of discrepancy also emerged, most likely because of Adam's limited introspective ability. More specifically, Adam did not fully grasp how his relational (submissive) stance actually invited the sort of interpersonal injuries he felt unable to cope with. This observation may highlight a clinically important issue regarding the assessment of Criterion B, which in research is predominantly questionnaire-based. It may well be that certain maladaptive areas remain unidentified when relying on self-report only. Our hypothesis is that especially in severe personality pathology, meaningful discrepancies may occur between self-report and clinical ratings. In fact, as argued many years ago by Grove & Tellegen (1991) sometimes the discrepancies may be the most informative pieces of evidence. Evaluating the correspondence between self- and informant report ratings is an interesting topic for future research. Marked discrepancies may point to potentially diagnostic limitations in self perception, and as such, serve to identify targets for treatment interventions (Hopwood & Bornstein, 2014). The interplay between Criteria A and B in conducting clinical AMPD model evaluations warrants discussion from the perspective of clinical utility. Some have argued that the conceptual distinction is blurry, and that Criteria A and B have poor incremental validity relative to one another (Bastiaansen et al., 2013; Bastiaansen et al., 2016; Berghuis et al., 2014). These observations call into question whether the model might be further reduced. However, these discussions tend to ignore aspects of clinical utility. Our case-analysis illustrated how Criteria A and B are intertwined and how information derived from both criteria added to a comprehensive clinical understanding of this patient's pathology. In this particular case, Criterion A information was essential to capture the severity of Adam's level of functioning; i.e., the severity of impairment of his self-esteem, and his pervasive inability to constructively connect to other people. It explained his strong tendency to avoid and withdraw from interpersonal contact, in order not to be overwhelmed by self-

esteem injury and the associated uncontrollable aggression. Without Criterion A it would be difficult to fully describe the severity of these impairments. Criterion B on the other hand, detailed in what ways his emotion regulation fell short (e.g. Emotional Lability, Hostility) and how he dealt with it interpersonally (e.g. Grandiosity, Attention Seeking, Submissiveness, Withdrawal). Along these lines, Huprich (2018) hypothesized that traits may ultimately be thought of as defenses against unpleasant ideas and motives rooted in Criterion A. From a different perspective: Criterion B can alert the clinician to the patient's relational style in therapy, while Criterion A captures the pervasiveness and rigidity with which this relational style will be expressed. Our patient may withdraw when he feels anxious or injured in personal interactions but reflecting the severity of his personality dysfunction, this withdrawal may not only be emotional, but also concrete: he may simply no longer show up. A case formulation like this, in which different aspects of personality functioning and traits are logically related and explained, is potentially more informative for treatment planning (than a summary of behavioral symptoms) and can offer patients a narrative understanding of their personality problems. Likewise, the AMPD model lends itself for giving feedback to patients in understandable, non-stigmatizing language, facilitating empathy in clinicians and fostering alliance early on.

The AMPD assessment can provide important clues for treatment planning. The specific areas and severity of dysfunction direct us to relevant interventions, in line with the integrated approach to treatment of PDs from Clarkin and colleagues (2015). For example, in the case of Adam, Criterion A points to his extreme sensitivity and inability to tolerate even minor injury. It alerts us that therapists will have to approach him with great sensitivity, and be extremely supportive and validating, continuously monitoring for any, even very small ruptures within the therapeutic relationship. Confrontations should be deferred for quite some time, and subsequently be done with great caution. Criterion B informs us on how Adam may approach the therapeutic relationship: he may initially defer and present himself rather

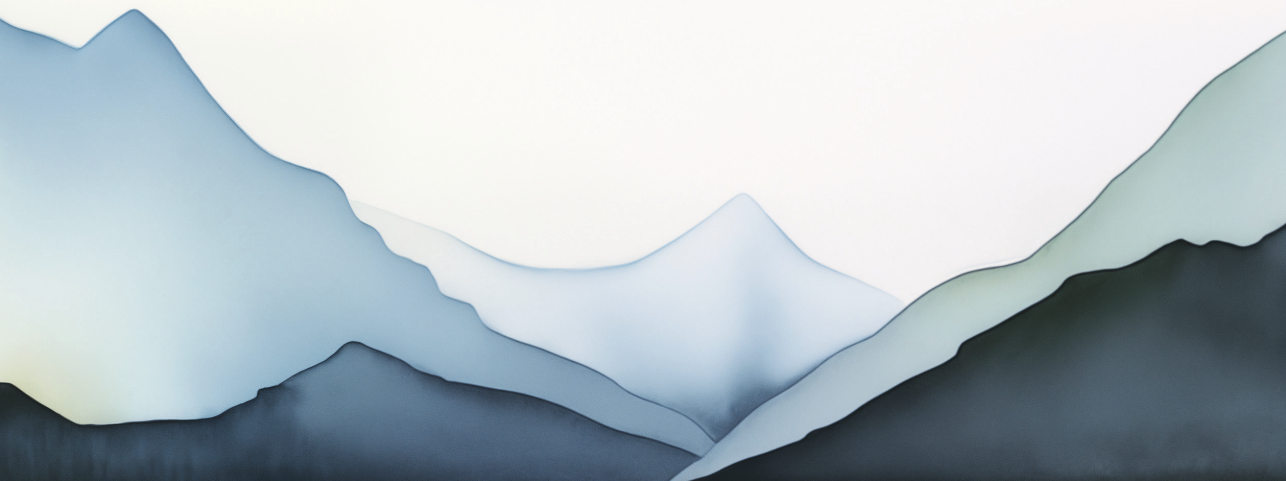
submissively, while trying to control negative feelings that may be stirred up in the therapy sessions. His affect may be initially somewhat detached, and he may be quite reluctant to disclose vulnerable emotions. However, establishing alliance with him is fraught with danger, as his increased self-revelations also increase the probability of feeling misunderstood or slighted, which in turn may lead to anger, displays of superiority, or even withdrawing completely. Building alliance will be an important goal in treatment and will probably take a considerable amount of time. For Adam, the therapeutic relationship may over time serve to help him better understand how he affects others (AMPD; Impact).

Finally, we want to highlight an interesting issue regarding the classification of narcissistic personality disorder. The section II PD criteria for NPD have been criticized for capturing only the grandiose types of narcissism (Levy, 2012; Skodol et al., 2014). In this respect, it is worth noting that Adam met criteria for NPD according to the AMPD, but did not according to the traditional section II PD criteria. Adam's narcissistic disturbance may be best conceptualized as a form of 'vulnerable' or 'covert' narcissism (Cain et al., 2008; Pincus et al., 2015). Hallmark of this type of narcissistic pathology is that the patient holds latent grandiose ideas, but initially expresses predominantly an avoidant personality style. Over time, the grandiose ideas become more overt as the therapeutic relationship deepens. The case of Adam thus illustrates that the AMPD assessment can detect covert narcissism as well.

Conclusion

As illustrated in this case analysis, we presented what we deem to be a clinically feasible multi-method, multi-informant procedure for assessing personality pathology according to the AMPD model. Within the scope of 3 hours of face-to-face assessment time, using readily available standardized instruments, the clinician can integrate the

MI-MM data into a comprehensive case-formulation that can readily serve both shared decision making with the patient and treatment planning. We anticipate that substantial aspects of this clinical procedure may be generalized to the ICD-11 classification of PDs, which soon awaits all WHO member countries.



Chapter 6

Client and clinical utility of the assessment of personality disorders

Laura C. Weekers, Joost Hutsebaut, Jan H. Kamphuis
Journal of Nervous and Mental Disease (2021), 209, 846-850

Abstract

Clinical utility and client utility are important desirable properties when developing and evaluating a new classification system for mental disorders. This study reports on four focus groups followed up by a Delphi study among clinicians working with clients with personality disorders (PD) and clients with PD themselves to harness both user groups' perspectives on the utility of PD diagnosis. Our findings show that the client and clinician views of the concept of utility were closely aligned and include aspects of transparency of communication and the ability of an assessment to enhance hope, curiosity, motivation, and insight into a client's personality patterns. Unique to clinicians' appraisal was the ability of an assessment to capture both vulnerabilities and resilience of clients and to give information about the prognosis in treatment. Unique to clients' appraisal was the ability of an assessment to be destigmatizing and collaborative. These findings may serve to expand our definition and measurement of clinical utility, in that collaborative and nonstigmatizing procedures likely promote client acceptability. To capture both aspects, we offer two preliminary questionnaires (i.e., item sets open to further empirical testing) based on the data derived from the Delphi procedure.

Introduction

Clinical utility has been identified as a top priority for personality disorder (PD) assessment, in both the new Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5) (American Psychiatric Association, 2013) and ICD-11 (World Health Organization, 2019) classification systems (First, 2005; Keeley et al., 2016; Reed, 2010; Skodol & Bender, 2009). First et al. (2004) defined clinical utility in terms of five core diagnostic functions of a diagnostic system: a) the way through which diagnostic entities are being conceptualized, b) the way through which clinically useful information is communicated to relevant others, c) the ease of use of the diagnostic categories and criteria, d) the extent to which the diagnostic system enables choosing effective interventions to improve clinical outcomes, and, finally, e) the capacity of the diagnostic system to predict and anticipate future clinical management needs.

DSM-5 (American Psychiatric Association, 2013) and ICD-11 (World Health Organization, 2019) PD classifications have shifted toward dimensional models of classification. However, several authors expressed concerns with respect to the clinical utility of the proposed dimensional models in DSM-5 and ICD-11 (Clarkin & Huprich, 2011). In fact, these concerns were a principal consideration in relegating the DSM-5 alternative model for personality disorders (AMPD) to Section III (Emerging Measures and Models) instead of Section II. At present, several empirical studies have documented superior clinician ratings of clinical utility of dimensional models relative to categorical models (Hansen et al., 2019; Lowe & Widiger, 2009.

Interestingly, although other aspects of the new classification systems of DSM-5 and ICD-11 have been studied extensively (e.g., reliability and validity; see Zimmermann et al., 2019, for a comprehensive overview), only few studies have focused on clinical utility

(Bornstein & Natoli, 2019; Milinkovic & Tiliopoulos, 2020). Clinical utility is comprehensively defined and measured. A six-item questionnaire, developed by Samuel and Widiger (2006), was used in several empirical studies (Hansen et al., 2019; Lowe & Widiger, 2009; Morey et al., 2014; Mullins-Sweatt & Widiger, 2011), and Kotelnikova and Clark (Kotelnikova Y, Clark LA Clinical Utility Rating Form [unpublished]) designed a 14-item questionnaire. Both instruments were developed "top-down" by expert clinical researchers based on the above-mentioned definition, which included acceptability, communication, ease of use, and value for treatment planning. Although these measures are definitely useful, it may prove beneficial to also enlist the perspective of the ultimate users of the assessment: the PD client. Therefore, we conducted two focus group procedures, followed by a Delphi study, to design an inductive definition and associated measures of utility of PD assessment based on input from professionals (i.e., clinical utility) and clients (i.e., client utility).

Methods

Participants

Participants—clinicians and clients—were initially recruited at Viersprong, a mental health care facility specializing in the assessment and treatment of adolescents and adults with PDs. A second group of clinicians and a second group of clients were recruited from other mental health care institutions to retrieve additional information. As no significant new information emerged from these groups, we considered the input to be "saturated," as in sufficiently comprehensive and representative.

Between December 2019 and March 2020, three live focus groups were organized: two focus groups with clinicians and one focus group with clients who had completed PD treatment. Because of the COVID-19 crisis, we had to reconsider the format of the second focus group and therefore asked them for written input. The first clinician focus

group consisted of seven clinicians working at the Viersprong. Their clinical experience ranged from 3 to 32 years (mean, 14.07; *SD*, 9.90), and they had been trained in a variety of treatment modalities, that is, mentalization-based treatment (MBT), schema-focused therapy (SFT), cognitive-behavioral therapy, transactional analysis, psychodynamic psychotherapy, and dynamic interpersonal therapy. The second focus group consisted of five experienced clinicians working in five other mental health care facilities specialized in treating PDs. Their clinical experience ranged from 20 to 32 years (mean, 26.20; *SD*, 4.82). Their theoretical background was also diverse; that is, these clinicians were trained in (one or more of) psychoanalytic psychotherapy, SFT, MBT, transference-focused psychotherapy, dialectical behavioral treatment, and/or indicated an eclectic orientation. Of the 12 participants, 8 responded to the subsequent Delphi rounds.

The first client focus group consisted of three clients who had completed their treatment at Viersprong and were recruited through the client board. Their ages ranged from 29 to 63 (mean, 42.33; *SD*, 18.15). Two female clients were included who had been treated for borderline PD with MBT (four, respectively, 10 years before the focus group), and a male client was included who had recently completed SFT for avoidant PD. Two of the three participants responded to the subsequent Delphi rounds. The second client group consisted of six clients (two female and four male), all treated for cluster C PD and recruited through the expertise center for PD (Kenniscentrum persoonlijkheidsstoornissen). Because of the onset of the COVID-19 crisis, no live focus groups could be organized, and their input therefore was collected in a written format. These clients were between 29 and 57 years old (mean, 48.33; *SD*, 10.50), and all responded to the first and second Delphi rounds. Five clients responded to the third Delphi round.

Procedure

The focus groups were organized as 2-hour sessions structured around a loose interview guideline and conducted by the two

primary researchers (L. W. and J. H.). The main objective was to define and operationalize "clinical utility and client utility of assessment procedures for personality disordered patients". The open group discussion was focused on the broad open-ended question "What makes an assessment procedure useful/helpful for you?" and was followed by more specific questions concerning clinical utility derived from the literature. More specifically, three aspects of clinical utility were explicitly checked for relevance: ease of use, communication, and treatment planning (First et al., 2004; Mullins-Sweatt & Widiger, 2009).

The focus groups were followed by a Delphi procedure to come to an agreement upon the definition and upon specific items for a "clinical and client utility" questionnaire in both groups (clients and clinicians). Based on the focus group input, both interviewers independently identified core themes and discussed until they agreed on the relevant themes for clinicians and clients separately. Participants from both samples then received an overview of identified themes and proposed definitions by e-mail and were asked to rate the degree to which they agreed with these constructs and definitions (completely disagree, disagree, agree, or completely agree). Constructs were revised when less than 75% of the participants agreed with the definition, and new feedback rounds were held until agreement met the 75% standard. After agreement on definitions, the researchers developed items to assess each of the core themes. These items were e-mailed to the participants and revised when there was less than 75% agreement on any given item. Again, feedback rounds were repeated until there was 75% agreement on all items.

Results

Part 1: Constituting of Client and Clinical Utility

Independent identification of core relevant themes from the clinicians' focus group discussion revealed a high level of convergence between

the primary researchers (L. W. and J. H.). Ninety percent of themes in the first focus group and 100% of themes in the second focus group were agreed upon. Disagreement was resolved by further discussing the themes. To design an accessible definition, the researchers set out to independently organize these themes into at most six overarching categories. Although wording was slightly different, the researchers were readily in agreement about six principal categories. The agreed upon categories that were deemed to constitute the clinicians' definition of clinical utility were a) process (subdivided into motivation and curiosity), b) insight in patterns, c) vulnerability/resilience (subdivided into severity of personality problems and resilience), d) prognosis, e) accessible language, and f) transparent communication.

The same procedure was used to infer common themes from both client focus groups. There was a high level of agreement between the researchers (five of six themes), and disagreement was resolved by further discussing the themes. The constructs deduced from the client focus groups were a) process (subdivided in to hope and motivation), b) insight in patterns, c) destigmatization, d) collaboration, and e) transparency.

Next, the appraisals of client and clinical utility, with their constituting core themes and related definitions, were sent back to both client and clinician groups separately to come to an agreement upon definitions. Both groups were asked for their agreement and for potential suggestions until sufficient consensus was reached, which required one round in both groups. The resulting definitions can be found in Tables 1 and 2.

Part 2: Preliminary Design of a Clinical Utility Questionnaire

Based on the agreed-upon definitions, the first two authors (i.e., L. W. and J. H.) formulated items that were deemed to capture as closely as possible the intended meaning of the pertinent constructs. Given the overlap between clinician's and client's definitions, we tried to formulate similar items if possible and appropriate. All items were presented in a Delphi procedure until sufficient (i.e., above 75%) consensus was

reached for each item. For the clinician questionnaire, agreement was achieved after the first feedback round. Some items were slightly altered based on the feedback provided by the participants. For the client questionnaire, a second feedback round was necessary to obtain sufficient consensus. Items for both questionnaires can be found in Tables 1 and 2. Of course, these items are in need of psychometric testing, and we provide these here as targets for future research.

Table 1 *Clinical Utility Definitions and Questions: The Clinician Perspective*

Construct	Definition	Items
Process-enhancing	The assessment starts a process in which clients begin to see their problems in a different light and get motivated for change in subsequent psychotherapy	1. The assessment stimulated the client to think more about the origin and background of his/her problems
		2. The assessment stimulated client awareness of what is needed to be able to change
		3. The assessment stimulated new insights and increased awareness in the client
		4. The assessment led the client to a better understanding of the core themes and their interrelatedness with respect to his/her problems
Curiosity	The client becomes curious about the origins of his/her symptoms and gets into an inquisitive, self-observing mode ('psychotherapy' mode)	5. The assessment stimulated the client to become more curious about the origin and interrelatedness of his/her problems
Motivation	The client becomes intrinsically motivated for treatment; the willingness to change is enhanced by the assessment	6. The assessment stimulated the client to become more willing to implement necessary changes for dealing with his/her problems
		7. The assessment stimulated the client to become more motivated to work on his/her problems in treatment

Table 1 *Clinical Utility Definitions and Questions: The Clinician Perspective* (continued)

Construct	Definition	Items
Core problem / patterns	The assessment generates information about the core of the client's problems and patterns, which allows for a coherent narrative of the client's history that integrates (often) seemingly diffuse or erratic problems and determines the focus for treatment	8. The assessment generated more clarity about the core of the client's problems
		9. The assessment generated a clear treatment focus
		10. The assessment clarified pervasive patterns in the clients' life history
Vulnerability/ resilience	The assessment provides a balanced view of both adaptive capacities and maladaptive characteristics of the patient	
Severity of personality problems	The assessment yields information regarding the severity of personality problems: e.g., defense mechanisms, ego-strength, level of identity integration, presence of (self-) destructive behavior	11. The assessment generated a clear indication of the severity of the personality problems
		12. The assessment clarified the nature of the client's vulnerabilities
Resilience	The assessment generates information on aspects of the client's adaptive or healthy functioning (e.g., mentalizing abilities, motivation to change, social network/ quality of interpersonal relationships)	13. The assessment clarified the client's adaptive potential and strengths
		14. The assessment clarified protective and adaptive factors in the client's environment

Table 1 *Clinical Utility Definitions and Questions: The Clinician Perspective* (continued)

Construct	Definition	Items
Prognosis	The assessment allows for predictions regarding: treatment, specifically to anticipate what the patient can tolerate in treatment, which interventions and therapeutic approach are likely to be helpful, what kind of critical interactional patterns can be expected, the probability of treatment success or failure (i.e., crisis or drop-out).	15. The assessment clarified which therapeutic approach and interventions are likely best suited in view of the client's coping ability
		16. The assessment allows for predictions regarding the probability of treatment success
		17. The assessment allows for predictions regarding possible pitfalls and risks the client may face during treatment
		18. The assessment allows for predictions regarding the nature of critical interactions between the patient and therapist, or group.
		19. The assessment allows for predictions regarding which therapeutic stance and interventions are helpful to the client
Accessible language	The results of the assessment as well as the interaction during the assessment are communicated in an accessible, readily understandable language. The assessment paints a vivid and concrete picture of the client.	20. The written report paints a clear, personal, and vivid picture of the client
		21. The written report is accessibly written and easy to understand
Transparent communication	The results of the assessment are communicated in a transparent way. The client receives all pertinent information from the assessment, and it becomes clear which parts of the diagnostic formulation are agreed upon by the client.	22. The results of the assessment are transparently shared
		23. It becomes clear which aspects of the clinical formulation the clinician and client agree and disagree on (if applicable)

Table 2 *Clinical Utility Definitions and Questions: The Patient Perspective*

Construct	Definition	Items
Destigmatizing	The assessment looks beyond the diagnosis and also allows for the person behind the diagnosis to be seen. As such, the client will recognize him/herself in the oral feedback and the written report. The assessment helps the client to not only see him/herself as merely a diagnosis, which enhances self-acceptance and reduces shame. The client is validated for the origins of the problems.	1. The written report showed the person behind the diagnosis, which helps me to not just 'be' the sum of my problems
		2. The written report described my problems in a respectful way
		3. The assessment helped me to better accept myself
		4. The assessment helped me to be less judgmental towards myself because of my problems and diagnosis
Process-enhancing – hope and motivation	The assessment allows the client to obtain insight into how patterns are related and strengthens the motivation and hope that treatment will help him/her to improve things. There is a focus on opportunities and potential change.	5. The assessment made me think more about the origin and background of my problems
		6. The assessment made me more aware of what is needed to change
		7. The assessment made me more curious about the origins of my problems as well as how they are interrelated
		8. I learned things about myself during the assessment that I was not clearly aware of before.
		9. The assessment made the primary themes behind my problems and my patterns of behavior clearer to me
		10. The assessment gave me hope that my current problems can change
		11. The assessment made me more motivated to work on my problems in treatment
		12. The assessment was not only focused on problems, but also on the potential to make positive changes

Table 2 *Clinical Utility Definitions and Questions: The Patient Perspective* (continued)

Construct	Definition	Items
Insight -core problem	The assessment generates insight into the core problems and serves the client to better understand him/herself. The assessment allows for the core problems to be discussed.	13. The assessment gave me more clarity about the core of my problems
		14. The assessment clarified recurring life patterns for me
		15. After the assessment it was clear to me what the focus of treatment should be
Collaborative	In the assessment the clinician and client work collaboratively, which instills in the client a sense of being understood and taken seriously; the clinician adjusts feedback to what the client can emotionally tolerate at that time.	16. During the assessment, there was a positive collaboration between the clinician and me
		17. During the assessment I felt I was taken seriously
		18. During the assessment, the clinician was attuned to my level of emotional tolerance
Transparent communication	The clinician is sincere and transparent about the assessment findings and their conclusions, and on the client's treatment prognosis.	19. The results of the assessment were shared in a transparent way
		20. The clinician explained what the conclusions were based on
		21. The clinician openly discussed with me which parts of the conclusion we agreed and (if applicable) we disagreed on
		22. The clinician discussed the expected result of treatment with me

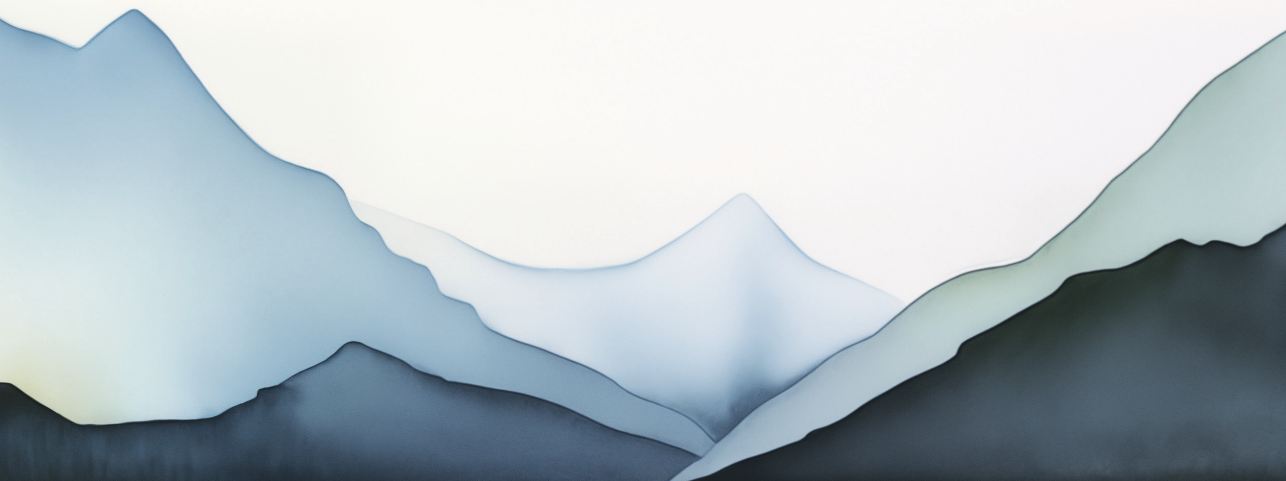
Discussion

In this study, we developed two preliminary questionnaires to capture the utility of PD assessments: a client and a clinician version. The item sets were based on discussion in four focus groups with subsequent Delphi rounds among clinicians with extensive experience working with clients with PD and (former) PD clients themselves. Interestingly, clients and clinicians agreed on several utility themes, with only few themes emerging that were unique to the client or clinician samples. Both groups highlighted the importance of transparency of communication and the ability to enhance hope, curiosity, motivation, and insight into patterns. Unique to clinicians' clinical utility definition was the ability of an assessment to capture both vulnerabilities and resilience of clients with PD and to give information about the prognosis in treatment. Unique to clients' clinical utility operationalization was the ability of an assessment to be destigmatizing and collaborative.

Several themes that emerged from the focus groups were similar to the clinical utility definitions from the existing literature, such as the importance of easily understandable language and transparent communication (communication; First et al., 2004), and the importance of prognostic information that can be used to select effective intervention (treatment planning; First et al., 2004). However, there were also several aspects nominated that add to the extant clinical utility definitions. For example, both clients and clinicians emphasized the importance of engaging the client (collaboratively) and emphasized the ability of an assessment to start a process in which the client becomes curious about him/herself, hopeful, and motivated to change (see also Kamphuis and Finn, 2018, for a discussion on epistemic trust in therapeutic assessment). Furthermore, clients highlighted the role an assessment can play in destigmatizing the client. Possibly, these aspects were highlighted given the specific nature of our samples: engaging clients, enhancing motivation, decreasing negative self-images (stigma), and increasing self-understanding may be especially

pivotal in clients with PD within a psychotherapeutic setting and therefore be specifically highlighted.

We believe the strength of the current definitions (and the proposed preliminary questionnaires) derives from the user-informed construction, using clients' and clinicians' appraisals of utility of assessment combined with information from the existing clinical utility literature. These client appraisals may serve to expand the more traditional definition and measurement of clinical utility in that collaborative and nonstigmatizing procedures likely promote client acceptability, which is crucial to the efficacy of any diagnostic procedure. Some caution is warranted with respect to the limits of the generalizability of our findings: all participating clients were or had been involved in psychotherapy



Chapter 7

Head to head comparison of the Alternative Model for Personality Disorders and Section II personality disorder model in terms of predicting patient outcomes one year later

Laura C. Weekers, Joost Hutsebaut, Jenneke M.C. Rovers, Jan H. Kamphuis

Personality Disorders: Theory, Research, and Treatment (2023)

Abstract

The present study investigated the predictive validity of Criterion A and B of the Alternative Model for Personality Disorders compared to the DSM-5 Section II personality disorder model in predicting patient outcomes one year after initial assessment, in a hetero-method longitudinal design. A clinical sample of 84 participants were administered both traditional Section II and AMPD interviews by two independent interviewers. One year after assessment, disability (WHODAS 2.0) and symptom severity (BSI) were assessed. The Section II PD model did not predict disability (R^2 = .01) nor symptom severity (R^2 = .03). The AMPD model on the other hand, predicted both disability (R^2 = .23) and symptom severity (R^2 = .29) one year post initial assessment. Both Criterion A and B were significant predictors, but when jointly combined only Criterion A remained significantly predictive of both disability and symptom severity while Criterion B did not. Criterion A thus appears to capture core vulnerabilities of personality disordered patients that are related to future functioning and symptom severity. Implications for clinical practice are discussed.

Introduction

Personality disorders (PDs) have traditionally been defined as enduring patterns of behaviors and experiences that express themselves in problems with emotion, cognition, impulse control, and interpersonal relatedness (American Psychiatric Association [APA], 2013). PDs impact strongly upon a person's functioning, affecting his or her liability for a range of mental disorders (Skodol et al., 2005), as well as reduced quality of life (Soeteman et al., 2008), life expectancy (Fok et al., 2012), and social and professional participation (Hastrup et al., 2019).

Although PDs are defined by their persistency and pervasiveness, longitudinal studies show a remarkable pattern of waxing and waning of specific PD symptoms (Zanarini et al., 2012). Zanarini and colleagues (2012) found that about half of the borderline PD patients no longer met diagnostic criteria over a two-year timeframe, calling into question whether the current criteria adequately capture enduring patterns of personality dysfunctioning. These findings may suggest that the criteria of PDs are more sensitive to temporary and situationally determined circumstances than would be suggested by the general definition (Hutsebaut et al., 2019). Possibly, the description of several criteria in terms of overt behavior and subjective experience may have made them more accessible for assessment but may also have induced state-like characteristics within the assessment of personality disorders (Wright et al., 2012). In turn, this may have contributed to a symptom-disability gap seen in longitudinal studies (Zanarini et al., 2012). Indeed, DSM-defined symptoms of PDs seem more flexible and changeable than the associated social and professional disability (Lenzenweger et al., 2004). Put differently, the current DSM-5 classification of PDs may not sufficiently capture the persistency and pervasiveness presumed to be inherent to personality disorders and their associated outcomes, as evidence suggests limited ability to predict long-term functional outcomes of individuals.

The Alternative Model for Personality Disorders (AMPD; APA, 2013) was proposed to address some of the shortcomings of the traditional symptom-based model. By defining PDs in terms of relatively stable underlying impairments in personality functioning (self- and interpersonal functioning) on the one hand, and in terms of maladaptive personality traits on the other hand, the AMPD model aims to capture the more enduring aspects of personality pathology (Zachar et al., 2015). For example: whereas the traditional model identifies the 'frantic efforts to avoid real or imagined abandonment' as a feature of borderline PD, the AMPD model instead defines these clinical issues in terms of impaired capacities to experience and tolerate emotional closeness (Criterion A), expressed in increased anxiety when experiencing separation (Criterion B). The later definition may allow this impairment to be still reflecting an individual's personality pathology, even in the absence of clinging behavior (e.g., when a person may avoid intimate relationships with other people).

Moreover, by describing personality pathology in relation to underlying impairments and traits, the AMPD criteria may be less sensitive to contextual influences and therefore better align with the core enduring vulnerability that may conceptually be characteristic of individuals with PDs. As the AMPD model focuses more on stable internal impairments and dispositions than on contextual effects, it may be better suited to explain and predict disability in the long term (see also Mulder & Tyrer, 2019). Indeed, underlying impairments in self- and interpersonal functioning as well as maladaptive traits will likely render individuals vulnerable for impaired social- and occupational functioning as well as the development of symptoms such as depression, anxiety, or somatization. When an individual has, for example, difficulties regulating emotions, low self-esteem, and deals with experienced insecurity in emotionally close relationships by socially isolating this will likely be associated with poor long-term outcomes such as inability to work, loneliness, depression, anxiety and somatic complaints.

Zimmerman and colleagues (2019) provided an excellent overview of the research on the AMPD model, including studies investigating the incremental validity of the AMPD model compared to the traditional Section II model of PDs. It has been shown that Criterion A predicts current social functioning, proposed treatment intensity, and estimated prognosis over and above categorical Section II diagnosis (Morey et al., 2013; Morey & Benson, 2016). Criterion B has shown incremental validity over the Section II PD diagnosis regarding general PD severity (Fossati et al., 2016), disability (Chmielewski et al., 2017), psychosocial functioning (Fowler et al., 2017; Simms & Calabrese, 2016), symptom severity (Fowler et al., 2017), social cognition deficits (Fossati et al., 2017), aggression (Somma et al., 2019), and predicting treatment planning (Morey & Benson, 2016). However, some negative findings were also reported. Creswel and colleagues (2016) did not find incremental validity of Criterion B when predicting problematic alcohol use. Furthermore, several studies assessing the incremental validity of Criterion A and B relative to each other have been conducted with mixed results. With regard to predicting Section II PDs, most studies found incremental validity of Criterion B compared to Criterion A (e.g. Anderson & Sellbom, 2018; Sleep et al., 2019; Sleep et al., 2020). Some studies, however, demonstrated incremental validity of Criterion A over B when predicting some Section II PDs (Cruijt et al., 2019; Wygant et al., 2016). Incremental validity of Criterion A compared to Criterion B has also been demonstrated for general functioning (Huprich et al., 2018), wellbeing, and symptom severity (Hutsebaut & Bach, 2018). Several studies found equal predictive validity of both Criteria (Few et al., 2013; Roche et al., 2018).

There are several limitations to these studies. First, all were based on self-report measures, like the widely used Personality Inventory for DSM-5 (PID-5) for the assessment of Criterion B. Moreover, to the best of our knowledge, only a few studies used an interview-based assessment of Criterion A to assess the level of personality functioning. Within a large group of respondents with PD diagnoses, Christensen and colleagues (2020) found that the level of personality

functioning as assessed by the Structured Clinical Interview for DSM-5 AMPD module I (SCID-AMPD-I) correlated with social disability (measured with the Work and Social Adjustment Scale and Global Assessment of Functioning). Moreover, level of personality functioning predicted social disability better than the sum of DSM-IV PD criteria did, indicating stronger predictive validity of the AMPD model. Second, none of these studies followed a longitudinal design, which makes it impossible to rule out that observed associations are confounded by the current state of participants. To evaluate the predictive potency of both the AMPD and the Section II PD model more stringently, it is essential to study longitudinal associations with long-term outcomes.

The present study therefore investigated the extent to which Criteria A and B of the AMPD model predict disability and symptom severity one year later. We utilized well-established self-report and interview measures to assess both criteria. Moreover, we compared the predictive validity of the AMPD model to the DSM-5 Section II PD model, as assessed by structured clinical interviews. Based upon the hypothesis that the AMPD model may better capture the structural personality pathology, we expected that the AMPD model would show stronger predictive validity than the Section II diagnoses in predicting disability and symptom severity after one year. Finally, we explored whether Criteria A and B have incremental validity relative to one another.

Method

Procedure and participants

This study was approved by the ethics committee of the University of Amsterdam. Prior to data collection a power analysis was conducted. With a power of .80, $\alpha = .05$, and two predictor variables in the linear regression model at least 68 participants were needed to detect the expected medium effect size ($f^2 = .15$).

Participants were adults who were referred for treatment to De Viersprong, a mental health-care institution specialized in the assessment and treatment of personality disorders. Participants in the present study had participated in a prior study (Weekers et al., 2021) and had indicated in the informed consent that they were willing to fill out questionnaires one year after their initial assessment. Of the 189 participants in the first study, 137 gave informed consent for follow-up assessment. All participants were contacted one year after their initial assessment and 84 filled out the questionnaires. There were no differences between the responders and nonresponders on the Section II PD variables (number of PD classifications (*d* = .23), number of PD criteria met (*d* = .24). There were significant differences between the responders and non-responders on the AMPD variables (Semi-structured Interview for DSM-5 Personality Functioning [STiP-5.1] total score – *d* = .40 and Structured Clinical Interview for DSM-5 Alternative Model for Personality Disorders, Module II [SCID-AMPD-II] total score – *d* = .36) , with responders having slightly lower (healthier) scores on these variables than nonresponders.

The standard admission procedure, consisting of the Structured Clinical Interview for DSM-5 Personality Disorders (SCID-5-PD; First et al., 2016 translated by Arntz et al., 2017) and the Structured Clinical Interview for DSM-5 Syndrome Disorders (SCID-5-S; First et al., 2018; Dutch translation by Arntz et al., 2018) was supplemented with a second interview session with a second interviewer. This second interviewer independently administered AMPD interviews as part of the assessment. These were the STiP 5.1 (Hutsebaut et al., 2017) and the SCID-AMPD-II (First et al., 2018). The SCID-5-PD interviews were administered by twenty interviewers and the AMPD interviews were administered by eight interviewers. Interviewers were trained psychologists working at De Viersprong. They were blind for information from the other interviewer. The data was collected between November 2018 and November 2020, part of the data was collected via online meetings because of COVID-19.

One year after initial assessment participants were asked to complete two questionnaires; the Brief Symptom Inventory (BSI; Derogatis, 1975) and the World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0; Üstün et al., 2010). They also indicated whether they had received treatment during the last year. Participants received the questionnaires via a secure email.

Table 1 *Clinical characteristics of the sample*

	N	*%*
Syndrome disorders		
Anxiety Disorders	15	17.9
Mood Disorders	32	38.1
Somatization Disorders	2	2.4
Eating Disorders	5	6.0
Substance use Disorders	9	10.7
Any Syndrome Disorder	45	53.6
Personality disorders		
Avoidant PD	16	19.0
Obsessive-compulsive PD	2	2.4
Narcisistic PD	4	4.8
Borderline PD	11	13.1
Other specified PD	28	33.3
Any PD	53	63.1
Number of PD criteria (*M, SD)*	5.9	2.9
WHODAS 2.0 total (*M, SD)*	30.7	20.0
BSI (*M, SD)*	1.2	0.8
STiP 5.1 mean total (*M, SD)*	1.9	0.6
SCID-AMPD-II mean total (*M, SD)*	0.8	0.3

Note. PD = personality disorder; WHODAS 2.0 = World Health Organization Disability Assessment Schedule 2.0; BSI = Brief Symptom Inventory; STiP 5.1 = Semistructured Interview for DSM-5 Personality Functioning; SCIDAMPD-II = Structured Clinical Interview for DSM-5 Alternative Model for Personality Disorders, Module II.

Sample characteristics

The final sample consisted of 84 participants, 62 of whom were female (73.8 %). The average age at assessment was 36.04 (*SD* = 11.76, range 19 – 65). Table 1 presents clinical information of the sample and descriptive statistics of all variables. Most participants met criteria for at least one DSM-5 personality disorder (63.1%), with other specified PD being the most prevalent.

Measures

The Structured Clinical Interview for DSM-5 Personality Disorders (SCID-5-PD; First et al., 2016; translated by Arntz et al., 2017) is a structured diagnostic interview for the assessment of the DSM-5 PDs (formerly SCID-II; First et al., 1997). To meet criteria, target behaviors must be pathological, pervasive, and persistent to be scored as present. No reliability or validity data is available for the SCID-5-PD. However, the SCID-5-PD is based on the SCID-II for which validity and reliability have been demonstrated in PD samples (Carcone et al., 2015; Lobbestael et al., 2011; Maffei et al., 1997; Weertman et al., 2003). The number of full PD classifications and the number of total PD criteria met were used in the present study to operationalize Section II PD.

The Structured Clinical Interview for DSM 5- Syndrome Disorders (SCID-5-S; First et al., 2018; Dutch translation by Arntz et al., 2018) is a structured diagnostic interview for the assessment of the DSM-5 syndrome disorders (formerly SCID-I; First et al., 1997). To date, no research on reliability and validity of the Dutch translation has been conducted, however good interrater reliability and validity were demonstrated in an American sample (Shankman et al., 2018).

The Semi-structured Interview for Personality Functioning DSM–5 (STiP 5.1, Hutsebaut et al., 2017) is a clinical interview based on the Level of Personality Functioning Scale (LFPS) as described in the Alternative Model of Personality Disorders in DSM-5, which was developed

to provide an estimate of the level of personality functioning in a standardized manner (Hutsebaut et al., 2017). The scale is a continuous scale, with higher scores indicating more severe disturbances in personality functioning. The interview yields a total severity score ranging from 0 (no impairments) to 4 (extreme impairments), and severity scores for each domain (Self- and Interpersonal functioning), element (identity, self-direction, empathy, intimacy) and facet of the LPFS separately. Internal consistency of the STiP-5.1 was excellent in an adult sample, with Cronbach's alpha of .97 for the total scale, and .94 for both the Self-Functioning and Interpersonal Functioning domains. Interrater reliability was not assessed in the present sample. Interrater reliability was good in a previous study, with intraclass correlation coefficients (ICCs) ranging from .58 to .81 in a clinical sample (Hutsebaut et al., 2017). In the present study the total STiP 5.1 score, calculated as the mean of all 12 facets, was used to operationalize Criterion A.

The Structured Clinical Interview for DSM-5 Alternative Model for Personality Disorders, Module II

(SCID-AMPD-II; First et al., 2018) is a structured interview, designed by the APA work group to assess maladaptive personality traits (Criterion B). A total of 25 traits are assessed within five trait domains: Negative affectivity, Detachment, Antagonism, Disinhibition, and Psychoticism. Traits are scored on a scale from 0 (very little or not at all descriptive) to 3 (very descriptive). Interrater reliability was not assessed in the present sample. A recent study, however, showed good interrater reliability with ICC's ranging from .54 to .89 for trait scores and .79 to .92 for domains scores (Somma et al., 2020). The mean of all facets of the SCID-AMPD-II was used to operationalize Criterion B.

The Brief Symptom Inventory

(BSI) was used to assess symptom severity (Derogatis, 1975; translated by de Beurs, 2011). The BSI is a self-report questionnaire containing 53 items (across 9 dimensions) that provides an overview of symptoms of

psychopathology in adults. The total BSI score is calculated by adding the severity of all symptoms (each coded on a 5-point Likert scale ranging from 0 (not at all) to 4 (extremely)). The Cronbach's α was high (α = .96) in this sample.

The World Health Organization Disability Assessment Schedule-2 (WHODAS-2; Üstün et al., 2010) is a generic assessment instrument for health and disability, which comprises 36 items with a recall period of 30 days. The WHODAS-2 contains six domains: Cognition, Mobility, Self-care, Getting along, Life activities, and Participation in society. The participants have response options from 1 (no difficulty) to 5 (extreme difficulty or cannot do). The total score is computed by summing all item responses to a summary score which ranges from 0 to 100, with higher scores indicating higher levels of disability. For participants with no occupation nor school activities, the total score was calculated without the Life activity items. In addition, two questions concerning treatment status were included to assess whether participants had received treatment the previous year, and whether they had completed at least 50% of the treatment program. Cronbach's α was high (α = .94) in the present sample.

Statistical analysis

First, the issue of missing data was addressed. The percentage of missing data ranged from 1.2% (for the BSI) to 8.3% (for treatment status), and 13 participants had missing data on one or more variables. For the self-report questionnaires missingness on the total scales was due to failure to complete multiple items on the questionnaire. For the interviews reasons for missingness were unknown. There were no significant differences between the participants with complete data and incomplete data regarding age and gender. We used multiple imputation to create and analyse 10 multiply imputed datasets. Incomplete variables were imputed under fully conditional specification, using the default settings of the mice 3.0 package (Van Buuren & Groothuis-Oudshoorn, 2011). The analyses were conducted

in each imputed dataset separately and combined using Rubin's rules. For comparison, we also performed the analyses using listwise deletion, with the 71 complete cases, and found no major differences.

As we hypothesized treatment status might influence the outcome variables, we explored the relationship between treatment status and WHODAS 2.0/BSI by means of a linear regression analysis (Model 1). Second, the relationship between the traditional section II model and WHODAS 2.0/BSI was assessed with the number of Section II PD diagnoses (Model 2a), and the number of Section II PD criteria (Model 2b) as predictors. Third, the relationship between the AMPD model and WHODAS 2.0/BSI was assessed with Criterion A (Model 3a) and Criterion B (Model 3b) as predictors. Furthermore, hierarchical linear regression analyses were performed comparing Model 3a (Criterion A) to Model 3c (Criterion A + Criterion B) and Model 3b (Criterion B) to Model 3c to assess the incremental validity of Criterion A and B relative to the other. For all models with more than one predictor adjusted R^2 are reported. Pearson correlations between all predictors and outcomes are presented in Table 2. As might be expected, the WHODAS 2.0 and BSI were highly related (r = .80).

Transparency and Openness

We report how the sample size for the present study was determined, how missing data was handled, all manipulations, and all measures in the study and Journal Article Reporting Standards were followed (Appelbaum et al., 2018). All data, analysis code, and research material is available upon request by contacting the first author. Data were analysed using RStudio. This study was not preregistered.

Table 2 *Pearson Correlations between Predictors and Outcomes*

Variable	Treatment status	Section II PDs	Section II PD criteria met	Criterion A	Criterion B	WHODAS
Treatment status						
Section II PDs	-0.13					
Section II PD criteria met	-0.09	0.80				
Criterion A (STiP 5.1)	-0.29	0.40	0.43			
Criterion B (SCID-AMPD-II)	-0.12	0.40	0.46	0.68		
WHODAS total scale	-0.21	0.09	0.05	0.48	0.36	
BSI total scale	-0.27	0.17	0.18	0.52	0.44	0.81

Note: PD = personality disorder; WHODAS = World Health Organization Disability Assessment Schedule; STiP 5.1 = Semistructured Interview for DSM-5 Personality Functioning; SCID-AMPD-II = Structured Clinical Interview for DSM-5 Alternative Model for Personality Disorders Module II; BSI = Brief Symptom Inventory.

Results

Treatment status

Treatment status, defined as having completed at least 50% of a treatment program in the previous year, significantly predicted symptom severity (BSI), albeit with a small effect-size (F = 5.53, p = .020, R^2 = .07). Those who had completed 50% or more of a treatment program had a lower symptom severity than participants who completed less than 50% of a treatment program (see Table 3). Treatment status did not significantly predict disability (WHODAS 2.0; F = 3.16, p = .076, R^2 = .04; see Table 4). As percentage of explained variance was low for both the BSI and WHODAS 2.0, we decided not to control for treatment status in the subsequent analyses.

Traditional (Section II) PD Model

The traditional Section II PD model did not predict disability (WHODAS 2.0) one year post initial assessment; neither the number of PD diagnoses (F = .61, p = .440, R^2 = .01) nor the number of PD criteria (F = 0.19, p = .670, R^2 = .003) significantly predicted disability one year after assessment (see Table 4). For symptom severity one year after assessment (BSI total score) a similar picture emerged: neither the number of PD diagnoses (F = 2.22, p =.136, R^2 = .03) nor the number of PD criteria (F = 2.40, p = .120, R^2 = .03) predicted symptom severity one year after assessment (see Table 3).

Alternative Model for Personality Disorders

Both Criterion A (F = 18.32, p < .001, R^2 = .23), and Criterion B (F = 10.82, p = .001, R^2 = .13) significantly predicted disability (WHODAS 2.0) one year after the initial assessment, such that a higher level of impairment in personality functioning (Criterion A) was associated with a higher level of disability one year post assessment, and a higher level of pathological personality traits (Criterion B) was associated with a higher level of disability one year post assessment (see Table 4). For symptom severity (BSI) we found similar results. Both Criterion A (F = 23.69, p < .001, R^2 =.27), and Criterion B (F = 17.65, p < .001, R^2 = .19) were significant predictors of symptom severity one year after assessment; again, higher level of impairment in personality functioning (Criterion A) and a higher level of pathological personality traits (Criterion B) were associated with higher symptom severity one year post assessment (see Table 3).

Incremental validity of Criterion A relative to Criterion B

We then conducted hierarchical regression analyses with Criterion A in the first model (3a) and Criterion B added in the second model (3c; see Table 3 and 4). Criterion A remained a strong predictor of disability (b = 15.69, p = .011), while Criterion B did not significantly predict disability when the effect of Criterion A was taken into account (b = 3.26, p = .737, ΔR^2 = .00). Also, for symptom severity Criterion A remained a strong

predictor (b = 0.56, p = .008), while Criterion B did not significantly predict symptom severity when the effect of Criterion A was taken into account (b = 0.35, p = .302, ΔR^2 = .00). Furthermore, when starting with Criterion B in the first model (3b) and adding Criterion A in the second model (3c) the proportion of explained variance increased both for disability (ΔR^2 = .08) and symptom severity (ΔR^2 = .08).

Table 3 *Impact of Section II PD and Criterion A and B of the AMPD on Symptom Severity (BSI)*

Model	*b*	*SE B*	*t*	*p*
Model 1 BSI (R^2 = .07; 95% CI = .002 – .22)				
Constant	1.38	0.11	12.72	<.001
Treatment status	-0.42	0.18	-2.39	.020
Model 2a BSI (R^2 = .03; 95% CI = .003 – .14)				
Constant	1.07	0.13	7.99	<.001
Number of Section II PDs	0.20	0.13	1.49	.139
Model 2b BSI (R^2 = .03; 95% CI = .002 – .15)				
Constant	0.96	0.19	5.04	<.001
Number of Section II PD criteria	0.05	0.03	1.56	.123
Model 3a BSI (R^2 = .27; 95% CI = .11 – .46)				
Constant	-0.09	0.27	-0.34	.734
Criterion A	0.70	0.14	4.96	<.001
Model 3b BSI (R^2 = .19; 95% CI = .06 – .36)				
Constant	0.40	0.21	1.94	.006
Criterion B	1.02	0.24	4.27	<.001
Model 3c BSI (R^2 = .27; 95% CI = .11 – .45)				
Constant	-0.12	0.27	-0.44	.660
Criterion A	0.56	0.20	2.79	.008
Criterion B	0.35	0.34	1.04	.302

Note. PD = personality disorder; AMPD = Alternative Model for Personality Disorders; BSI = Brief Symptom Inventory; CI = confidence interval.

Table 4 *Impact of Section II PDs and Criterion A and B of the AMPD on Disability (WHODAS 2.0)*

Model	*b*	*SE B*	*t*	*p*
Model 1 WHODAS (R^2 = .04; 95% CI = .00 – .18)				
Constant	33.70	2.86	11.80	<.001
Treatment status	-8.62	4.76	-1.81	.075
Model 2a WHODAS (R^2 = .01; 95% CI = .01 – .09)				
Constant	28.16	3.55	7.94	<.001
Number of Section II PDs	2.82	3.57	0.79	.432
Model 2b WHODAS (R^2 = .003; 95% CI = .03 – .07)				
Constant	28.25	5.08	5.56	<.001
Number of Section II PD criteria	0.36	0.78	0.46	.648
Model 3a WHODAS (R^2 = .23; 95% CI = .07 – .41)				
Constant	-1.47	7.63	-0.19	.848
Criterion A	16.92	3.96	4.27	<.001
Model 3b WHODAS (R^2 = .13; 95% CI = .02 – .29)				
Constant	12.79	5.68	2.25	.027
Criterion B	21.76	6.54	3.33	.001
Model 3c WHODAS (R^2 = .21; 95% CI = .06 – .40)				
Constant	-1.75	7.58	-0.23	.818
Criterion A	15.69	5.85	2.68	.011
Criterion B	3.26	9.65	0.34	.737

Note. PD = personality disorder; AMPD = Alternative Model for Personality Disorders; WHODAS 2.0 = World Health Organization Disability Assessment Schedule 2.0; CI = confidence interval.

Discussion

The current study investigated the predictive validity of the traditional Section II PD model and AMPD model, in a hetero-method longitudinal design, by assessing disability and symptom severity one year after initial assessment. The Section II PD model did not predict disability (WHODAS 2.0), nor symptom severity (BSI); not at

the level of diagnoses, nor at the criterion count level (i.e., number of PD criteria met). The AMPD model, on the other hand, was a strong predictor of both disability and symptom severity one year after the initial assessment. When analysed separately, Criterion A and B were both significant predictors of disability and symptom severity. When taken the effect of Criterion A into account however, Criterion B was no longer a significant predictor of disability and symptom severity. Furthermore, treatment status was not related to disability but showed a small effect on symptom severity.

These results are in line with previous research showing the incremental validity of the AMPD model over the Section II PD model in terms of psychosocial functioning. In an early study, Morey and colleagues (2013) demonstrated incremental validity of Criterion A over Section II PD in predicting current psychosocial functioning. Of note, the study by Morey and colleagues (2013) was a retrospective study based on clinician-ratings. Using interview data to assess Criterion A (SCID-AMPD module I), Christensen and colleagues (2020) also found incremental validity of Criterion A over the Section II PD model in predicting current psychosocial functioning, as did Fowler and colleagues (2017). Contrary to our findings however, all of these studies reported small but significant correlations between Section II PD and psychosocial functioning. Together, these findings suggest that Section II PD appears to be related to current psychosocial functioning, but not to future disability, at least in a relatively homogeneous sample of PD patients.

The present results could provide preliminary evidence for the fundamental conceptual difference between the Section II PD and AMPD model. Section II PD features seem to capture specific behavioral and symptomatic outcomes of impaired personality functioning, while the AMPD rather seems to capture the impairments themselves in terms of underlying vulnerability and an associated range of outcomes. The first appears less stable than initially intended – also reflected in several longitudinal studies showing low to moderate

stability of Section II PD diagnoses - while the latter appears to capture more the enduring vulnerability that should conceptually probably be inherent to the notion of PDs. In a sense, one could argue that the AMPD model reconceptualizes PDs as a failure to achieve several universal developmental tasks (e.g. stable sense of self, establish intimate relationships with others; Livesley, 1998) and by doing so, captures more a broader and more stable disposition for a number of negative outcomes. Of note, symptom severity and disability were highly related, with a higher level of disability being associated with a higher level of symptom severity. This makes theoretical and intuitive sense as both constructs are related to the burden of disease. Analyses were done for each of the outcomes separately, but were highly similar in pattern of associations.

Although our study demonstrated substantial overlap between Criterion A and B, with a correlation of r = .68, our study showed that Criterion A had incremental validity over Criterion B with respect to the prediction of both symptom severity as well as disability. Previous studies showed mixed results. For example, Bach and Hutsebaut (2018) found incremental validity of Criterion A over B in predicting wellbeing and symptom severity, while Ohse and colleagues (2022) found incremental validity of Criterion B over A in predicting psychosocial functioning (WHODAS). Of note, the results of Ohse and colleagues (2022) might be biased in favor of Criterion B by a method factor, as both Criterion B and psychosocial functioning were measured by self-report while Criterion A was measured by an interview. Our study was the first to adhere to a longitudinal design, allowing for prediction over time. Our results suggest, in line with the AMPD model, that Criterion A is a stronger indicator of general severity than Criterion B. Criterion A is thus a useful tool for assessing core vulnerabilities and predicting future functioning and disability.

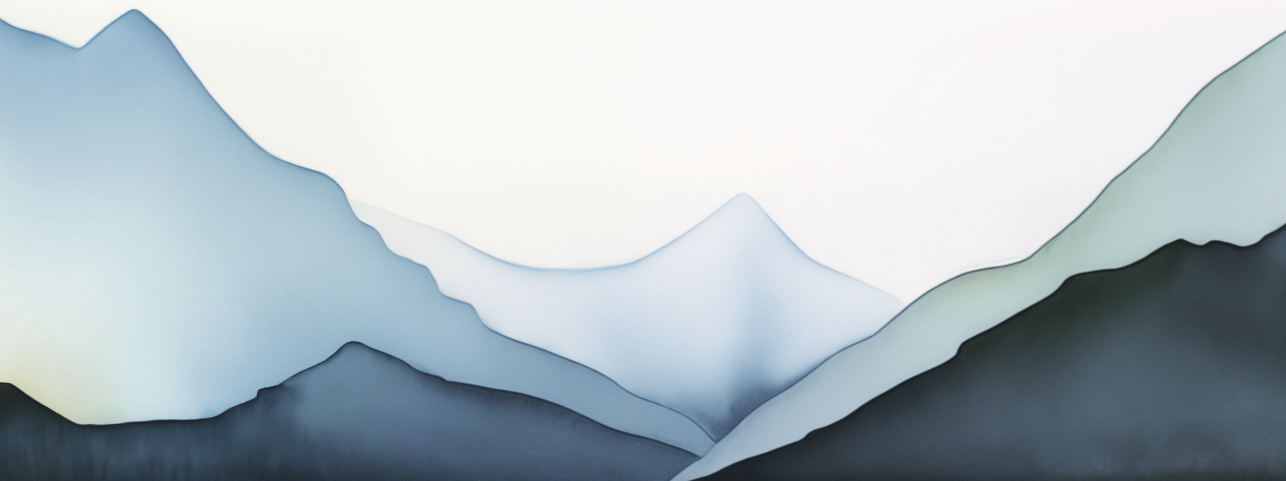
These results may put some research findings on the changeability of personality disorders into perspective and contribute to the debate on the stable and changeable components of personality pathology. Treatment outcome studies have supported the notion

that (Section II) features of PD are responsive to psychotherapy (see e.g., Cristea et al., 2017), leading to the optimistic stance that PD are 'curable'. However, longitudinal studies suggest that remission from (borderline) PD features may be much easier to obtain compared to full psychosocial recovery (Zanarini et al., 2012). Our findings suggest that long term psychosocial disability may be underpinned by chronic impairments in personality functioning. Whereas classic PD symptoms may be more sensitive to actual situational factors, impairments in personality functioning may be capturing more the long-term core vulnerability characteristic of personality pathology. Change in these latter features of personality pathology may be more predictive of full future recovery. Implications for treatment and research may be substantial. If treatment success implies psychosocial recovery, treatment progress may have to be monitored in terms of changes in personality functioning, while remission of PD symptoms may only have limited relevance regarding sustained social and occupational recovery. In addition, treatment assignment and planning may rather be informed by severity of personality functioning than by severity in terms of classic PD diagnosis or features. For example, patients with more severe impairments in personality functioning might need more intensive and specialized treatment, with a focus on improving self- and interpersonal functioning. Even more important: if the aim is to prevent future psychosocial impairment, e.g., in young persons, treatment assignment may be informed by impairments in Criterion A, rather than by the mere presence (or absence) of Section II PD features. These implications should be tested in future studies.

The current study has several notable strengths and limitations. A key strength was the longitudinal design, which allowed for a true comparison of predictive validity of both models. To our knowledge, the current study is the first to use such a design to compare the predictive validity of both diagnostic models. Furthermore, we found strong associations using a hetero-method longitudinal design, with clinician rated interviews at initial assessment and patient self-report one year later. Also, the interviews were administered by trained

professionals who had experience in working with PD patients. Several limitations should be mentioned. Our clinical sample does not cover the whole range of personality pathology. There was a small difference between the responders and non-responders in terms severity. Also, it is unclear whether our results generalize to, for example, patients with Cluster A or antisocial PDs. Furthermore, interrater reliability of the interview schedules was not assessed in the present sample. To mitigate this concern, previous studies have demonstrated adequate to good interrater reliability of all interviews employed in this study. Lastly, the present sample was rather homogeneous sample in that most patients were assessed with moderate or severe impairments in personality functioning. Future research should therefore replicate these results in a more heterogeneous (general psychiatric) sample.

Taken together, the current study showed that the AMPD model, especially Criterion A, has predictive validity with respect to both symptom severity and disability one year after initial assessment. Accordingly, it demonstrated incremental value of the AMPD model over the Section II PD model in predicting future disability.



Chapter 8

Comparing the clinical utility of the Alternative Model for Personality Disorders to the Section II personality disorder model: A Randomized Controlled Trial

Laura C. Weekers, Joost Hutsebaut, Hilde De Saeger,
Jan H. Kamphuis

Personality Disorders: Theory, Research, and Treatment (2024)

Abstract

The Alternative Model for Personality Disorders has been extensively studied over the past decade, but to date there is no direct comparison of the clinical utility of the AMPD model relative to the Section II personality disorder (PD) model in an ecologically valid design. The current study examined the clinical utility of an AMPD-informed assessment procedure and Section II PD assessment procedure as assessed by both patients and clinicians in a randomized controlled trial. A sample of 119 patients were randomly assigned to either an AMPD or a Section II PD assessment procedure. At the end of the assessment patients filled out questionnaires pertaining to clinical utility, satisfaction, motivation for treatment and general experience of the assessment. Clinicians who subsequently started treatment with these patients also completed two clinical utility questionnaires. There were no significant differences between the AMPD and Section II PD assessment procedure on patients' reported clinical utility, motivation for treatment, satisfaction and general experience of the assessment, nor were there significant differences between the models on clinician reported clinical utility. Explorative analyses revealed that, for patients, a positive relationship with the assessor was predictive of experienced utility. This study shows no superiority of the AMPD in terms of clinical utility, but suggests that the alliance with the assessor is a particularly salient factor in clinical utility.

Introduction

Diagnostic and Statistical Manual of Mental Disorders, fifth edition (DSM-5) introduced the alternative model for personality disorders (AMPD) in its Section III ("emerging measures and models"; American Psychiatric Association, 2013). It provides a major shift in the operationalization of personality disorders (PDs) by describing PDs as a combination of (a) impairments in self and interpersonal functioning (Criterion A) and (b) pathological personality traits (Criterion B). Since its publication, numerous studies concerning the reliability and validity of the model have been conducted (e.g. Zimmermann et al., 2019). Several self-report, informant-report, and interview-schedules have been constructed to reliably assess both Criterion A and B (e.g. First et al., 2018; Goth et al., 2018; Huprich et al., 2018; Hutsebaut et al., 2017; Krueger et al., 2012; Morey, 2017; Thylstrup et al., 2016; Weekers et al., 2019). Validity of the model has been demonstrated by meaningful associations with other measures of PD severity and related constructs (e.g. Zimmermann et al., 2019). Head-to-head comparisons of the AMPD model and Section II PD model are still scarce; however, they are needed to warrant such a major paradigm shift. The current study compares the AMPD- and Section II PD model in terms of clinical utility as rated by both patients and clinicians.

The Section II PD model has been criticized for its lack of reliability and validity (e.g. Krueger et al., 2014; Samuel, 2015; Skodol, 2014). Although reliability and validity touch upon the conceptual soundness of a diagnostic system, in clinical practice clinical utility may be at least as important. Clinical utility refers to the extent to which a diagnostic system assists professionals in fulfilling five core diagnostic functions (First et al., 2004): (a) conceptualizing diagnostic entities, (b) communicating clinical information to relevant others, (c) using diagnostic categories and criteria sets in clinical practice (including for diagnostic interviewing and differential diagnosis), (d) choosing effective interventions to improve clinical outcomes, and (e) predicting

future clinical management needs. Limited clinical utility of the Section II PD model has been demonstrated by the fact that clinicians do not use the Section II PD model as intended by the manual (Keeley et al., 2013; Keeley et al., 2015), for example by using theoretical assumptions to weigh the importance of different PD criteria (Kim & Ahn, 2002). Furthermore, the utility of the Section II PD model for assessing adolescents and old age has been questioned, with some criteria being less suited for these age groups (Videler et al., 2019). Lastly, the Section II model has a risk of stigmatizing patients, by differentiating arbitrarily between abnormal and normal categories of personality and by using concepts that have become connotative, like narcissistic and histrionic personality disorder (Waugh, 2019).

Research on clinical utility of the AMPD model is rather limited and has almost exclusively focused on utility from a clinician perspective. An overview has been provided by Bach and Tracy (2022). A meta-analysis showed clinicians favor dimensional PD models over categorical PD models, particularly in respect to its usefulness in communicating with patients, comprehensiveness in describing the patient's personality problems and usefulness in formulating therapeutic interventions (Bornstein & Natoli, 2019). Research on the utility of the AMPD model from a client perspective is even more scarce. One study by Lengel and Mullins-Sweatt (2017) demonstrated that computerized feedback concerning (mal)adaptive personality traits helped patients' understanding of their personality characteristics and problems in living. Several authors hypothesize that the AMPD model can improve utility of assessment for clients. For example, the use of understandable, less-jargonized language can enhance communication with clients and can potentially work to avoid stigma (Waugh, 2019). Feedback that is framed in experience near language might further promote a therapeutic working alliance from the start and enhance epistemic trust (i.e. trust in the authenticity and personal relevance of information from others; Sperber et al., 2010; Kamphuis & Finn, 2018).

Early studies on the clinical utility of the AMPD model bear promise but are also limited in many ways. Most importantly, the issue of clinical utility is almost exclusively approached from the perspective of the clinician, focusing on rather formal aspects like ease of use, or ease of communication. These studies did not assess an essential ingredient of clinical utility: how may an assessment help clients and professionals to come to a better understanding of the core problems, and how does this understanding inform treatment planning and enhance treatment readiness? Moreover, these studies make use of case vignettes (Garcia et al., 2018) or involve hypothetical assessments of already familiar patients (Morey et al., 2014). None of these studies represent personality assessment in daily clinical practice. In order to assess the AMPD model in terms of clinical utility, as compared to the Section II PD model, a comprehensive ecologically valid assessment procedure using established instruments and resulting in a case formulation that reflects the core aim of the classification system is needed. Also, clinical utility should be studied from the patients' as well as the professionals' perspective in order to establish its (assumed) clinical usefulness for both groups of stakeholders.
The current study constitutes a rigorous test of the clinical utility of the AMPD model compared to the Section II PD model in an ecologically valid randomized controlled trial (RCT). We employed well-established interview measures to assess the AMPD- and Section II PD criteria. Both patient- as well as clinician-rated utility were assessed. Based upon our review of the literature, we hypothesized that the AMPD assessment procedure would have superior clinical utility for both patients and clinicians.

Method

Design

This study is a RCT comparing the clinical utility of the AMPD assessment to the Section II PD assessment in a parallel group design.

Patients were randomly allocated to either the AMPD assessment or Section II PD assessment and outcome questionnaires were administered after the assessment was completed.

Ethics

This study was approved by the Medical Ethics Research Committee of the Academic Medical Centre, Amsterdam (NL75676.018.20).

Sample Size and Power Calculation

A priori sample size calculation was based on the clinical utility patient questionnaire, the primary outcome measure in this study. We conducted a short pilot study with 17 patients. All participants were assessed using the Section II PD assessment procedure. Internal consistency of the total clinical utility score was high (α = .88) and mean total score was 3.71 (*SD* = 0.48). Based on a prior study by De Saeger and colleagues (2014), in which therapeutic assessment was compared to a short structured motivational treatment, we expected an effect size of *d* = .50. Based on information from the pilot study, with a sample size of at least 64 per assessment procedure an effect size of *d* = 0.50 can be detected with two-sided testing, with α = .05 and a power of .80 (G-power; Faul et al., 2007). We therefore aimed at a sample size of *N* = 128.

Participants and Eligibility Criteria

Participants were treatment seeking adults who were referred to De Viersprong, a mental health care facility for the assessment and treatment of personality disorders. Inclusion and exclusion criteria were the same as the in- and exclusion criteria for assessment at De Viersprong. Inclusion criteria were: (suspected) personality pathology. Exclusion criteria were: a diagnosis of autism spectrum disorder, chronic psychotic disorder or organic brain disorder, and intellectual disability.

Procedure

Patients on the waitlist for initial assessment were contacted by one of the researchers and informed about the study. An information letter was sent after this call and a (virtual) meeting was arranged 1 week later to answer questions and sign informed consent. If patients consented, they were randomized into either the traditional Section II PD assessment or AMPD assessment. The randomization file was a block (4x2) randomization constructed by an independent statistician and managed by an employee that was not part of the research team. After randomization, participants were invited for the assessment procedure, the result of the randomization was not shared with participants. Patients who wanted to be informed about the assessment procedure were informed after the assessment and after they had filled out all questionnaires.

After the assessment procedure was complete and patients had received a written report of the assessment, they were contacted to fill out the questionnaires. If patients were referred for subsequent treatment within the institution, the treating therapist was asked to read the written report of the assessment and to complete a clinical utility questionnaire after their first contact with the patient. All data were collected between February 2022 and May 2023.

Assessment Procedures and Clinicians

Section II PD assessment consisted of a clinical interview, a structured interview session for assessing DSM-5 personality disorders (SCID-5-PD) and a feedback session informing patients about their diagnoses and treatment options. The AMPD assessment procedure consisted of a clinical interview, a structured interview session for assessing DSM-5 section III (AMPD) personality disorders (Semi-structured interview for personality functioning DSM-5 (STiP 5.1) and SCID-AMPD module II) and a feedback session informing patients about their (dimensional) diagnoses and treatment options. In both conditions, patients were additionally administered a structured clinical interview assessing major symptom disorders (SCID-5 clinician version) by a different

clinician. This instrument was included to have an assessment of other (comorbid) mental disorders needed for treatment planning in this real-life assessment procedure. Each assessment procedure was manualized (available upon request) and reviewed by an international expert (dr. A.E. Skodol). Prior to the start of the study, clinicians in both assessment procedures received a one-day training (including video demonstrations of the administration of the interviews) to familiarize them with the manual of their respective assessment procedure. The Section II PD training was led by two of the authors (Hilde De Saeger and Jan Henk Kamphuis) and the AMPD training was led by Laura Weekers and Joost Hutsebaut. Four clinicians were trained in the AMPD assessment procedure, and seven in the Section II PD assessment procedure. All clinicians had experience in working with PD patients. Hilde De Saeger supervised the Section II PD clinicians and Joost Hutsebaut supervised the AMPD clinicians; all patients were discussed, and adherence to the models was monitored during these sessions.

Adherence

To assess adherence to the assessment procedures, 20 written reports (10 AMPD and 10 Section II PD) were rated on adherence by raters who were not involved in the present study but had experience with PD assessment. We constructed an adherence scale (available upon request) based on the manuals of the assessment procedures. Rates were asked to rate individual elements of the written report on adherence and provide an overall rating of adherence ranging from 0 (*all elements are missing, not adherent at all)* to 4 (*exceeded adherence criteria*). For the AMPD assessment, the average adherence score was 3.40 (*SD* = 0.52), and 100% of the written reports were scored as adherent or exceeding adherence criteria. For the Section II PD assessment, average adherence score was 2.80 (*SD* = 0.92), and 70% of the written reports were scored as adherent or exceeding adherence criteria.

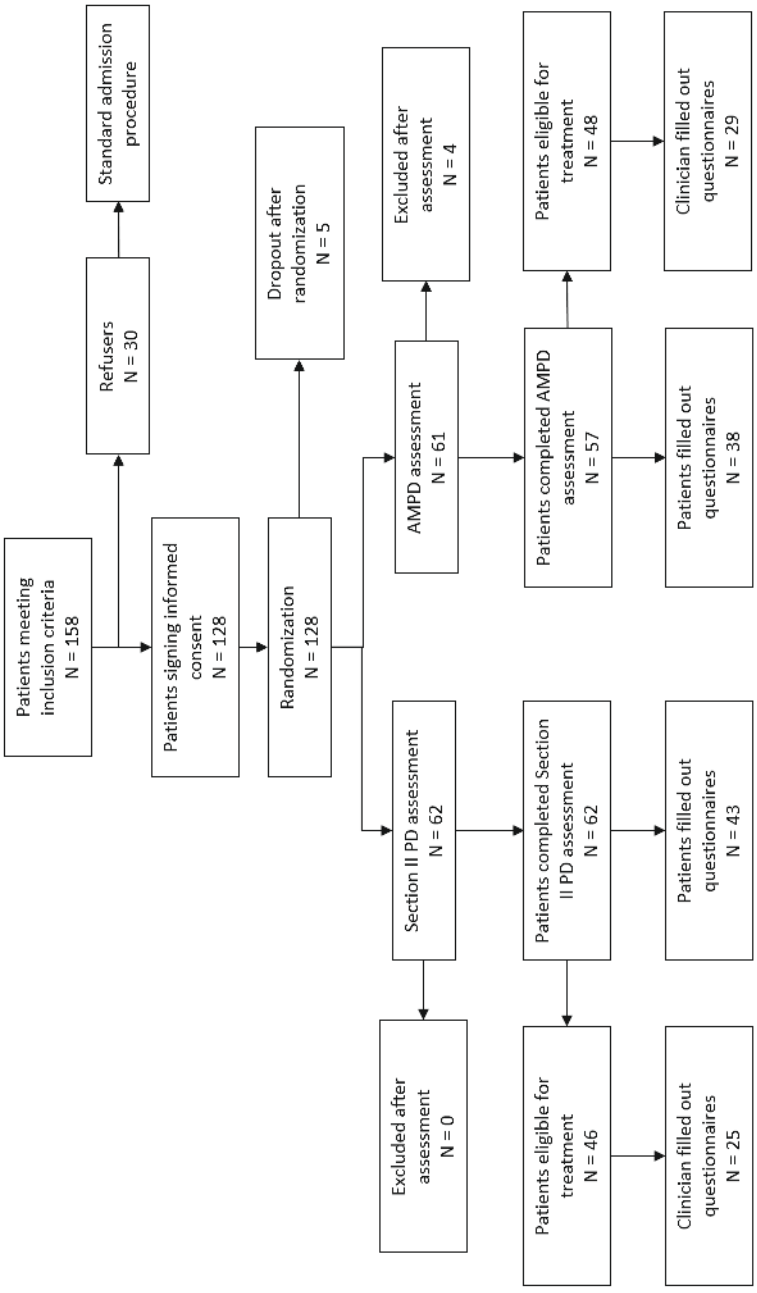


Figure 1
Flowchart of Patient Allocation and Dropout

Sample Characteristics

A total of 128 patients signed informed consent (see Figure 1). Nine participants were excluded because of varying reasons: four were excluded because later assessment revealed a total IQ score below 80, two were excluded because they never started the assessment procedure and sought treatment elsewhere, three were excluded because they were not available on days when the research assessments took place.

The final sample consisted of 119 participants. Age ranged from 19 to 60 years old (M = 36.16, SD = 11.23), 86 were female (72.3%). Of the 119 participants, 57 participants were randomized into the AMPD assessment procedure and 62 participants into the Section II PD assessment procedure. The average time from start to end of assessment was 26.96 days (SD = 18.18). There were no significant differences between the assessment procedures on age (d = .24), gender (φ = .17), or duration of assessment (d = .27). After assessment, 94 participants were eligible for treatment at De Viersprong (79.0%), their treating clinician was contacted to fill out clinical utility questionnaires. Table 1 presents clinical characteristics of the sample.

Table 1 *Clinical Characteristics of the Sample (N = 119)*

	AMPD assessment (N=60)		Section II assessment (N=59)	
Disorders	*N*	*%*	*N*	*%*
Syndrome disorders				
Mood Disorders	33	55.0	31	52.5
Anxiety Disorders	24	40.0	20	33.9
Eating Disorders	5	8.3	7	11.9
Somatization Disorders	3	5.0	5	8.5
Substance use Disorders	4	6.7	9	15.3
Any Syndrome Disorder	47	78.3	45	76.3

Table 1 *Clinical Characteristics of the Sample (N = 119)* (continued)

	AMPD assessment (*N*=60)		Section II assessment (*N*=59)	
Disorders	*N*	*%*	*N*	*%*
Personality disorders				
Avoidant PD	11	18.3	11	18.6
Obsessive-compulsive PD	3	5.0	3	5.1
Paranoid PD	-	-	1	1.7
Narcisistic PD	1	1.7	1	1.7
Borderline PD	20	33.3	19	32.2
Antisocial PD	0	0	2	3.4
Other/trait specified PD	25	41.7	22	37.3
Any PD	57	95.0	51	86.4

Note. PD = personality disorder; AMPD = Alternative Model for Personality Disorders.

Measures

Primary Outcome Measure

Clinical Utility Questionnaire – Patient version. A specific outcome instrument measuring clinical utility for patients was developed by our research group for the purpose of this study (Weekers et al., 2021). Given the lack of such an instrument, a focus group was organized to collect implicit patient knowledge on clinical utility of PD assessment. A group of patients was asked to brainstorm about the concept of clinical utility of assessment; this procedure was repeated with other patient groups until no new information arose (suggesting saturation). The resulting themes were described and returned to all participants in a Delphi procedure until sufficient consensus (at least 75%) was reached. Following up on consensus on the definition, specific items were formulated to assess the aspects of clinical utility that had emerged from the focus groups, until sufficient consensus was reached. We found that patients defined clinical utility of assessment as the ability of an assessment procedure to (a) be destigmatizing, (b) start a process in

which the patient starts to get more insight into patterns and become hopeful and motivated to change, (c) summarize the core patterns which underly the patients' problems, (d) collaboratively work with the patient, and (e) communicate transparently with the patient about the results of the assessment. The resulting 22 item questionnaire has to be rated on a 5-point Likert scale ranging from 1 (*completely disagree*) to 5 (*completely agree*). The resulting total and domain scores range from 1 to 5. A description of the definitions of each domain, including an example question can be found in Table 2. Internal consistency was good for the total score (α = .89), and acceptable for the domains "destigmatizing" (α = .74), "process enhancing" (α = .78), "core patterns" (α = .74), "collaborative" (α = .78), and "transparent communication" (α = .70).

Secondary Outcome Measures

Clinical Utility Questionnaire – Clinician version. In a similar way as for patients, we used several focus groups and a subsequent Delphi procedure to define clinical utility of PD assessment from a clinician perspective and to formulate items to assess each of these aspects. Clinicians defined clinical utility of assessment as the ability of an assessment procedure to (a) start a process in which the patient becomes curious about the problems he is facing and gets motivated to change, (b) summarize the core patterns which underly the patients' problems, (c) give a balanced view of both vulnerabilities and resilience, (d) make predictions (prognostic) useful in treatment (i.e. risks, expected treatment success, expected interactional patterns, useful treatment interventions), (e) use accessible and easy to understand language and paint a vivid picture of the patient, and (f) communicate transparently with the patient about the results of the assessment. The resulting 23-item questionnaire is rated on a 5-point Likert scale ranging from 1 (*completely disagree*) to 5 (*completely agree*). The resulting total and domain scores range from 1 to 5. A description of the definitions of each domain, including an example question can be found in Table 2. Internal consistency in the present sample was good for the total

score (α = .83) and the "prognostic" domain (α =.81), and acceptable for the domains "process-enhancing" (α =.79), "core patterns" (α = .69), "vulnerability and resilience" (α = .71), and "transparent communication" (α = .68). The internal consistency of the domain "accessible language" was poor (α = .36); this domain was excluded from the analyses.

Expectancy for Future Treatment Scale (EFTS). The EFTS is a one-item visual analogue scale for patients to rate their expectancy regarding future treatment ("To what extent do you believe this intervention will benefit your future treatment?").

Client Satisfaction Questionnaire (CSQ; Larsen et al., 1979). For assessing general satisfaction with the assessment procedure questions from the CSQ (Larsen et al., 1979) were used. The questionnaire consists of 8 items, rated on a 4-point Likert scale. Internal consistency in the present sample was good with α = .89, and comparable to previous research in a Dutch sample (de Wilde & Hendriks, 2005).

Assessment Questionnaire (AQ; Finn et al., 1994). The AQ was used to assess different aspects of patient's experience of the assessment procedure. The AQ is a 48-item self-report questionnaire, with a 5-point Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). It has been used in research concerning therapeutic/ collaborative assessment (Holst et al., 2009; Allen et al., 2003). The questionnaire consists of a total score and four factors: new self-awareness/ understanding, positive accurate mirroring, positive relationship with the examiner, and negative feelings about the assessment. The total and domain scores range from 1 to 5. Internal consistency was good for the total AQ score (α = .88) and ranged from acceptable (α = .74) to good (α = .86) for the domain scores.

Motivation for Treatment Questionnaire (MTQ-8; van Beek & Verheul, 2008). Motivation was assessed by MTQ-8, an eight-item self-report questionnaire composed of two factors: need for help and readiness to change. Internal consistency of the MTQ-8 is fair to good, with α's ranging from .63 to .77.

Clinical Utility Scale Clinicians. Morey and colleagues (2014) developed a 6-item questionnaire to assess different aspects of clinical utility as rated by clinicians. Five of these items were relevant for the present study (Items 2-6) and were used as a secondary outcome to enhance comparability with other international studies. The items were rated on a visual analogue scale ranging from 0 (*not useful*) to 100 (*extremely useful*).

Table 2 *Clinical Utility Definitions and Example Items*

Clinical utility construct	Definition	Example item
Patient questionnaire		
Destigmatizing	The assessment looks beyond the diagnosis and also allows for the person behind the diagnosis to be seen. As such, the client will recognize him/herself in the oral feedback and the written report. The assessment helps the client to not only see him/herself as merely a diagnosis, which enhances self-acceptance and reduces shame. The client is validated for the origins of the problems.	The written report showed the person behind the diagnosis, which helps me to not just 'be' the sum of my problems
Process enhancing – hope and motivation	The assessment allows the client to obtain insight into how patterns are related and strengthens the motivation and hope that treatment will help him/her to improve things. There is a focus on opportunities and potential change.	The assessment gave me hope that my current problems can change
Core problem/ patterns	The assessment generates insight into the core problems and serves the client to better understand him/herself. The assessment allows for the core problems to be discussed.	The assessment clarified recurring life patterns for me

Table 2 *Clinical Utility Definitions and Example Items* (continued)

Clinical utility construct	Definition	Example item
Collaborative	In the assessment the clinician and client work collaboratively, which instills in the client a sense of being understood and taken seriously; the clinician adjusts feedback to what the client can emotionally tolerate at that time.	During the assessment I felt I was taken seriously
Transparent Communication	The clinician is sincere and transparent about the assessment findings and their conclusions, and on the client's treatment prognosis.	The results of the assessment were shared in a transparent way
Clinician questionnaire		
Process-enhancing	The assessment starts a process in which clients begin to see their problems in a different light and get motivated for change in subsequent psychotherapy	The assessment stimulated the client to think more about the origin and background of his/her problems
Core problem/ patterns	The assessment generates information about the core of the client's problems and patterns, which allows for a coherent narrative of the client's history that integrates (often) seemingly diffuse or erratic problems and determines the focus for treatment	The assessment generated more clarity about the core of the client's problems
Vulnerability/ resilience	The assessment provides a balanced view of both adaptive capacities and maladaptive characteristics of the patient	The assessment clarified the nature of the client's vulnerabilities
Prognosis	The assessment allows for predictions regarding: treatment, specifically to anticipate what the patient can tolerate in treatment, which interventions and therapeutic approach are likely to be helpful, what kind of critical interactional patterns can be expected, the probability of treatment success or failure (i.e., crisis or drop-out).	The assessment clarified which therapeutic approach and interventions are likely best suited in view of the client's coping ability

Table 2 *Clinical Utility Definitions and Example Items* (continued)

Clinical utility construct	Definition	Example item
Accessible language	The results of the assessment as well as the interaction during the assessment are communicated in an accessible, readily understandable language. The assessment paints a vivid and concrete picture of the client.	The written report paints a clear, personal, and vivid picture of the client
Transparent communication	The results of the assessment are communicated in a transparent way. The client receives all pertinent information from the assessment, and it becomes clear which parts of the diagnostic formulation are agreed upon by the client.	The results of the assessment are transparently shared

Statistical Analyses

First, missing data were examined. Of the 119 patients, 37 (31.1%) did not fill out any questionnaires after the assessment. Reasons for not responding were unknown. Comparisons between the responders and non-responders showed no significant differences on number of syndrome disorders (d = .09), PD status (d = .08), age (d = .22), or gender (φ =.01). Of the 82 patients that did fill out questionnaires, seven had partially missing data on one or more questionnaires. Of the 94 clinicians that were contacted to fill out questionnaires, 54 responded (57.4%). Of these 54 clinicians, two had partially missing data on the questionnaires. Multiple imputation was used to create and analyse 30 multiply imputed datasets. Incomplete variables were imputed under fully conditional specification, using the default settings of the mice 3.0 package (Van Buuren & Groothuis-Oudhoorn, 2011). The analyses were conducted in each imputed dataset separately and combined using Rubin's rules. For comparison, we also performed the analysing using listwise deletion, and found no major differences.

Independent samples t-tests were performed to assess differences between the assessment groups on all measures. As an exploratory

part of our research question, we assessed the relationship between "positive relationship with the examiner" (AQ domain) and clients' clinical utility, satisfaction, and motivation for treatment ratings using linear regression analysis.

Transparency and Openness

Sample size calculations for the present study were reported, as well as an account of how missing data were handled. All measures and statistical analyses were reported, and Journal Article Reporting Standards were followed (Appelbaum et al., 2018). All data, analysis code, and research material are available upon request by contacting Laura C. Weekers. Data were analyzed using RStudio. This study was preregistered at trialsearch.who.int (Identifier NL9191).

Results

Primary Outcome: Clinical Utility Questionnaire Patients

There were no significant differences between the Section II PD and AMPD assessment on the total clinical utility score nor on the subscales (i.e., destigmatization, process enhancing, core patterns, collaborative, and transparent communication; see Table 3). Patients in the AMPD assessment did not rate the assessment as more useful than patients in the Section II PD assessment did.

Secondary Outcomes: Patient Questionnaires

No significant differences between the Section II PD and AMPD assessment were found for expectancy for future treatment (EFTS), satisfaction (CSQ), motivation for treatment (MTQ-8), and general experience of the assessment (AQ; Table 3).

Secondary Outcomes: Clinical Utility According to Clinicians

There were no significant differences between the Section II PD and AMPD groups on total clinical utility score as rated by clinicians

nor were significant differences observed between the groups on subscales of the clinical utility questionnaire (i.e., process enhancing, core problems, vulnerability and resilience, prognostic, and transparent communication; see Table 4). Moreover, on the Morey clinical utility questions, no differences were found between the groups, however several (non-significant) trends emerged in favor of the AMPD. Clinicians rated the AMPD model and Section II PD model as equally useful in terms of communicating with other professionals, communicating with the client, comprehensiveness, describing the global personality of the patient and treatment planning (see Table 4).

Table 3 *Independent Samples t-tests for the Patient Questionnaires (N = 119)*

	Section II M(SD)	AMPD M(SD)	*t* value	*p* value	Cohen's *d*
CUQ Total	3.69 (0.40)	3.66 (0.52)	-0.32	.750	0.07
CUQ Destigmatizing	3.73 (0.70)	3.63 (0.78)	-0.78	.438	0.14
CUQ Process enhancing	3.54 (0.67)	3.50 (0.64)	-0.38	.703	0.06
CUQ Insight	3.35 (0.78)	3.48 (0.97)	0.79	.433	0.15
CUQ Colllaborative	4.37 (0.62)	4.29 (0.72)	-0.61	.546	0.12
CUQ Transparant communication	3.69 (0.71)	3.69 (0.73)	0.05	.963	0.00
EFTS	67.70 (25.03)	68.95 (24.68)	0.27	.789	0.05
CSQ Total	24.48 (4.90)	25.44 (5.28)	1.04	.302	0.19
AQ Total	3.38 (0.40)	3.41 (0.36)	0.35	.725	0.08
AQ New self-awareness	3.14 (0.53)	3.10 (0.50)	-0.48	.635	0.08
AQ Positive Accurate Mirroring	3.09 (0.58)	3.08 (0.66)	-0.03	.974	0.02
AQ Positive Relationship	3.64 (0.56)	3.73 (0.56)	0.75	.458	0.16
AQ Negative Feelings	2.29 (0.86)	2.21 (0.70)	-0.57	.572	0.10
MTQ Total	50.00 (7.19)	49.70 (7.67)	-0.22	.827	0.04

Note. CUQ = Clinical Utility Questionnaire; EFTS = Expectancy for Future Treatment Scale; CSQ = Client Satisfaction Questionnaire; AQ = Assessment Questionnaire; MTQ = Motivation for Treatment Questionnaire.

Table 4 *Independent Samples t-tests for the Clinician Questionnaires (N = 94)*

	Section II M(SD)	AMPD M(SD)	*t* value	*p* value	Cohen's *d*
CUQ Total	3.62 (0.40)	3.55 (0.36)	-0.80	.426	0.19
CUQ Process enhancing	3.48 (0.52)	3.27 (0.61)	-1.84	.071	0.37
CUQ Core patterns	3.62 (0.85)	3.66 (0.93)	0.25	.802	0.05
CUQ Vulnerability & resilience	3.68 (0.81)	3.51 (0.80)	-1.02	.314	0.21
CUQ Prognostic	3.41 (0.72)	3.59 (0.64)	1.25	.216	0.27
CUQ Transparant communication	3.92 (1.02)	3.73 (1.08)	-0.80	.429	0.18
How useful do you feel this PD model is for communicating information about this individual with other mental health professionals?	78.36 (16.78)	78.11 (16.78)	-0.08	.935	0.02
How useful are these concepts for comprehensively describing all the important personality problems the individual has?	76.18 (19.31)	83.57 (18.83)	1.78	.084	0.39
How useful do you feel this PD model is for communicating information about the individual to him or herself?	73.81 (18.69)	79.95 (17.85)	1.65	.106	0.34
How useful is this PD model for helping you to formulate an effective intervention for this individual?	68.22 (21.12)	75.53 (21.85)	1.59	.119	0.34
How useful is this PD model for describing the individual's global personality?	66.37 (24.81)	72.91 (20.22)	1.39	.172	0.29

Positive Relationship and Clinical Utility

Positive relationship with the examiner (AQ) was a significant predictor of the total clinical utility score, $F = 10.82$, $p = .001$, $R^2 = .12$. Patients who reported a more positive relationship with the examiner rated their assessment as more useful. Furthermore, positive relationship with the examiner was a small but significant predictor of satisfaction ratings (CSQ-8), $F = 4.46$, $p = .04$, $R^2 = .06$. Positive relationship was not predictive of expectancy for future treatment, $F = 0.55$, $p = .554$, $R^2 = .01$, nor for motivation for treatment, $F = 0.46$, $p = .460$, $R^2 = .01$.

Discussion

The present study compared the clinical utility of the AMPD model to the Section II PD model as rated by both patients and clinicians in an ecologically valid randomized controlled trial. No differences between the models were observed in terms of patient-rated clinical utility or clinician-rated clinical utility. Moreover, no differences were observed between the models for motivation for treatment, satisfaction, or expectancy for future treatment. We did find a significant relationship between positive relationship with the examiner and patient-rated clinical utility and satisfaction, with patients who reported a more positive relationship with their assessor rating the assessment as more useful and more satisfying.

Our results are partially in line with but, with regard to some aspects, also markedly different from previous research on clinician-rated clinical utility. Morey and colleagues (2014) compared the clinical utility of Criterion A and B of the AMPD model separately with the Section II PD model and reported superior clinical utility ratings for the Section II PD model on ease of use and communication with other professionals, while no differences were found on clinical utility ratings between the Section II PD and Criterion A for communicating with patients, comprehensiveness, treatment planning, and global descriptive ability. Criterion B was superior compared to the Section II PD model on all

clinical utility ratings except professional communication. Bornstein and colleagues (2019) also demonstrated in a meta-analysis that traits (AMPD or Five Factor Model) were generally rated by clinicians as more useful than the Section II PD model. Although we did not find significant differences between the models, a consistent trend in favor of the AMPD model was observed for several clinical utility questions Morey and colleagues (2014) formulated (with effect sizes around 0.30 - 0.40) pertaining to communication with the patient, comprehensiveness, treatment planning, and global descriptive ability, which might suggest slight preference in terms of utility for the AMPD assessment. Of note, power analysis was based on the patient variables, but our study was slightly underpowered to detect significant differences in clinicians, as not all patients were allocated for subsequent treatment.

Previous studies are different from the present study in several respects. In our study the complete AMPD assessment was rated on clinical utility instead of separate elements of the model. Also, the clinicians rating the assessment model did not perform the assessment themselves, but rated the utility of the assessment after reading the written report and seeing the patient for the first time. Furthermore, with our RCT design, clinicians were asked to rate only one model on experienced clinical utility and were not able to compare the assessment models relative to each other. The study by Morey and colleagues (2014) asked clinicians to rate already familiar patients on both models and compare the utility of these models. Our design constitutes a more stringent test of the clinical utility of the AMPD model.

To our knowledge the present study is the first to compare the Section II PD and full AMPD assessment in terms of patient-rated clinical utility and experience of the assessment. Contrary to our expectations, patients rated both models as equally useful in all respects. Not the employed assessment model, but the relationship with the assessor was related to experienced utility and satisfaction. We hypothesized that the AMPD model might provide the clinician with

more tools to form a positive relationship with clients than the Section II PD model, however, no such differences emerged. One explanation is that we had experienced clinicians who had been working with PD patients for several years. One of the major challenges and goals in working with PD patients is forming and maintaining an alliance. It makes sense that all clinicians working in this specialized setting are skilled in building a strong alliance, independent of what 'language' or model they have to use. Our results are in line with the body of research on treatment alliance and outcome in psychotherapy (i.e. Norcross & Lambert, 2018). Although therapeutic alliance is a broader concept than we assessed in the current study, forming an emotional bond (e.g., a positive relationship) is one aspect of therapeutic alliance (Bordin, 1979). Therapeutic alliance has consistently been associated with treatment outcome across a range of psychotherapies (Baier et al., 2020; Flückiger et al., 2018; Martin et al., 2000). Our results suggest that alliance in the assessment phase - often ignored in assessment research - is also related to experiencing an assessment as useful and a higher satisfaction with the assessment. Holst and colleagues (also using the Assessment Questionnaire) (2009) found similar results in the context of neuropsychological assessments: a positive relationship was related to a higher level of satisfaction.

However, other reasons for a lack of superiority of the AMPD model from a patient's perspective may also be related to features of the AMPD itself. Aspects of the assessed client utility of an assessment related to, for example, the destigmatizing nature of the assessment procedures and outputs, or the increased insight into one's own core patterns. The AMPD still uses a typological approach, including potentially stigmatizing labels like borderline PD. New Criterion B concepts, including "antagonism" or "negative affectivity," are likely not less stigmatizing, while some Criterion A descriptors, like "experience of a unique self… organized around perceived external persecution," may be complex and in need for "translation" to be fully useful for patients. Therefore, from a patient's perspective, the AMPD may benefit from rewording and simplification.

We did not find the expected superiority of the AMPD model with respect to clinician-rated clinical utility. However, when taken into account that the (end user) clinicians were very familiar with the Section II PD model and less so with the AMPD model, these results are encouraging. Clinicians (again, not involved in the assessment) rated the models as equally useful for communicating with other professionals and all other aspects of clinical utility, even though the AMPD constitutes a whole set of new constructs with which professionals were not familiar yet. Familiarizing clinicians with the AMPD model will conceivably aid experienced utility of the model in the future. Furthermore, most clinicians conducting the AMPD assessment had to familiarize themselves with a new model, while the clinicians conducting the Section II assessment had been working with this model for years. That both groups provide equally useful assessments according to the end users, despite the differences in familiarity between the models, is at least reassuring. Of note, there was a slight difference in adherence between the models, the Section II PD assessors were less adherent to the model than the AMPD assessors, although still 70% of our random sample was rated as adherent. One explanation could be that the Section II PD assessors were previously used to incorporating different theories (schema-, psychodynamic, etc) into their case formulations, which may have made it harder to be adherent to the (a-theoretical) Section II PD model. Lastly, our results are in line with many treatment intervention studies in which two well-described and protocolized interventions generally perform equally well, independent of their theoretical background (Cristea et al., 2017).

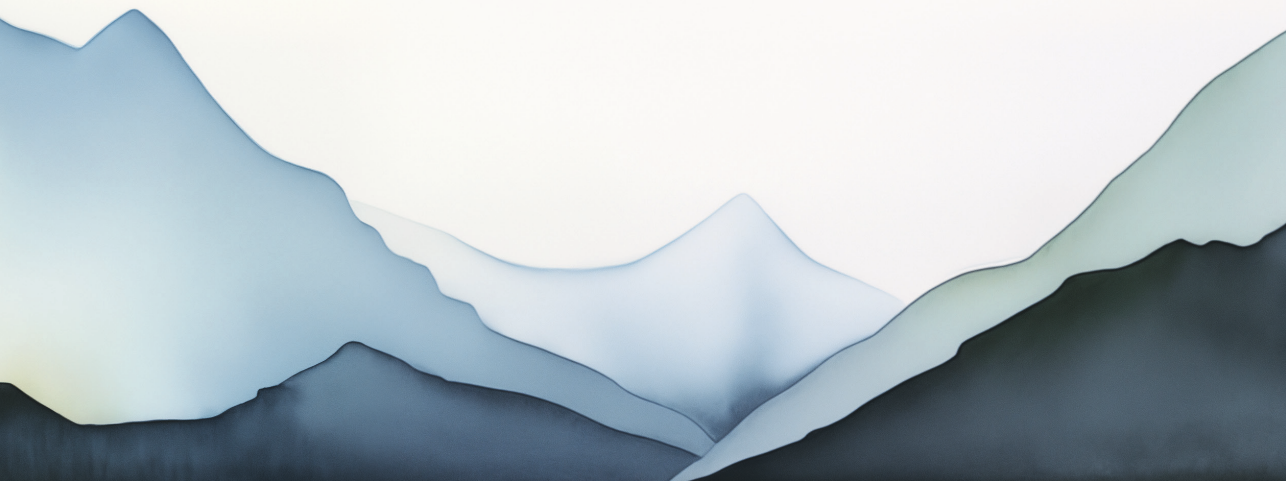
Our study has several strengths and limitations. First, a major strength is the RCT design which randomized patients over the assessment models and made direct comparisons possible. Second, the study was integrated into routine clinical practice, which aids ecological validity. Third, we conducted both assessment procedures according to detailed protocols and carefully trained the clinicians in their respective models. Most limitations of the current study pertain to the clinician variables. We included only four assessment clinicians

in the AMPD assessment procedure and seven in the Section II assessment procedure, which may well have impacted the respective results of both assessment models. Accordingly, we were not able to meaningfully assess (differences in) experienced utility as rated by the assessment clinicians working with the AMPD or Section II PD models. Moreover, assessment clinicians were not randomized over the assessment models. Further, most clinicians had years of experience working with the Section II PD model, while the experience of the clinicians in the AMPD assessment with the model ranged from none to only few years. A replication of our study with a larger number of assessment clinicians is warranted. Lastly, the current study may have been a too stringent test of clinical utility, as we compared the AMPD assessment to a Section II PD assessment conducted in a highly specialized centre with clinicians with years of experience in working with PD patients. Future studies may compare the utility of the AMPD model to 'assessment as usual' in a more general psychiatric unit. Lastly, a substantial percentage of data was missing, although to mitigate this concern, we handled missing data with multiple imputation, which is regarded as a state-of-the-art flexible technique in which the imputations stay close to the data. Lastly, this study used the AMPD assessment approach. Though highly similar to the AMPD, a replication of this study using the *International Classification of Diseases*, 11th revision framework would be timely and helpful for World Health Organization member states that are about to implement a new dimensional classification of PDs.

In conclusion, from a clinical utility perspective, the AMPD assessment was not superior to the Section II PD assessment, in terms of clinician- or patient-rated utility. There are, of course, many other reasons to choose one classification system over the other, such as general acceptability of the model in clinical practice and validity.

Acknowledgements

We would like to thank dr. Andrew E. Skodol for reviewing the manuals for the Section II PD and AMPD assessment.



Chapter 9

General discussion

The overarching objective of this dissertation was to investigate the validity and utility of operationalizations of the Alternative Model for Personality Disorders (AMPD) and to test and facilitate implementation of the assessment model in clinical practice. The following aims were formulated:

1. Develop specific instruments for assessing Criterion A and evaluate the impact of using the AMPD model in terms of case-identification of PD diagnoses;
2. Describe an AMPD assessment procedure which can be implemented in clinical practice and define what constitutes clinical utility of PD assessment;
3. Compare the AMPD model to the traditional Section II PD model in terms of predictive validity and clinical utility.

This chapter reviews and reflects on the main findings, and follows up with implications of these findings for policy making, theory development, and clinical practice, as well as recommendations for future research.

Review and discussion of the main findings

Instrument development and psychometric evaluation

The AMPD model radically changed the general criteria for the classification of personality disorders, by describing personality disorders as a combination of impairments in personality functioning (Criterion A) and maladaptive personality traits (Criterion B). Upon publication of the DSM-5, no instruments were available to reliably assess Criterion A. The DSM-5 provided the Level of Personality Functioning Scale, a description of 12 facets of personality functioning, each with five severity levels, resulting in a total of 60 descriptions. Initially, the work group claimed the level of personality functioning would be easy to assess in an unstructured clinical interview (Skodol et al., 2011). The LPFS was criticized, however, for being theory-laden and

abstract (Clarkin & Huprich, 2011; Pilkonis et al., 2011; Pincus, 2011; Tyrer, 2012). Critics were concerned that assessment with the LPFS would require extensive training, years of clinical experience with PD patients, and assessment over a longer period of time (Zimmermann et al., 2014). Indeed, several studies showed ratings based on unstructured clinical impressions or interviews not specifically designed to assess the LPFS had low to acceptable interrater reliability (Cruitt et al., 2019; Few et al., 2013; Preti et al., 2018; Roche et al., 2018; Morey, 2019; Young & Beazley, 2023; Zimmermann et al., 2014). Over the years, several instruments were developed to assess the LPFS (Bender et al., 2018; Goth et al., 2018; Huprich et al., 2018; Hutsebaut et al., 2016; Hutsebaut et al., 2017; Morey, 2017; Thylstrup et al., 2016). The present thesis investigated the psychometric properties of both a self-report questionnaire (LPFS-BF 2.0; Chapter 2) as well as an interview schedule (STiP 5.1; Chapter 3) for assessing the LPFS. Results were largely supportive for subsequent use of these instrument in clinical practice. Our results are in line with several other studies demonstrating that specifically designed instruments for assessing the LPFS generally improve validity and reliability (Young & Beazley, 2023). Taken together, the initial claim of the workgroup regarding the use of unstructured clinical interviews may have been ill advised, as the current evidence underscores the need to use structured interviews specifically designed to assess the LPFS for valid and reliable assessment (Young & Beazley, 2023).

Clinicians now have a considerable number of instruments and assessment methods to choose from when assessing the LPFS (i.e., interview, self-report, informant-report, brief vs. lengthy questionnaires), leading to the dilemma of instrument selection. Generally, substantial correlations have been found between interview schedules and self-report measures of personality functioning (Heissler et al., 2021; Nelson et al., 2018; Ohse et al., 2022; Somma et al., 2020; Roche & Jaweed, 2021). Although questionnaires are less time-consuming and more convenient than interview schedules, the extent to which (some) patients can self-assess their personality functioning is an important issue to consider. Some findings may indeed question the

ability of patients to self-assess their personality impairments. In a forensic setting, self-reported personality functioning and expert-rated personality functioning (interview) did not correlate at all (Hutsebaut et al., 2021). Furthermore, the Empathy element may be less accessible for self-report. Goth and colleagues (2018), for example, found that the Empathy scale of their Level of Personality Functioning Questionnaire for adolescents did not differentiate between a healthy- and a patient sample, whereas the other elements of the LPFS did. Moreover, the Empathy scale of the LPFS Self Report also showed lower (healthier) mean scores compared to the other elements in an adult sample (Morey, 2017). This observation raises the question whether at least some aspects of the LPFS are harder to self-report on than others, at least for certain subgroups of patients. In this light, clinician-rated interviews may provide a valuable alternative for specific elements of the LPFS or in specific populations.

With regard to questionnaire (e.g. brief/comprehensive) selection, a critique specific to the LPFS-BF 2.0 is that the questionnaire may not be able to capture the whole range of PD pathology (Paap et al., 2023). Paap and colleagues found weak associations between the LPFS-BF 2.0 and the number of Avoidant PD traits, which may complicate detecting these patients when using this instrument (alone). It may be that the brevity of the LPFS-BF 2.0 has made it specifically accurate for detecting the presence or absence of severe levels of personality pathology, as the items generally refer to more severe levels of the LPFS. A more comprehensive questionnaire, like the LPFS-SR (Morey, 2017), which includes questions pertaining to all severity levels of the LPFS, may be more able to detect a broader range of personality pathology. Moreover, these recent findings may highlight the importance of assessing both severity and traits to be able to detect the full range of personality pathology, as traits such as social avoidance, intimacy avoidance, and anxiousness for example are indicative of Avoidant PD.

Lastly, the factor structure of the LPFS remains subject to debate. Although the LPFS was developed and presented as a

unidimensional construct, a two-factor structure with two correlated factors as presented in the current thesis, corresponding to self- and interpersonal functioning domains, was corroborated by several research groups (Bach & Hutsebaut, 2018; Lakuta et al., 2022; Le Corff et al., 2022; Paap et al., 2023; Natoli et al., 2022; Rossi & Diaz-Batanero, 2023; Spitzer et al., 2021). Other studies found support for a single factor (Goth et al., 2018; Morey et al., 2017; Weekers et al., 2022; Zimmermann et al., 2020). Studies that did support a two-factor structure, all reported two highly correlated factors, which is in line with the notion of a single higher order severity dimension. For clinical practice however, the distinction of the two domains may have clinical utility in terms of formulating specific treatment recommendations.

Continuity and discontinuity between the Section II PD- and AMPD model

Chapter 4 investigated to what extent the AMPD assessment would yield PD prevalence similar to the Section II PD assessment. One might argue that 100% convergence is neither expected nor even desirable (as the latter raises the question why to change models), but large discrepancy would also be problematic, as we would lose connection with accumulated research and clinical findings. Continuity and discontinuity between the classification models was assessed by diagnosing patients using both models. Results were largely reassuring as acceptable to good convergence was found between the AMPD- and Section II PD models for specific PD types. Of note, although the AMPD model retained several PD types, it is conceivable that the PD field will eventually abandon types altogether, in line with the PD chapter in ICD-11 (World Health Organization, 2019).

An important difference between both models emerged with respect to PD prevalence in general. Assessment according to the AMPD model classified more patients as having a PD than the Section II PD model did. In adults, the AMPD model identified a substantially larger group of patients as having a PD (trait specified) than the Section II PD model did. This raises the question which model is "correct",

which remains essentially unanswerable in the absence of a golden standard. A next question then becomes which model is more useful for identifying patients in need of PD treatment. Is the AMPD model perhaps too sensitive for practical use? Or is the Section II PD model too stringent in its reliance on behavioral criteria and missing patients struggling with personality pathology? Of note, we observed the same pattern of findings in adolescents (Chapter 3). A higher number of individuals met the threshold for a PD when using the LPFS compared to the Section II PD model, reflected by adolescents generally meeting the threshold of moderate or more severe impairments in personality functioning required for a PD diagnosis, but a much lower number of adolescents meeting criteria for a full Section II PD diagnosis. By shifting towards a model that focusses on underlying psychic processes, maladaptive personality processes may be detected earlier, perhaps before (Section II PD) behavioral criteria are met. This would have several advantages as it might, for example, lead at-risk patients towards the right treatment program earlier and prevent a series of unsuccessful treatments. There is evidence that especially Borderline PD traits, which are consistently associated with more severe levels of personality functioning in the present thesis, are a reliable predictor of treatment non-response in adolescents (Ranøyen et al., 2018; Kivuruusu et al., 2020). The present thesis did not investigate whether these (new) PD patients would indeed benefit more from PD treatment (instead of a symptom-focused treatment), and this question poses an important topic for future research. When the purpose of case-identification was the prediction of future problems and disability, the AMPD model appeared superior (Chapter 7). In sum, the AMPD model appears to capture a vulnerability that is stable over time and relates to general malfunctioning. Early detection of these impairments may help direct at risk patients to PD-oriented treatment early, but the value of such efforts remains to be tested.

The benefits of early detection should be balanced by an appraisal of the issue of stigmatization. Lowering the threshold for PD classification may potentially stigmatize a larger group of patients, with

a risk of discriminatory attitudes from others, devaluation, and rejection (Catthoor et al., 2015; Corrigan et al., 2007). Psychiatric disorders, particularly PDs have consistently been associated with a high level of stigma in both adolescents and adults (e.g. Catthoor et al., 2015; Corrigan et al., 2007; Farrington 1977; Martin et al., 2007; Mukolo & Heflinger, 2011; Perry et al., 2007; Ring & Lawn, 2019; Sheehan et al., 2016). Although the AMPD model, at least its dimensional aspects, was presumed to lower stigma, we did not find support for this notion in our data. One might expect the AMPD model to have less stigma associated with it, as it represents a universal model of human personality, with every person falling somewhere on the continuum of healthy to extremely impaired personality functioning (Sharp & Wall, 2021). However, the AMPD model also employes several terms that could be considered stigmatizing, such as 'severe impairments' or traits such as 'hostility' or 'attention seeking'. Future research might address these issues, in order for the field to make an informed decision regarding early classification, more specifically the potential benefits of detecting PDs early in terms of treatment response, prognosis, and cost-effectiveness should be assessed.

Beyond psychometrics: PD assessment in clinical practice

Implementing a new model in clinical practice entails more than just administering different instruments. Translating a new model into a comprehensive assessment is challenging. No specific guidelines for how to use the model in clinical practice are available yet. In Chapter 5 we provided an example of how we translated the full model in clinical practice by presenting the case of Adam. We hold that the field can benefit most from the AMPD model when all components are embedded into a comprehensive case-formulation. Such a case-formulation ideally describes underlying impairments in personality functioning, hypotheses on why and how these impairments developed, and how they are related to maladaptive traits and presenting symptoms. In other words, trying to capture the patient in terms of her underlying vulnerabilities, tendencies in thinking, feeling,

and acting, and trying to understand why this particular individual developed these specific problems and how they are related to her current context. Of note, this is one way of embedding the AMPD model in clinical practice, not explicitly dictated as such by DSM-5. Others may choose a more succinct description of severity level and traits.

Employing the AMPD model to make a comprehensive case-formulation is consonant with the current trend towards de-medicalization of mental health care, in which advocates plead for less stigmatizing categorical classifications and more individualized assessment and treatment with a focus on the therapeutic relationship (Scheepers, 2021). Classifying a PD according to the AMPD model is still a categorical decision however. Loosing this categorical component, for example by only describing the patient's personality functioning and traits, may be an alternative and advocates of abandoning categories altogether may indeed feel this is enough for treatment planning. On the other hand, to treat or not to treat is also a categorical decision. It may be difficult to allocate patients to the right treatment when losing the PD category altogether. Research has shown for example that treatment of patients classified as having a PD is most effective when certain common factors are present (e.g., structure, focus on affect and alliance; Hutsebaut et al., 2021). Knowing whether a patient meets the PD threshold may thus be informative. At the same time, ideally our assessments should not (solely) focus on classification but pave the way for successful treatment by seeing and treating the patient as a person, beyond a diagnosis or set of symptoms, working collaboratively with the patient to develop first insights into core patterns and enhancing hope and motivation for treatment. As described in Chapter 6, these elements were underscored as most important in PD assessment by patients. Feeling respected and truly 'seen' is important for everyone but may be especially important for patients suffering from PDs as they tend to have negative experiences with others and are hypervigilant as a consequence (De Saeger et al., 2014; Siewerda et al., 2007; Pretzer, 1990).

How to optimize assessment with these aspects in mind? As our randomized controlled trial suggests, it was not the differential assessment models, but rather the quality of the relationship with the assessor that impacted the patient utility of assessment (Chapter 8). This finding fits well into the body of research on the importance of the therapeutic alliance in psychotherapy (Baier et al., 2020; Fluckiger et al., 2018; Martin et al., 2000; Norcross & Lambert, 2018). Therapeutic alliance has consistently been shown to relate to treatment outcome but has received rather scant attention in the assessment literature. The present thesis demonstrates that alliance in the assessment phase is also critically related to experiencing an assessment as useful, i.e., feeling less stigmatized, getting more insight into personal patterns, and for enhancing hope and motivation, all of which may pave the way for a positive treatment outcome.

Useful lessons may be drawn from Therapeutic Assessment (TA), a semi-structured method of collaborative assessment that aims for direct therapeutic impact (Finn, 2007; Kamphuis & Finn, 2018). Whereas regular assessment typically has been a top-down, unilateral endeavor with an 'expert' (the clinician) evaluating the patient, TA invites the patient to work collaboratively throughout all assessment phases. Its core values are grounded in humanistic psychology (Fisher, 1994; Fisher, 2000) and focus on respect, reducing the power imbalance between patient and assessor, and dialoguing with patients about test results instead of insisting that the assessor knows best. Several meta-analyses have shown this approach to be efficacious (Aschieri et al., 2023; Durosini & Aschieri, 2021; Poston & Hanson, 2010), and more specifically TA has demonstrated to foster therapeutic alliance in both the assessment phase and subsequent psychotherapy (De Saeger et al., 2014; Hilsenroth et al., 2004), as well as promote treatment readiness (De Saeger et al., 2014). Recent theorizing in TA puts less emphasis on specific techniques or tests, and more on its core values along with the therapeutic relationship and attitude of the assessor (Kamphuis & Finn, 2018). Specifically, the assessor demonstrates and models curiosity, and aims to be maximally transparent, non-judgmental, and

collaborative. These interpersonal elements may play a role in the restoration of epistemic trust: trust in the authenticity and personal relevance of information from others (Fonagy et al., 2015; Sperber et al., 2010). Epistemic trust has been proposed as a central mechanism of change in PD treatment, as relaxing epistemic hypervigilance restores the ability for social learning and accurate interpretations of other people's intentions. That said, TA is a fully individualized and typically high-dose form of psychological assessment, and it is unrealistic to provide all patients who come for PD assessment with a standard full TA (which may not involve PD assessment at all). We can, however, learn from the basic interpersonal principles of TA and adopt some of these principles in regular PD assessment. Helping assessors to take on a non-judgmental, collaborative attitude with a focus on building a genuine relationship with the patient and using accessible, non-stigmatizing language may foster utility of PD assessment.

Direct comparison of validity and utility of the two models of assessment

The introduction of the AMPD model represented a major shift from a phenomenological, descriptive model, describing behavioral manifestations of PD, towards a structural model, focused on underlying impairments and disposition. At the start of this thesis no direct comparisons between the models in clinical practice existed. Hence, there was a clear need for comparing the models on various validity and utility domains; a radical paradigm shift is only warranted when the new classification system improves validity and utility in clinical practice. To our knowledge, the present thesis presents the first head-to-head comparison of predictive validity between the assessment models in a true longitudinal design. Core finding was that the AMPD assessment, especially the LPFS, was related to long term general disability and symptom severity, while the Section II PD assessment was not. Accordingly, it may be theorized that the LPFS reflects the core PD vulnerability better than the Section II PD model, as impairments in personality functioning tend to be more

stable indicators of long term general (mal)functioning than the Section II PD behavioral criteria. For example, previous longitudinal studies demonstrated that Section II PD is less stable than initially thought, while the associated general disability remains stable over time (e.g. Gunderson et al., 2011; Zanarini et al., 2010; Zanarini et al., 2012). Indeed, a substantial number of patients with Borderline PD achieve symptom recovery but never achieve full recovery in terms of social- and occupational functioning (Gunderson et al., 2011; Zanarini et al., 2012). This poses a large symptom-disability gap, which begs the question whether remission in terms of Section II PD should be labelled as recovery, if patients are still disabled in most areas of their lives. If indeed impairments in personality functioning are more solid predictors of future functioning, this may have some implications for treatment planning. When the ultimate goal of treatment is not symptom remission but recovery of general wellbeing and functioning, treatment should focus on changing (and monitoring changes in) these underlying impairments: impairments in self- and interpersonal functioning. In sum, reconceptualizing PDs in terms of their underlying impairments may have a higher prognostic value, an important aspect of clinical utility. Clearly, as our study was the first to administer assessments according to both models, these findings are in need for replication and we recommend further trial with diverse samples.

Another way of comparing the AMPD- and Section II PD model is to address the notion that these models generate different information. Specifically, the AMPD assessment yields information regarding underlying vulnerabilities related to a variety of negative outcomes, while the Section II PD assessment informs about observable behavior. While the shift towards underlying impairments aids clinical utility in several ways, we may also lose some clinically relevant information. Diagnostic criteria tapping behaviors such as self-harm and suicidality, for example, are also quite informative for treatment planning. In sum, the AMPD model may provide us with information that is useful for long-term treatment planning, while information on dysfunctional (diagnostic) behaviors can be crucial for short-term treatment planning.

In terms of clinical utility as assessed by both patients and clinicians our randomized controlled trial demonstrated, contrary to expectation, that the Section II PD- and AMPD assessment did not differ (Chapter 8). The familiarity of both assessors and clinicians with the Section II PD model compared to the AMPD model complicates interpretation of this finding. When taking this (lack of) familiarity into account, the results could be interpreted as rather reassuring for the new model. This study also led us to an alternative hypothesis as described above: alliance may be most important for patients for experiencing a PD assessment as useful.

The hypothesized superiority of the AMPD in fostering alliance (and thus utility) in the assessment phase was not supported by our data. Perhaps the specific classification system used is less important for patients than initially thought, which would be in line with psychotherapy research showing that common factors, such as alliance, are more important for treatment outcome than the specific treatment method employed (Gelso et al., 2018; Wampold, 2015). Indeed, all clinicians in our study were experienced in working with PD patients and building alliance. Furthermore, most were also trained in Therapeutic Assessment and more than likely adopted a therapeutic stance that may have promoted alliance irrespective of the model they had to use.

Strengths and limitations

The current dissertation has several strengths that deserve mention. First, the presented studies constitute ecologically valid tests of the AMPD assessment model with true PD patients. They were implemented into regular practice and results should be generalizable to other setting specialized in the assessment and treatment of PD patients. Second, to test the predictive validity of the AMPD assessment, a true prospective study was conducted using a hetero-method design (with clinician rated interviews at start and self-report questionnaires at follow-up), which is rather rare in clinical practice. Third, we presented the first clinical test of differential

clinical and patient utility using a RCT design. Fourth, multi-method operationalizations of Criterion A and B were used in most studies. Lastly, a large part of this thesis focused on (defining and testing) the utility of assessment, which is in line with the current focus and research agenda in the larger field of clinical assessment (Kamphuis et al., 2021).

Several limitations also warrant mentioning. Given the setting of the comprising research projects, most patients had moderate or severe impairments in personality functioning. Patients with either healthier levels of personality functioning or more extreme impairments were underrepresented. Replication of our findings in a sample with more variability in severity level should be considered. Furthermore, the RCT comparing clinical and patient utility had several limitations. First, only few assessors administered the assessments, which may have impacted the results of both assessment models. Moreover, the difference in familiarity of the assessors with their respective models may have also impacted results. Finally, whereas our study was adequately powered to test patient-rated utility, differential clinician-rated utility could not be adequately tested as, not all patients were allocated to a treatment program within the institution after assessment, which led to a lower *N*.

Implications for theory development

Traditionally, research on treatment outcome has focused on symptom remission. Our findings (described in Chapter 7) suggest however that (Section II PD) symptoms and (long-term) social- and occupational functioning are not necessarily related. This may hold important implications. A shift towards a focus on recovery in terms of general functioning in treatment outcome studies instead of symptom remission seems warranted. The main goal of treatment is not the absence of symptoms, but rather the ability of patients to live a life worth living (Linehan, 2021). Indeed, many patients with PD can learn to create a stable environment for themselves, for example by limiting social activities and working a job that is less stressful

and not compatible with their level of cognitive functioning. Such an environment provides little stress and may contribute to the absence of PD symptoms. As a personal anecdote, a patient several years after symptom remission once told me: "I feel like I'm living my life with the brakes on." It seems there still is a lot to learn in terms of how to aid patients to function in society and create meaningful lives. The present thesis provides preliminary support for the notion that impairments in personality functioning may be an important underlying mechanism that hinders full recovery in terms of general functioning.

The discrepancy may also shed some new light on the controversial topic of the changeability of PDs. In the past, PDs were seen as 'untreatable chronic conditions' (Campbell et al., 2016; Coolidge & Segal, 1998). This view changed radically with the emergence of several evidence-based treatments for PD, leading to the optimistic stance that PDs are curable (Cristea et al., 2017; Mehlum et al., 2020; Stoffers et al., 2012). This position is, however, largely based on studies demonstrating (Section II PD) symptom remission. Longitudinal studies demonstrate that recovery in terms of social- and occupational functioning is much harder to attain, especially for patients with Borderline PD who, as the present thesis demonstrated, are patients with more severe impairments in personality functioning. Perhaps we are moving towards a more nuanced view, such that impairments in personality functioning may render patients vulnerable throughout their lives. Numerous studies have demonstrated that PD patients have generally endured negative childhood experiences and trauma (Battle et al., 2004; Jonhson, 1999; Laporte & Guttman, 1996). These early experiences may hinder the development of healthy self- and interpersonal functioning and leave the person with a lifelong vulnerability for psychopathology. Addressing these underlying impairments in personality functioning may well be essential. In fact, most evidence-based treatments for PDs are already implicitly or explicitly focused on improving self- and interpersonal functioning. Future research should investigate how these impairments evolve during and after treatment, as this may help us understand why some

patients seem to recover after treatment but come back several years later with a relapse in symptoms. If the underlying vulnerability (to some extent) remains, it stands to reason that PD patients return for treatment more than once in their lives, especially after life events or stressors, and denying patients treatment because they already underwent an evidence-based treatment in the past may not be appropriate.

Implications and recommendations for clinical practice

We will now articulate some of the major implications and recommendations for clinical practice. First, use of specifically designed instruments to assess Criterion A and B is recommended as it greatly enhances reliability of assessment. Interview schedules may be particularly warranted in specific samples (e.g., forensic samples) and when assessing specific elements that appear more difficult to self-assess (e.g., empathy). Questionnaires are especially useful for screening purposes and routine outcome monitoring. Since its publication, the LPFS-BF 2.0 has been translated into a number of languages and studied extensively in a variety of samples (e.g. András & Béla, 2023; Bach & Hutsebaut, 2018; Lakuta et al., 2022; Le Corff et al., 2022; Minarčíková et al., 2019; Natoli et al., 2022; Paap et al., 2023; Rossi & Diaz-Batanero, 2023; Spitzer et al., 2021; Stone et al., 2021; Zimmermann et al., 2020; Weekers et al., 2022). The international consortium for health outcomes measurement (ICHOM) has adopted the LPFS-BF 2.0 in their recommendation for a standard set of outcomes in personality disorder research, which will aid comparability of outcome studies. The LPFS-BF 2.0 may be especially useful for quick screening purposes and monitoring changes during and after treatment, while the LPFS-SR paints a more comprehensive picture of self-reported personality functioning during initial assessment. The STiP 5.1 is a valid and reliable tool for assessing Criterion A in a comprehensive PD assessment in both adults and adolescents. Currently there is no Dutch interview schedule for assessing maladaptive traits, which hinders the adoption of the full

AMPD model in clinical practice. The SCID-AMPD module II is available in English but has not been published in Dutch at the time of writing this thesis. For the time being, combining the STiP 5.1 with the PID-5 questionnaire (Krueger et al., 2012) seems like a best practice option.

Second, for classification and clinical decision making, investigating impairments in personality functioning may be most important. Although the AMPD defines PDs as a combination of (at least) moderate impairments in personality functioning (A) and at least one maladaptive personality trait (B), in practice it is highly unlikely that a patient meets Criterion A without meeting Criterion B. In other words, when moderate or more severe impairments in personality functioning are present, there usually (if not always) is at least one maladaptive personality trait present. Maladaptive traits, however, are informative as they provide the clinician with detailed information on the stylistic presentation of the PD; i.e., general ways of thinking, feeling, and acting of the patient which may consequently inform us on for example the relational style we can expect in treatment. Impairments in personality functioning, however, may be the most salient for clinical decision making. More severe impairments in personality functioning are associated with poorer outcomes and perhaps warrant more resource intensive treatments (Hopwood, 2018). The more severely impaired patients will struggle the most in terms of attaining social- and occupational recovery and general functioning.

Third, the implementation of the AMPD model in clinical practice remains an important subject to consider. It is a complex model with several (theoretical) concepts that are still unfamiliar to a large group of clinicians. What are the training requirements necessary to aid clinicians in using this model in clinical practice? A one- or two-day training will likely not suffice. A qualitative study by Heltne and colleagues (2022) demonstrated that, although clinicians favored the use of an LPFS interview over a Section II PD interview, they stressed the need of specific theoretical knowledge in order to administer the LPFS interview adequately. Experience with using the model and supervision in groups, headed by an experienced clinician with

respect to using the AMPD assessment model is likely necessary for clinicians to be able to use the model with confidence. Future DSM version might even consider simplifying several aspects of the model, for example replacing the rather complex formulations of the severity levels of the LPFS with more general descriptions of each severity level and general descriptions of what each facet of the LPFS entails.

Finally, the present thesis underscores the importance of a topic that is understudied in the clinical assessment literature: building a therapeutic alliance with PD patients in assessment. Guiding (novice) clinicians in how to build a genuine relationship with patients may be as important as teaching them the specific skills to reliably administer assessment instruments.

Implications for policy

The results of the present thesis provide preliminary support for the feasibility, as well as validity and utility of the AMPD model in clinical practice. Replication of these findings in similar and more heterogeneous samples is necessary, but pending these replications it seems justified to advice *policy makers* (e.g. managerial boards, guideline makers) to start implementing (aspects) of the model, more specifically the LPFS. Reliable instruments are now available to assess the LPFS and implementation of the model in practice is possible and seems acceptable to both patients and clinicians. We demonstrated adequate convergence between the AMPD- and Section II PD model, suggesting a not too radical change in prevalence rates of specific PD types. Furthermore, the AMPD model may have several advantages. First, the LPFS is related to (long-term) general disability and symptom severity, which can aid clinical decision making. Second, the LPFS seems especially useful for detecting personality pathology early (both in adolescents and adults) and identifying patients at risk, which can aid treatment allocation and planning. Taken together, there seems enough evidence to warrant implementation of the LPFS in clinical practice and educational programs.

More specifically, we would advise *policy makers* to implement the STiP 5.1 in personality disorder assessment both in adults and adolescents, particularly in specialized settings. Administering the STiP 5.1 could either be in addition to- or as a replacement of Section II PD measures. In the latter case, adding the PID-5 to the assessment battery is recommended. Moreover, we would advise *graduate training programs* in clinical psychology to train students on the AMPD model and use of the LPFS and *post-master educational programs* may want to include the AMPD model into their standard educational program. This may enhance awareness of personality pathology and impairments in personality functioning beyond specialized settings, as in general mental health care, where most PD patients will (initially) be treated. Psychotherapy for PD patients requires emphasis on specific clinical skills, such as a focus on the therapeutic alliance and ability to repair ruptures. More severe impairments in personality functioning may put a higher strain on the therapeutic alliance, which has consistently been shown to predict treatment outcome. The AMPD model might provide a useful framework for enhancing awareness of personality pathology in general mental health care, as it is a more universal model of personality that is applicable to all patients (and non-patients for that matter).

Lastly, in line with the recent ICHOM recommendations, we recommend *personality disorder researchers* to include the LPFS in outcome batteries as changes in personality functioning may be more informative for general functioning of patients than changes in (Section II PD) symptoms.

Recommendations for future research

Our body of work suggests several next steps for future research. First and foremost, direct comparisons between the Section II PD and AMPD model in clinical practice (as presented in the present thesis) are scarce and our results are therefore in need for replication in both similar and more heterogeneous samples. Second, the assessment of general impairment over time as a function of

personality functioning deserves more research attention. Assessing impairments in personality functioning over the course of several years in a longitudinal design would lend itself well to further test the relationship between general functioning/disability and personality functioning e.g., are changes in general functioning related to changes in personality functioning. Furthermore, this design should also include the short- and long-term impact of PD treatment on impairments in personality functioning, as well as how changes in personality functioning during and after treatment are related to changes in general functioning/disability. Finally, such research designs may also incorporate the impact of life events and personality functioning on general functioning, to investigate the hypothesis that life events will impact general functioning (more) when personality functioning is more severely impaired.

A third clinically relevant research question is how the AMPD model may inform treatment allocation. The model fits well with a shift towards a transdiagnostic approach, like the Hierarchical Taxonomy of Psychopathology model (HiTOP; Kotov et al., 2017), in which categorical classification is replaced by general (severity of) psychopathology and several trait clusters. Current treatment allocation, however, is still largely based on categorical diagnoses. The question *which treatment works best for whom* in terms of severity level and trait profile is thus an important one to address. A step towards answering this question may be to investigate the impact of (severity) of impairments in personality functioning on the treatment of other common mental disorders. Hence, instead of investigating the impact of comorbidity on treatment of common disorders such as depression or attention deficit hyperactivity disorder, it may be more useful to investigate the impact of impairments in personality functioning and the efficacy of adding common factors (such as a focus on alliance) that have shown to aid treatment outcome. Level of personality functioning may have a moderating effect on the relationship between treatment type and treatment outcome, where patients with healthy levels of personality functioning may benefit equally from treatment as usual and alliance-

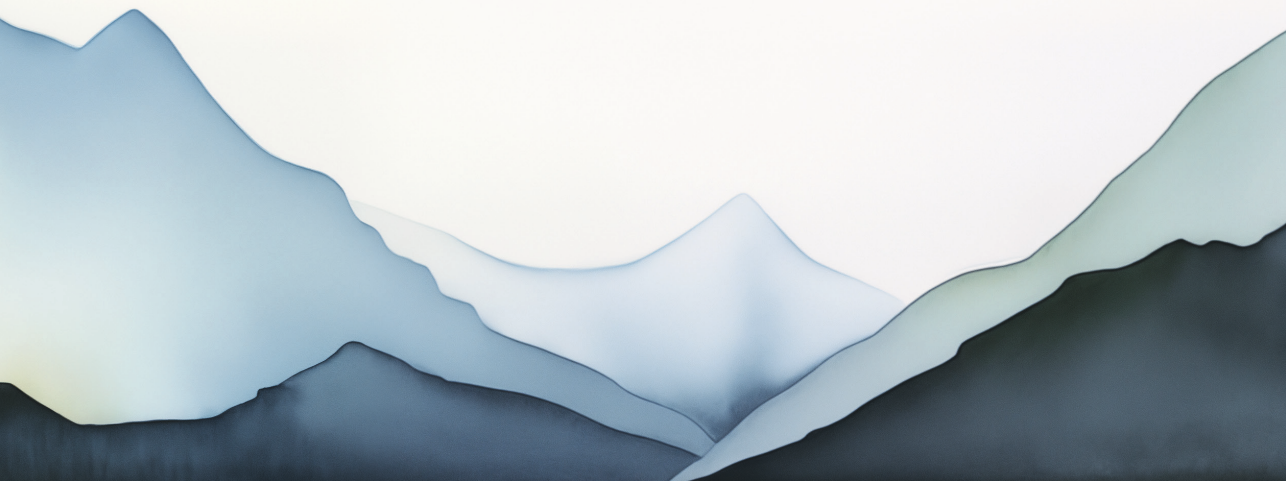
focused treatment, while more severely impaired patients might benefit more from the alliance-focused treatment.

Lastly, extending the line of speculative reasoning, with respect to early detection of impairments in personality functioning, we recommend investigating whether treatment allocation based on these impairments, i.e., treatment focused on self- and interpersonal functioning instead of symptoms such as anxiety and depression, leads to better patient outcomes. This could be tested in an adolescent sample in a randomized controlled trial (or more specifically, a manipulated assessment design), in which adolescents with moderate or more severe impairments in personality functioning, and a symptom disorder such as depression or anxiety, are randomly allocated to either a) a symptom focused treatment or b) a treatment focused on personality problems. Primary outcome in such a design could be general functioning/disability both short- and long term. Furthermore, cost-effectiveness of either treatment may be assessed, which is of course highly relevant for policy making.

Conclusions

The AMPD model constitutes a new and radically different PD classification system. It is a shift towards dimensionality, describing underlying impairments and traits, and fits into a transdiagnostic framework. The present thesis provided largely favorable psychometric results for both self-report and clinician-rated interview methods for assessing the AMPD. Compared to the Section II PD assessment, the AMPD assessment, in particular the LPFS is a promising tool for detecting PD early. Furthermore, compared to the Section II PD model, the LPFS may be superior in capturing vulnerabilities of patients that are predictive of general malfunctioning and may therefore be especially relevant for clinical decision making. Together, these findings add to the growing body of research demonstrating several important advantages of the AMPD model, as well as its applicability in clinical practice. With regard to utility of assessment, the patient perspective has largely been ignored. The present thesis provides a

first insight into what makes PD assessment useful for patients: feeling less stigmatized, providing insight, enhancing hope and motivation for treatment, collaboration, and transparency. Alliance seems a promising mechanism to foster these elements of utility and should be an important focus of all clinicians working with (PD) patients.



Addendum

Summary

Towards a New Perspective on Personality Disorder Classification:
The Alternative Model for Personality Disorders in Clinical Practice

DSM-5 introduced the Alternative Model for Personality Disorders (AMPD) to remedy many of the shortcomings of the existing categorical personality disorder (PD) model. The AMPD model presented a major paradigm shift and challenge to the field, as upon publication several gaps were notable in the research base that needed to be addressed before the model might be implemented in clinical practice. First, there were no reliable measures to assess Criterion A of the AMPD model. Furthermore, continuity between the AMPD- and existing Section II PD model in terms of PD prevalence had not been tested in clinical practice. The first aim of this thesis was to evaluate two measures for assessing Criterion A in an adult and adolescent sample and to evaluate the impact of the AMPD model in terms of case-identification (prevalence estimates) of PD diagnoses. Second, guidelines for implementation of the model in clinical practice were lacking. Moreover, for assessing clinical utility of the model in practice, sound clinical utility definitions and questionnaires were missing, specifically from a patient perspective. The second aim of this thesis was thus to describe an AMPD assessment procedure which could be implemented in clinical practice and to define what constitutes clinical utility of PD assessment from both a patient and a clinician perspective. Lastly, head-to-head comparisons were needed between the traditional Section II PD assessment and AMPD assessment in terms of validity and utility to justify the change towards a new model for PD classification. Hence, our third aim was to compare the predictive validity and clinical utility of the AMPD assessment to the traditional Section II PD assessment.

The first part of this thesis focused on the development and psychometric evaluation of two instruments for assessing the Level of Personality Functioning Scale (LPFS) and the impact of using the AMPD model on case-identification of PD. **Chapter 2** described the development and psychometric evaluation of the Level of Personality Functioning Scale Brief Form 2.0 (LPFS-BF 2.0), a short self-report questionnaire for assessing level of personality functioning. The LPFS-BF 2.0 was developed to give a global indication of personality functioning. Confirmatory factor analyses revealed a two-factor structure, representing Self- and Interpersonal functioning. Internal consistency of the questionnaire was acceptable and construct validity was supported by theoretically predicted associations with other measures of personality functioning and symptom severity. A significant difference between patients with and without a Borderline PD was found, with Borderline PD patients reporting more severe disturbances in personality functioning. Furthermore, the LPFS-BF 2.0 was sensitive to change, as evidenced by a high effect size after three months of inpatient PD treatment. Together, these results support the reliability and validity of the LPFS-BF 2.0 as a brief instrument to get a global impression of impairments in personality functioning and monitor changes in personality functioning over the course of treatment.

Chapter 3 described the psychometric evaluation of the Semi-structured interview for personality functioning DSM-5 (STiP 5.1), an interview for assessing the LPFS, in an adolescent sample. The STiP 5.1 could be reliably administered to adolescents, and demonstrated generally good interrater reliability, similar to reliability findings in an adult sample (Hutsebaut et al., 2017). Its construct validity was supported by theoretically predicted associations with other measures of personality functioning and personality traits, and the STiP 5.1 distinguished clinical- from non-clinical youngsters. Contrary to findings in adults, however, the STiP 5.1 did not differentiate between clinical youngsters with and without a Section II PD, with Borderline PD being the exception to this finding. The STiP 5.1 potentially identified

a higher number of adolescents as having a PD than the Section II PD model did, reflected by adolescents generally meeting the threshold of moderate or more severe impairments in personality functioning required for a PD diagnosis. We argue that the LPFS seems promising for detecting personality pathology early.

Chapter 4 investigated to what extent the AMPD model would yield PD prevalence similar to the extant Section II PD model. Continuity and discontinuity between the classification models was assessed by diagnosing (adult) patients using both models. Borderline and Avoidant PD showed stability of prevalence rates, suggesting continuity between the PD models. The AMPD assessment had a lower threshold for classifying PD than the Section II PD assessment resulting in more PD (trait specified) classifications when using the AMPD assessment compared to the Section II PD assessment.

The second part of this thesis described the implementation of the AMPD assessment in clinical practice and provided client- and clinician-definitions of utility of PD assessment. In **Chapter 5** a standardized multi-informant multi-method approach to assessing personality pathology using the AMPD model was presented and illustrated by a case-presentation. The use of several standardized instruments was discussed and their integration into a comprehensive case formulation was illustrated. Furthermore, in **Chapter 6,** we investigated the clinical utility construct by developing bottom-up definitions of client and clinician defined clinical utility of PD assessment. Clients and clinicians had considerable overlap in their definitions of what constitutes clinical utility of PD assessment. They both highlighted the importance of transparent communication, enhancing hope, curiosity and motivation and providing insight into patterns. Unique to clinicians' clinical utility definition was the importance of capturing both vulnerabilities and resilience of patients, and information on prognosis in treatment. Patients also highlighted the importance of an assessment to be collaborative and destigmatizing.

The third part of this thesis focused on direct comparisons between the AMPD assessment and the Section II PD assessment in ecologically

valid designs. A direct, head-to-head comparison of the ability of the AMPD vs the Section II PD model to predict disability and symptom severity one year after assessment was presented in **Chapter 7**. PD status and number of PD criteria according to the Section II PD assessment did not predict patient outcomes. In contrast, the AMPD assessment predicted disability and symptom severity one year later. More specifically, Criterion A was a strong predictor of functioning one year later, with patients with more severe impairments in personality functioning showing higher levels of disability and symptom severity.

Chapter 8 described a randomized controlled trial comparing the clinical utility of the Section II PD- and AMPD model as rated by patients and clinicians. Patients were randomized into either a) a Section II PD assessment or b) an AMPD assessment. Contrary to expectations, no differences between the models were observed for both patient- and clinician rated utility. Not the employed assessment model but experiencing a positive relationship with the assessor was predictive of patient-rated utility.

In **Chapter 9** the main findings, implications, and directions for future research were discussed. Taken together, the present thesis highlighted several advantages of using the AMPD model in clinical practice. Reliable instruments are available for assessment of the LPFS. Furthermore, the LPFS seems a promising tool for clinical decision making as personality pathology may be detected earlier both in adults and adolescents and level of personality functioning is related to future (general) functioning. Replications of these findings and more research is still needed, for example in more heterogeneous samples and using longitudinal designs, to further investigate the usefulness of the model in different populations and settings. Pending these results, it seems warranted to use the AMPD model, more specifically the LPFS, at least in settings for assessment and treatment of PDs, as a means of identifying personality pathology earlier and identifying adults at risk for future disability.

Nederlandse samenvatting

Richting een Nieuw Perspectief op de Classificatie van Persoonlijkheidsstoornissen:
Het Alternatieve Model voor Persoonlijkheidsstoornissen in de Klinische Praktijk

DSM-5 introduceerde het Alternatieve Model voor Persoonlijkheidsstoornissen (AMPD) om veel van de tekortkomingen van het bestaande categoriale model voor persoonlijkheidsstoornissen (PS) te verhelpen. Het AMPD model bracht een belangrijke paradigma verschuiving met zich mee en vormde een uitdaging voor het vakgebied, aangezien er nog verschillende hiaten in het onderzoeksveld waren bij publicatie van het model die moesten worden aangepakt voordat het model in de klinische praktijk kon worden geïmplementeerd. Ten eerste waren er geen betrouwbare meetinstrumenten om Criterium A van het AMPD model in kaart te brengen. Bovendien was de continuïteit tussen het AMPD model en bestaande Sectie II PS model in termen van de prevalentie van PS nog niet onderzocht in de klinische praktijk. Het eerste doel van dit proefschrift was om twee meetinstrumenten te evalueren voor het in kaart brengen van Criterium A in een volwassenen- en adolescentensteekproef, en om de impact van het AMPD model te evalueren op de identificatie van PS diagnoses (prevalentie-schattingen). Ten tweede ontbraken richtlijnen voor de implementatie van het model in de klinische praktijk. Bovendien ontbraken solide definities en vragenlijsten om klinische bruikbaarheid in kaart te brengen, met name vanuit het perspectief van de patiënt. Het tweede doel van dit proefschrift was daarom het beschrijven van een AMPD intakeprocedure die in de klinische praktijk kon worden geïmplementeerd, en om te definiëren wat klinische bruikbaarheid van een PS intake inhoudt, zowel vanuit het perspectief van de patiënt als vanuit het perspectief van de clinicus. Ten slotte waren vergelijkingen nodig tussen de traditionele Sectie-II PS intakeprocedure en de

AMPD intakeprocedure wat betreft validiteit en bruikbaarheid om de verandering naar een nieuw model voor PS classificatie te rechtvaardigen. Daarom was ons derde doel het vergelijken van de predictieve validiteit en klinische bruikbaarheid van de AMPD intakeprocedure en traditionele Sectie II PS intakeprocedure.

Het eerste deel van dit proefschrift richtte zich op de ontwikkeling en psychometrische evaluatie van twee instrumenten voor het beoordelen van de *Level of Personality Functioning Scale* (LPFS) en de impact van het gebruik van het AMPD model op de identificatie van PS diagnoses. **Hoofdstuk 2** beschreef de ontwikkeling en psychometrische evaluatie van de *Level of Personality Functioning Scale Brief Form 2.0* (LPFS-BF 2.0), een korte zelfrapportagevragenlijst voor het in kaart brengen van het niveau van persoonlijkheidsfunctioneren. De LPFS-BF 2.0 is ontwikkeld om een globale indicatie van persoonlijkheidsfunctioneren te geven. Uit de Confirmatory Factoranalyse kwam een twee-factorstructuur, die Zelf- en Interpersoonlijk functioneren vertegenwoordigden. De interne consistentie van de vragenlijst was acceptabel en de constructvaliditeit werd ondersteund door theoretisch voorspelde associaties met andere maten van persoonlijkheidsfunctioneren en symptoomernst. Er werd een significant verschil gevonden tussen patiënten met en zonder een Borderline PS, waarbij patiënten met een Borderline PS ernstigere beperkingen in het persoonlijkheidsfunctioneren hadden. Bovendien was de LPFS-BF 2.0 gevoelig voor verandering, zoals blijkt uit een grote effect size na drie maanden klinische behandeling voor PS. Samenvattend ondersteunen deze resultaten de betrouwbaarheid en validiteit van de LPFS-BF 2.0 als een kort instrument om een globale indruk te krijgen van beperkingen in het persoonlijkheidsfunctioneren en veranderingen in het persoonlijkheidsfunctioneren te monitoren gedurende de behandeling.

Hoofdstuk 3 beschreef de psychometrische evaluatie van het Semi-gestructureerde interview voor persoonlijkheidsfunctioneren DSM-5 (STiP 5.1), een interview voor het beoordelen van de LPFS, in een adolescentensteekproef. De STiP 5.1 kon betrouwbaar worden

afgenomen bij adolescenten en vertoonde over het algemeen goede interbeoordelaarsbetrouwbaarheid, vergelijkbaar met de betrouwbaarheidsbevindingen in een volwassen steekproef (Hutsebaut et al., 2017). De constructvaliditeit werd ondersteund door theoretisch voorspelde associaties met andere maten van persoonlijkheidsfunctioneren en persoonlijkheidstrekken, en de STiP 5.1 onderscheidde klinische jongeren van niet-klinische jongeren. In tegenstelling tot bevindingen bij volwassenen kon de STiP 5.1 echter geen onderscheid maken tussen klinische jongeren met en zonder een PS volgens Sectie II, waarbij Borderline PS de uitzondering vormde op deze bevinding. De STiP 5.1 identificeerde potentieel een groter aantal adolescenten met een PS dan het Sectie II PS model, adolescenten voldeden over het algemeen aan de drempelwaarde van matige of ernstigere beperkingen in het persoonlijkheidsfunctioneren die vereist is voor een PS diagnose. De LPFS lijkt veelbelovend voor het vroegtijdig detecteren van persoonlijkheidspathologie.

Hoofdstuk 4 onderzocht in hoeverre het AMPD model PS prevalenties zou opleveren die vergelijkbaar zijn met het bestaande Sectie II PS model. Continuïteit en discontinuïteit tussen de classificatiemodellen werd in kaart gebracht door (volwassen) patiënten te diagnosticeren met behulp van beide modellen. Borderline en Vermijdende PS vertoonden stabiliteit in prevalentiecijfers, wat wijst op continuïteit tussen de PS modellen. De AMPD intakeprocedure hanteerde een lagere drempel voor het classificeren van PS dan de Sectie II PS intakeprocedure, wat resulteerde in meer (trek-gespecificeerde) PS classificaties bij gebruik van de AMPD-intakeprocedure in vergelijking met de Sectie II PS intakeprocedure.

De tweede helft van dit proefschrift beschreef de implementatie van de AMPD-intakeprocedure in de klinische praktijk en gaf definities van klinische bruikbaarheid van een PS intake vanuit het perspectief van de cliënt en de clinicus. In **Hoofdstuk 5** werd een gestandaardiseerde multi-informant multi-methode benadering voor het beoordelen van persoonlijkheidspathologie met behulp van het AMPD model gepresenteerd en geïllustreerd aan de hand van een

gevalsbeschrijving. Het gebruik van verschillende gestandaardiseerde instrumenten werd besproken en hun integratie in een uitgebreide casusformulering werd geïllustreerd. Bovendien onderzochten we in **Hoofdstuk 6** het construct van klinische bruikbaarheid door bottom-up definities te ontwikkelen van klinische bruikbaarheid van een PS intake zoals gedefinieerd door cliënten en clinici. Cliënten en clinici hadden aanzienlijke overlap in hun definities van wat klinische bruikbaarheid van een PS intake inhoudt. Ze benadrukten allebei het belang van transparante communicatie, het versterken van hoop, nieuwsgierigheid en motivatie, en het bieden van inzicht in patronen. Uniek voor de definitie van klinische bruikbaarheid door clinici was het belang van het in kaart brengen van zowel kwetsbaarheden als veerkracht van patiënten, en informatie over de prognose in behandeling. Patiënten benadrukten ook het belang van een intake die collaboratief en destigmatiserend is.

Het derde deel van dit proefschrift richtte zich op directe vergelijkingen tussen de AMPD intakeprocedure en de Sectie II PS intakeprocedure in ecologisch valide onderzoeksdesigns. Een directe vergelijking van de capaciteit van het AMPD model versus het Sectie II PS model om algemeen functioneren en symptoomernst een jaar na intake te voorspellen, werd gepresenteerd in **Hoofdstuk 7**. De PS-status en het aantal PS-criteria volgens de Sectie II PS intakeprocedure voorspelden de patiëntresultaten niet. Daarentegen voorspelde de AMPD intakeprocedure algemeen functioneren en symptoomernst een jaar later. Meer specifiek was Criterium A een sterke voorspeller van functioneren een jaar later, waarbij patiënten met ernstigere beperkingen in het persoonlijkheidsfunctioneren hogere niveaus van algemeen disfunctioneren en symptoomernst vertoonden.

Hoofdstuk 8 beschreef een randomized controlled trial waarin de klinische bruikbaarheid van het Sectie II PS model en het AMPD model werd vergeleken, beoordeeld door patiënten en clinici. Patiënten werden willekeurig toegewezen aan a) een Sectie II PS intakeprocedure of b) een AMPD intakeprocedure. In tegenstelling tot de verwachtingen werden er geen verschillen tussen de modellen

gevonden voor zowel de bruikbaarheid zoals beoordeeld door de patiënt als door de clinicus. Niet het gebruikte intakemodel, maar het ervaren van een positieve relatie met de intaker voorspelde de door de patiënt beoordeelde bruikbaarheid van de intake.

In **Hoofdstuk 9** werden de belangrijkste bevindingen, implicaties en richtingen voor toekomstig onderzoek besproken. Samenvattend benadrukte dit proefschrift verschillende voordelen van het gebruik van het AMPD model in de klinische praktijk. Betrouwbare instrumenten zijn beschikbaar voor de beoordeling van de LPFS. Bovendien lijkt de LPFS een veelbelovend hulpmiddel te zijn voor klinische besluitvorming, aangezien persoonlijkheidspathologie zowel bij volwassenen als adolescenten eerder kan worden opgespoord en het niveau van persoonlijkheidsfunctioneren gerelateerd is aan toekomstig (algemeen) functioneren. Replicaties van deze bevindingen en meer onderzoek zijn nog steeds nodig, bijvoorbeeld in meer heterogene steekproeven en met longitudinale designs, om de bruikbaarheid van het model in verschillende populaties en contexten verder te onderzoeken. In afwachting van deze resultaten lijkt het gerechtvaardigd om het AMPD model, en specifiek de LPFS, in ieder geval in settings gericht op diagnostiek en behandeling van PS, te gebruiken om persoonlijkheidspathologie eerder te detecteren en volwassenen die risico lopen op toekomstig disfunctioneren te identificeren.

References

Allen, A., Montgomery, M., Tubman, J., Frazier, L., & Escovar, L. (2003). The effects of assessment feedback on rapport-building and self-enhancement processes. *Journal of Mental Health Counseling, 25*(3), 165-182. https://doi.org/10.17744/mehc.25.3.lw7h84q861dw6ytj

Allport, G. W. (1937). *Personality: a psychological interpretation*. Holt.

American Psychiatric Association, (2013). *Diagnostic and statistical manual of mental disorders: DSM-5(5th edition).* American Psychiatric Association.

Anderson, J. L., Sellbom, M. (2018). Evaluating the DSM-5 Section III personality disorder impairment criteria. *Personality Disorders: Theory, Research, and Treatment, 9*(1), 51–61. https://doi.org/10.1037/per0000217

Anderson, J. L., Snider, S., Sellbom, M., Krueger, R., & Hopwood, C. (2014). A comparison of the DSM-5 Section II and Section III personality disorder structures. *Psychiatry Research*, *216*(3), 363–372. https://doi.org/10.1016/j.psychres.2014.01.007

András, L., & Béla, B. (2023). A Személyiségműködés Színvonala – Rövid Változat 2.0 kérdőív magyar változatának (LPFS-BF 2.0 H) pszichometriai jellemzői egy Supetemista mintán. *Mentálhigiéné és Psichoszomatika, 24*(2), 100-112. doi: https://doi.org/10.1556/0406.2023.00031

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist, 73*(1), 3-25. https://doi.org/10.1037/amp0000191

Arntz, A. R., Kamphuis, J. H., & Derks, J. L. (2017). *SCID-5-P: Gestructureerd klinisch interview voor DSM-5 persoonlijkheidsstoornissen.* Boom.

Arntz, A., Kamphuis, J. H., & Derks, J. L. (2018). *SCID-5-S: Gestructureerd klinisch interview voor DSM-5 Syndroomstoornissen: Nederlandse vertaling van Structured Clinical Interview for DSM-5 Disorders-Clinician Version (SCID-5-CV), first edition, en User's Guide for the Structured Clinical Interview for DSM-5 Disorders-Clinician Version (SCID-5-CV), first edition* en delen van *Structured Clinical Interview for DSM-5 Disorders-Research Version (SCID-5-RV).* Boom.

Aschieri, F., van Emmerik, A. A., Wibbelink, C. J., & Kamphuis, J. H. (2023). A systematic research review of collaborative assessment methods. *Psychotherapy, 60*(3), 355-369. https://doi.org/10.1037/pst0000477

Bach, B., & Bernstein, D. P. (2019). Schema therapy conceptualization of personality functioning and traits in ICD-11 and DSM-5. *Current Opinion in Psychiatry, 32*(1), 38-49. https://doi.org/10.1097/yco.0000000000000464

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

Bach, B., & First, M. B. (2018). Application of the ICD-11 classification of personality disorders. *BMC psychiatry, 18*(1), 1-14. https://doi.org/10.1186/s12888-018-1908-3

Bach, B., Markon, K., Simonsen, E., & Krueger, R. F. (2015). Clinical utility of the DSM-5 Alternative Model of Personality Disorders: Six cases from practice. *Journal of Psychiatric Practice, 21*(1), 3–25. https://doi.org/10.1097/01.pra.0000460618.02805.ef

Bach, B., & Sellbom, M. (2016). Continuity between DSM-5 categorical criteria and traits criteria for borderline personality disorder. *Canadian Journal of Psychiatry, 61*(8), 489–494. https://doi.org/10.1177/0706743716640756

Bach, B., Sellbom, M., & Simonsen, E. (2018). Personality inventory for DSM-5 (PID-5) in clinical versus nonclinical individuals: Generalizability of psychometric features. *Assessment, 25*(7), 815-825. https://doi.org/10.1177/1073191117709070

Bach, B., & Tracy, M. (2022). Clinical utility of the alternative model of personality disorders: A 10th year anniversary review. *Personality Disorders: Theory, Research, and Treatment, 13*(4), 369–379. https://doi.org/10.1037/per0000527

Baier, A. L., Kline, A. C., & Feeny, N. C. (2020). Therapeutic alliance as a mediator of change: A systematic review and evaluation of research. *Clinical psychology review, 82*, 101921. https://doi.org/10.1016/j.cpr.2020.101921

Balsis, S., Lowmaster, S., Cooper, L. D., & Benge, J. F. (2011). Personality disorder diagnostic thresholds correspond to different levels of latent pathology. *Journal of Personality Disorders, 25*(1), 115-127. https://doi.org/10.1521/pedi.2011.25.1.115

Bastiaens, T., Smits, D., De Hert, M., Vanwalleghem, D., & Claes, L. (2016). DSM-5 section III personality traits and section II personality disorders in a Flemish community sample. *Psychiatry Research*, *238*, 290–298. https://doi.org/10.1016/j.psychres.2016.02.056

Bastiaansen, L., De Fruyt, F., Rossi, G., Schotte, C., & Hofmans, J. (2013). Personality disorder dysfunction versus traits: structural and conceptual issues. *Personality Disorders: Theory, Research, and Treatment, 4*(4), 293-303. https://doi.org/10.1037/per0000018

Bastiaansen, L., Hopwood, C. J., Van den Broeck, J., Rossi, G., Schotte, C., & De Fruyt, F. (2016). The twofold diagnosis of personality disorder: How do personality dysfunction and pathological traits increment each other at successive levels of the trait hierarchy?. *Personality Disorders: Theory, Research, and Treatment, 7*(3), 280- 292. https://doi.org/10.1037/per0000149

Battle, C. L., Shea, M. T., Johnson, D. M., Yen, S., Zlotnick, C., Zanarini, M. C., & Morey, L. C. (2004). Childhood maltreatment associated with adult personality disorders: findings from the Collaborative Longitudinal Personality Disorders Study. *Journal of Personality Disorders, 18*(2), 193-211. https://doi.org/10.1521/pedi.18.2.193.32777

Bender, D. S., Morey, L. C., Skodol, A. E. (2011). Toward a model for assessing level of personality functioning in DSM–5, Part I: A review of theory and methods. *Journal of Personality Assessment, 93*(4), 332–346. doi:10.1080/00223891.2011.583808.

Bender, D. S., & Skodol, A. E. (2007). Borderline personality as a self-other representational disturbance. *Journal of Personality Disorders, 21*(5), 500-517. https://doi.org/10.1521/pedi.2007.21.5.500

Bender, D. S., Skodol, A. E., First, M. B., & Oldham, J. M. (2018). Module I: Structured Clinical Interview for the Level of Personality Functioning Scale. In M. B. First, A. E. Skodol, D. S. Bender, & J. M. Oldham (Eds.), *Structured Clinical Interview for the DSM-5 Alternative Model for Personality Disorders (SCID-AMPD).* American Psychiatric Association Publishing.

Berghuis, H., Kamphuis, J. H., & Verheul, R. (2012). Core features of personality disorder: Differentiating general personality dysfunctioning from personality traits. *Journal of Personality Disorders, 26*(5), 704-716. https://doi.org/10.1521/pedi.2012.26.5.704

Berghuis, H., Kamphuis, J. H., Verheul, R., Larstone, R., & Livesley, J. (2013). The General Assessment of Personality Disorder (GAPD) as an instrument for assessing the core features of personality disorders. *Clinical Psychology & Psychotherapy, 20*(6), 544- 557. https://doi.org/10.1002/cpp.1811

Berne, E. (1996). Principles of transactional analysis. *Indian Journal of Psychiatry, 38*(3), 154-159.

Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, research & practice, 16*(3), 252-260. https://doi.org/10.1037/h0085885

Bornstein, R. F., & Natoli, A. P. (2019). Clinical utility of categorical and dimensional perspectives on personality pathology: A meta-analytic review. *Personality Disorders:Theory, Research, and Treatment*, *10*(6), 479–490. https://doi.org/10.1037/per0000365

Braun-Scharm, H. (1996). Suicidality and personality disorders in adolescence. *Crisis: Journal of Crisis Intervention & Suicide, 17*(2), 64-68. https://doi.org/10.1027/0227-5910.17.2.64

Brown, T. A. (2014). *Confirmatory factor analysis for applied research*. Guilford Press.

Buwalda, V. J. A., Nugter, M. A., Swinkels, J. A., & Mulder, C. L. (2011). *Praktijkboek ROM in de ggz. Een leidraad voor gebruik en implementatie van meetinstrumenten*. De Tijdstroom.

Cain, N. M., Pincus, A. L., & Ansell, E. B. (2008). Narcissism at the crossroads: Phenotypic description of pathological narcissism across clinical theory, social/personality psychology, and psychiatric diagnosis. *Clinical Psychology Review, 28*(4), 638-656. https://doi.org/10.1016/j.cpr.2007.09.006

Campbell, K., Clarke, K. A., Massey, D., & Lakeman, R. (2020). Borderline Personality Disorder: To diagnose or not to diagnose? That is the question. *International Journal of Mental Health Nursing, 29*(5), 972-981. https://doi.org/10.1111/inm.12737

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81–105. https://doi.org/10.1037/h0046016

Carcone, D., Tokarz, V., & Ruocco, A. (2015). A systematic review on the reliability and validity of semistructured diagnostic interviews for borderline personality disorder. *Canadian Psychology, 56*(2), 208-226. https://doi.org/10.1037/cap0000026

Carnovale, M., Carlson, E. N., Quilty, L. C., & Bagby, R. M. (2019). Discrepancies in self- and informant-reports of personality pathology: Examining the DSM–5 Section III trait model. *Personality Disorders: Theory, Research, and Treatment, 10*(5), 456-467. https://doi.org/10.1037/per0000342

Catthoor, K., Feenstra, D. J., Hutsebaut, J., Schrijvers, D., & Sabbe, B. (2015). Adolescents with personality disorders suffer from severe psychiatric stigma: evidence from a sample of 131 patients. *Adolescent Health, Medicine and Therapeutics, 6,* 81-89. https://doi.org/10.2147/ahmt.s76916

Chanen, A. M., & McCutcheon, L. K. (2008). Personality disorder in adolescence: The diagnosis that dare not speak its name. *Personality and Mental Health, 2*(1), 35-41. https://doi.org/10.1002/pmh.30

Chanen, A. M., & McCutcheon, L. K. (2013). Prevention and early intervention for borderline personality disorder: current status and recent evidence. *The British Journal of Psychiatry, 202*(s54), s24-s29. https://doi.org/10.1192/bjp.bp.112.119180

Chen, H., Cohen, P., Kasen, S., & Johnson, J. G. (2006). Adolescent axis I and personality disorders predict quality of life during young adulthood. *The Journal of adolescent health: official publication of the Society for Adolescent Medicine, 39*(1), 14–19. https://doi.org/10.1016/j.jadohealth.2005.07.005

Chen, H., Cohen, P., Kasen, S., Johnson, J. G., Berenson, K., & Gordon, K. (2006b). Impact of adolescent mental disorders and physical illnesses on quality of life 17 years later. *Archives of pediatrics & adolescent medicine, 160*(1), 93–99. https://doi.org/10.1001/archpedi.160.1.93

Chmielewski, M., Clark, L. A., Bagby, R. M., & Watson, D. (2015). Method matters: Understanding diagnostic reliability in DSM-IV and DSM-5. *Journal of Abnormal Psychology, 124*(3), 764-769. https://doi.org/10.1037/abn0000069

Chmielewski, M., Ruggero, C. J., Kotov, R., Liu, K., Krueger, R. F. (2017). Comparing the dependability and associations with functioning of the DSM-5 Section III trait model of personality pathology and the DSM-5 Section II personality disorder model. *Personality Disorder, 8*(3), 228–236. https://doi.org/10.1037/per0000213

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

Christensen, T. B, Eikenaes, I., Hummelen, B., Pedersen, G., Nysæter, T. E., Bender, D. S., Skodol, A. E., & Selvik, S. G. (2020). Level of personality functioning as a predictor of psychosocial functioning—Concurrent validity of criterion A. *Personality Disorders: Theory, Research, and Treatment, 11*(2), 79-90. https://doi.org/10.1037/per0000352

Clarkin, J. F., Cain, N., & Livesley, W. J. (2015). An integrated approach to treatment of patients with personality disorders. *Journal of Psychotherapy Integration, 25*(1), 3-12. https://doi.org/10.1037/a0038766

Clarkin, J. F., & Huprich, S. K. (2011). Do DSM-5 personality disorder proposals meet criteria for clinical utility?. *Journal of personality disorders, 25*(2), 192-205. https://doi.org/10.1521/pedi.2011.25.2.192

Coolidge, F. L., & Segal, D. L. (1998). Evolution of personality disorder diagnosis in the Diagnostic and Statistical Manual of Mental Disorders. *Clinical Psychology Review, 18*(5), 585-599. https://doi.org/10.1016/s0272-7358(98)00002-6

Corrigan, P. W., & Watson, A. C. (2007). How children stigmatize people with mental illness. *International Journal of Social Psychiatry, 53*(6), 526-546. https://doi.org/10.1177/0020764007078359

Creswell, K. G., Bachrach, R. L., Wright, A. G. C., Pinto, A., Ansell, E. (2016). Predicting problematic alcohol use with the DSM-5 alternative model of personality pathology. *Personality Disorders, 7*(1), 103–111. https://doi.org/10.1037/per0000131

Cristea, I. A., Gentili, C., Cotet, C. D., Palomba, D., Barbui, C., & Cuijpers, P. (2017). Efficacy of psychotherapies for borderline personality disorder: a systematic review and meta-analysis. *Jama psychiatry, 74*(4), 319-328. https://doi.org/10.1001/jamapsychiatry.2016.4287

Cruitt, P. J., Boudreaux, M. J., King, H. R., Oltmanns, J. R., & Oltmanns, T. F. (2019). Examining criterion a: Dsm-5 level of personality functioning as assessed through life story interviews. *Personality Disorders: Theory, Research, and Treatment*, *10*(3), 224–234. https://doi.org/10.1037/per0000321

Daley, S. E., Rizzo, C. J., & Gunderson, B. H. (2006). The longitudinal relation between personality disorder symptoms and depression in adolescence: The mediating role of interpersonal stress. *Journal of Personality Disorders, 20*(4), 352-368. https://doi.org/10.1521/pedi.2006.20.4.352

De Beurs, E. (2011). *Brief Symptom Inventory 18 (BSI): Handleiding*. PITS.

De Clercq, B., De Fruyt, F., De Bolle, M., Van Hiel, A., Markon, K. E., & Krueger, R. F. (2014). The hierarchical structure and construct validity of the PID-5 trait measure in adolescence. *Journal of Personality, 82*(2), 158-169. https://doi.org/10.1111/jopy.12042

Debast, I., Rossi, G., Feenstra, D. F., & Hutsebaut, J. (2017). Developmentally sensitive markers of personality functioning in adolescents: Age-specific and age-neutral expressions. *Personality Disorders: Theory, Research and Treatment, 8*(2), 162-171. https://doi.org/10.1037/per0000187

De Saeger, H., Kamphuis, J. H., Finn, S. E., Smith, J. D., Verheul, R., van Busschbach, J. J., & Horn, E. K. (2014). Therapeutic assessment promotes treatment readiness but does not affect symptom change in patients with personality disorders: Findings from a randomized clinical trial. *Psychological Assessment, 26*(2), 474-483. https://doi.org/10.1037/a0035667

De Wilde, E. F., & Hendriks, V. M. (2005). The Client Satisfaction Questionnaire: Psychometric Properties in a Dutch Addict Population. *European Addiction Research, 11*(4), 157-162. https://doi.org/10.1159/000086396

Dereboy, F., Dereboy, Ç., & Eskin, M. (2018). Validation of the DSM-5 Alternative Model Personality Disorder Diagnoses in Turkey, Part 1: Lead validity and reliability of the personality functioning ratings. *Journal of Personality Assessment*, *100*(6), 603–611. https://doi.org/10.1080/00223891.2018.1423989

Derogatis, L. R. (1975). *Brief symptom inventory.* Clinical Psychometric Research.

Delgadillo, J., Huey, D., Bennett, H., & McMillan, D. (2017). Case complexity as a guide for psychological treatment selection. *Journal of consulting and clinical psychology, 85*(9), 835–853. https://doi.org/10.1037/ccp0000231

Deyo, R. A., Diehr, P., & Patrick, D. L. (1991). Reproducibility and responsiveness of health status measures statistics and strategies for evaluation. *Controlled Clinical Trials*, *12*(4), S142-S158. https://doi.org/10.1016/s0197-2456(05)80019-4

Di Pierro, R., Gargiulo, I., Poggi, A., Madeddu, F., & Preti, E. (2020). The Level of Personality Functioning Scale Applied to Clinical Material From the Structured Interview of Personality Organization (STIPO): Utility in Detecting Personality Pathology. *Journal of Personality Disorders*, *34*(Supplement C), 62-76. https://doi.org/10.1521/pedi_2020_34_472

Durosini, I., & Aschieri, F. (2021). Therapeutic assessment efficacy: A meta-analysis. *Psychological Assessment, 33*(10), 962–972. https://doi.org/10.1037/pas0001038

Eaton, N. R., Krueger, R. F., South, S. C., Simms, L. J., & Clark, L. A. (2011). Contrasting prototypes and dimensions in the classification of personality pathology: Evidence that dimensions, but not prototypes, are robust. *Psychological Medicine, 41*(6), 1151-1163. https://doi.org/10.1017/s0033291710001650

Evans, C. M., & Simms, L. J. (2018). Assessing inter-model continuity between the section II and section III conceptualizations of borderline personality disorder in DSM-5. *Personality Disorders: Theory, Research, and Treatment, 9*(3), 290–296. https://doi.org/10.1037/per0000243

Farrington, D. P. (1977). The effects of public labelling. *British Journal of Criminology, 17*(2), 112-125. https://doi.org/10.1093/oxfordjournals.bjc.a046802

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods, 39*(2), 175-191. https://doi.org/10.3758/bf03193146

Fay, M. P. (2010). Two-sided exact tests and matching confidence intervals for discrete data motivating. *R Journal, 2*(1), 53–58.

Feenstra, D. J., Hutsebaut, J., Verheul, R., & Busschbach, J. J. (2011). Severity Indices of Personality Problems (SIPP–118) in adolescents: Reliability and validity. *Psychological Assessment*, *23*(3), 646-655. https://doi.org/10.1037/a0022995

Feenstra, D. J., Hutsebaut, J., Verheul, R., & van Limbeek, J. (2014). Identity: empirical contribution. Changes in identity integration of adolescents in treatment for personality disorders. *Journal of Personality Disorders, 28*(1)*, 101-112*. https://doi.org/10.1521/pedi.2014.28.1.101

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*(4), 1057-1069. https://doi.org/10.1037/a0034878

Finn, S. E. (2007). *In our clients' shoes: Theory and techniques of therapeutic assessment*. Routledge.

Finn, S. E., Schroeder, D. G., & Tonsager, M. E. (1994). The Assessment Questionnaire (AQ): A measure of clients' experiences with psychological assessment. *Unpublished manuscript*.

First, M. B. (2005). Clinical utility: A prerequisite for the adoption of a dimensional approach in DSM. *Journal of Abnormal Psychology, 114*(4), 560–564. https://doi.org/10.1037/0021-843X.114.4.560

First, M. B., Pincus, H. A., Levine, J. B., Williams, J. B. W., Ustun, B., & Peele, R. (2004). Clinical utility as a criterion for revising psychiatric diagnoses. *American Journal of Psychiatry, 161*(6), 946-954. https://doi.org/10.1176/appi.ajp.161.6.946

First, M. B., Skodol, A. E., Bender, D., & Oldham, J. M. (2018). *Structured Clinical Interview for DSM-5 Alternative Model for Personality Disorders (SCID-5-AMPD).* American Psychiatric Association.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1997). *Structured Clinical Interview for DSM-IV Axis I disorders.* American Psychiatric Press.

First, M. B., Gibbon, M., Spitzer, R. L., Benjamin, L. S., & Williams, J. B. (1997). *Structured clinical interview for DSM-IV axis II personality disorders (SCID-II).* American Psychiatric Press.

First, M. B., Williams, J. B., Benjamin, L. S., & Spitzer, R. L. (2016). *SCID-5-PD: Structured Clinical Interview for DSM-5 Personality Disorders*. American Psychiatric Association.

First, M. B., Williams, J. B., Karg R. S., & Spitzer, R. L. (2018). *SCID-5-CV: Structured Clinical Interview for DSM-5 Clinician Version.* American Psychiatric Association.

Fischer, C. T. (1994). *Individualizing psychological assessment.* Mahwah, NJ: Erlbaum.

Fischer, C. T. (2000). Collaborative, individualized assessment. *Journal of Personality Assessment*, 74(1), 2–14. doi:10.1207/S15327752JPA740102

Flückiger, C., Del Re, A. C., Wampold, B. E., & Horvath, A. O. (2018). The alliance in adult psychotherapy: A meta-analytic synthesis. *Psychotherapy, 55*(4), 316-340. https://doi.org/10.1037/pst0000172

Fok, M. L. Y., Hayes, R. D., Chang, C. K., Stewart, R., Callard, F. J., & Moran, P. (2012). Life expectancy at birth and all-cause mortality among people with personality disorder. *Journal of Psychosomatic Research, 73*(2), 104-107. https://doi.org/10.1016/j.jpsychores.2012.05.001

Fonagy, P., Luyten, P., & Allison, E. (2015). Epistemic petrification and the restoration of epistemic trust: A new conceptualization of borderline personality disorder and its psychosocial treatment. *Journal of Personality Disorders, 29*(5), 575-609. https://doi.org/10.1521/pedi.2015.29.5.575

Fossati, A., Krueger, R. F., Markon, K. E., Borroni, S., & Maffei, C. (2013). Reliability and validity of the Personality Inventory for DSM-5 (PID-5): Predicting DSM-IV personality disorders and psychopathy in community-dwelling Italian adults. *Assessment*, *20*(6), 689–708. https://doi.org/10.1177/1073191113504984

Fossati, A., Somma, A., Borroni, S., Maffei, C., Markon, K. E., & Krueger, R. F. (2016). A head-to-head comparison of the Personality Inventory for DSM-5 (PID-5) with the Personality Diagnostic Questionnaire-4 (PDQ-4) in predicting the general level of personality pathology among community dwelling subjects. *Journal of Personality Disorders, 30*(1), 82–94. https://doi.org/10.1521/pedi_2015_29_184

Fossati, A., Somma, A., Borroni, S., Pincus, A. L., Markon, K. E., & Krueger, R. F. (2017). Profiling pathological narcissism according to DSM–5 domains and traits: A study on consecutively admitted Italian psychotherapy patients. *Psychological Assessment, 29*(11), 1400-1411. http://dx.doi.org/10.1037/pas0000348

Fossati, A., Somma, A., Krueger, R. F., Markon, K. E., Borroni, S. (2017). On the relationships between DSM-5 dysfunctional personality traits and social cognition deficits: A study in a sample of consecutively admitted Italian psychotherapy patients. *Clinical Psychology & Psychotherapy, 24*(6), 1421–1434. https://doi.org/10.1002/cpp.2091

Fowler, J. C., Patriquin, M. A., Madan, A., Allen, J. G., Frueh, B. C., & Oldham, J. M. (2017). Incremental validity of the PID-5 in relation to the five factor model and traditional polythetic personality criteria of the DSM-5. *International Journal of Methods in Psychiatric Research, 26*(2), e1526. https://doi.org/10.1002/mpr.1526

Frances, A. (2012). Two who resigned from DSM-5 explain why. *Psychology Today*.

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

Gamache, D., Savard, C., Leclerc, P., Payant, M., Berthelot, N., Côté, A., & Tremblay, M. (2021). A Proposed Classification of ICD-11 Severity Degrees of Personality Pathology Using the Self and Interpersonal Functioning Scale. *Frontiers in Psychiatry*, *12*, 628057. https://doi.org/10.3389/fpsyt.2021.628057

Garcia, D. J., Skadberg, R. M., Schmidt, M., Bierma, S., Shorter, R. L., & Waugh, M. H. (2018). It's not that difficult: An interrater reliability study of the DSM–5 section III Alternative Model for Personality Disorders. *Journal of Personality Assessment, 100*(6), 612-620. https://doi.org/10.1080/00223891.2018.1428982

Gelso, C. J., Kivlighan Jr, D. M., & Markin, R. D. (2018). The real relationship and its role in psychotherapy outcome: A meta-analysis. *Psychotherapy, 55*(4), 434-444. https://doi.org/10.1037/pst0000183

Gibbon, S., Khalifa, N. R., Cheung, N. H-Y., Völlm, B. A., McCarthy, L. (2020). Psychological interventions for antisocial personality disorder. C*ochrane Database of Systematic Reviews*, *2020*(9). https://doi.org/10.1002/14651858.CD007668.pub3

Goth, K., Birkhölzer, M., & Schmeck, K. (2018). Assessment of Personality Functioning in Adolescents With the LoPF-Q 12-18 Self-Report Questionnaire. *Journal of Personality Assessment*, *100*(6), 680–690. https://doi.org/10.1080/00223891.2018.1489258

Grant, B. F., Hasin, D. S., Stinson, F. S., Dawson, D. A., Patricia Chou, S., June Ruan, W., & Huang, B. (2005). Co-occurrence of 12-month mood and anxiety disorders and personality disorders in the US: Results from the national epidemiologic survey on alcohol and related conditions. *Journal of Psychiatric Research, 39*(1), 1–9. https://doi.org/10.1016/j.jpsychires.2004.05.004

Grant, B. F., Stinson, F. S., Dawson, D. A., Chou, S. P., & Ruan, W. J. (2005). Co-occurrence of DSM-IV personality disorders in the United Stats: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Comprehensive Psychiatry, 46*(1), 1-5. https://doi.org/10.1016/j.comppsych.2004.07.019

Grilo, C. M., McGlashan, T. H., Quinland, D. M., Walker, M. L., Greenfeld, D., & Edell, W. S. (1998). Frequency of personality disorder in two age cohorts of psychiatric inpatients. *American Journal of Psychiatry, 155*(1), 140-142. https://doi.org/10.1176/ajp.155.1.140

Groenestijn, M. A. C., Akerhuis, G. W., Kupka, R. W., Schneider, N., & Nolen, W. A. (1999). *Gestructureerd klinisch interview voor de vaststelling van DSM-IV As-I stoornissen.* Swets test publishers.

Grove, W. M., & Tellegen, A. (1991). Problems in the classification of personality disorders. *Journal of Personality Disorders, 5*(1), 31-41. https://doi.org/10.1521/pedi.1991.5.1.31

Gunderson, J. G., Stout, R. L., McGlashan, T. H., Shea, M. T., Morey, L. C., Grilo, C. M., & Ansell, E. (2011). Ten-year course of borderline personality disorder: Psychopathology and function from the Collaborative Longitudinal Personality Disorders study. *Archives of general psychiatry, 68*(8), 827-837. https://doi.org/10.1001/archgenpsychiatry.2011.37

Guyatt, G., Walter, S., & Norman, G. (1987). Measuring change over time: assessing the usefulness of evaluative instruments. (2), *Journal of Chronic Diseases, 40*171-178. https://doi.org/10.1016/0021-9681(87)90069-5

Hansen, S. J., Christensen, S., Kongerslev, M. T., First, M. B., Widiger, T. A., Simonsen, E., & Bach, B. (2019). Mental health professionals' perceived clinical utility of the ICD- 10 vs. ICD-11 classification of personality disorders. *Personality and Mental Health, 13*(2), 84-95. https://doi.org/10.1002/pmh.1442

Hastrup, L. H., Jennum, P., Ibsen, R., Kjellberg, J., & Simonsen, E. (2019). Societal costs of Borderline Personality Disorders: A matched-controlled nationwide study of patients and spouses. *Acta Psychiatrica Scandinavica, 140*(5), 458-467. https://doi.org/10.1111/acps.13094

Heissler, R., Doubková, N., Hutsebaut, J., & Preiss, M. (2021). Semi-structured interview for personality functioning DSM-5 (STiP-5.1): Psychometric evaluation of the Czech version. *Personality and Mental Health*, *15*(3), 198–207. https://doi.org/10.1002/pmh.1508

Heltne, A., Bode, C., Hummelen, B., Falkum, E., Selvik, S. G., & Paap, M. C. (2022). Norwegian clinicians' experiences of learnability and usability of scid-ii, scid-5-pd and scid-5-ampd-i interviews: A sequential multi-group qualitative approach. *Journal of personality assessment, 104*(5), 599-612. https://doi.org/10.1080/00223891.2021.1975726

Heumann, K. A., & Morey, L. C. (1990). Reliability of categorical and dimensional judgements of personality disorder. *The American journal of psychiatry, 147*(4), 498- 500. https://doi.org/10.1176/ajp.147.4.498

Hilsenroth, M. J., Peters, E. J., & Ackerman, S. J. (2004). The development of therapeutic alliance during psychology assessment: Patient and therapist perspectives across treatment. *Journal of Personality Assessment, 83*(3), 332–344. doi:10.1207/s15327752jpa8303_14

Holst, Y., Nyman, H., & Larsson, J. O. (2009). Predictors of patient satisfaction with feedback after a neuropsychological assessment. *The Open Psychiatry Journal, 3*(1), 50 –55. https://doi.org/10.2174/1874354400903010050

Hopwood, C. J. (2018). A framework for treating DSM-5 alternative model for personality disorder features. *Personality and mental health, 12*(2), 107-125. https://doi.org/10.1002/pmh.1414

Hopwood, C. J., & Bornstein, R. F. (2014). *Multimethod clinical assessment*. Guilford Publications.

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

Hopwood, C. J., Malone, J. C., Ansell, E. B., Sanislow, C. A., Grilo, C. M., McGlashan, T. H., & Morey, L. C. (2011). Personality assessment in DSM-5: Empirical support for rating severity, style, and traits. *Journal of Personality Disorders, 25*(3), 305-320. https://doi.org/10.1521/pedi.2011.25.3.305

Huprich, S. K. (2018). Moving beyond categories and dimensions in personality pathology assessment and diagnosis. *The British Journal of Psychiatry, 213*(6), 685-689. https://doi.org/10.1192/bjp.2018.149

Huprich, S. K., Bornstein, R. F., & Schmitt, T. A. (2011). Self-report methodology is insufficient for improving the assessment and classification of Axis II personality disorders. *Journal of Personality Disorders, 25*(5), 557-570. https://doi.org/10.1521/pedi.2011.25.5.557

Huprich, S. K., Nelson, S. M., Meehan, K. B., Siefert, C. J., Haggerty, G., Sexton, J., & Baade, L. (2018). Introduction of the DSM-5 levels of Personality Functioning Questionnaire. *Personality Disorders: Theory, Research, and Treatment*, *9*(6), 553–563. https://doi.org/10.1037/per0000264

Hutsebaut, J., Bachrach, N., Kindt, K. C. M., & van Dam, L. J. H. (2021). Hoe bewezen effectief is de guideline-informed treatment for personality disorders (GIT-PD)?. *Tijdschrift voor Psychiatrie, 64*(1), 000-000.

Hutsebaut, J., Berghuis, H., De Saeger, H., Kaasenbrood, A., & Ingenhoven, T. (2014). *Semi-structured interview for personality functioning DSM-5 (STiP 5).* The Podium DSM-5 research Group of the Netherlands Centre of Expertise on Personality Disorders. Trimbos Institute.

Hutsebaut, J., Feenstra, D. J., & Kamphuis, J. H. (2016). Development and preliminary psychometric evaluation of a brief self-report questionnaire for the assessment of the DSM–5 level of Personality Functioning Scale: The LPFS brief form (LPFS-BF). *Personality Disorders: Theory, Research, and Treatment, 7*(2), 192-197. https://doi.org/10.1037/per0000159

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

Hutsebaut, J., Videler, A. C., Verheul, R., & van Alphen, S. P. J. (2019). Managing Borderline Personality Disorder from a life course perspective: clinical staging and health management. *Personality Disorders: Theory, Research, and Treatment, 10*(4), 309. https://doi.org/10.1037/per0000341

Hutsebaut, J., Videler, A. C., Willemsen, E., & Aalders, H. (2023). Guideline-informed treatment for personality disorders (GIT-PD): Tegenover, naast, op, onder of in plaats van CGT? *Tijdschrift voor Gedragstherapie, 56*(2).

Hutsebaut, J., Weekers, L. C., Tuin, N., Apeldoorn, J. S., & Bulten, E. (2021). Assessment of ICD-11 personality disorder severity in forensic patients using the Semi-structured Interview for Personality Functioning DSM-5 (STiP-5.1): Preliminary findings. *Frontiers in Psychiatry, 12,* 617702. https://doi.org/10.3389/fpsyt.2021.617702

Johnson, J. G., Cohen, P., Brown, J., Smailes, E. M., & Bernstein, D. P. (1999). Childhood maltreatment increases risk for personality disorders during early adulthood. *Archives of General Psychiatry, 56*, 600-606. https://doi.org/10.1001/archpsyc.56.7.600

Johnson, J. G., First, M. B., Cohen, P., Skodol, A. E., Kasen, S., & Brooks, J.S. (2005). Adverse outcomes associated with personality disorder not otherwise specified in a community sample. *American Journal of Psychiatry, 162*(10), 1926-1932. https://doi.org/10.1176/appi.ajp.162.10.1926

Johnson, J. G., Cohen, P., Kasen, S., Skodol, A. E., Hamagami, F., & Brooks, J. S. (2000). Age-related change in personality disorder trait levels between early adolescence and adulthood: a community-based longitudinal investigation. *Acta Psychiatrica Scandinavica, 102*(4), 265-275. https://doi.org/10.1034/j.1600-0447.2000.102004265.x

Johnson, J. G., Cohen, P., Skodol, A. E., Oldham, J. M., Kasen, S., & Brook, J. S. (1999). Personality disorders in adolescence and risk of major mental disorders and suicidality during adulthood. *Archives of General Psychiatry, 56*(9), 805–811. https://doi.org/10.1001/archpsyc.56.9.805

Kamphuis, J. H., & Finn, S. E. (2018). Therapeutic assessment in personality disorders: Toward the restoration of epistemic trust. *Journal of Personality Assessment*, *101*(6), 662-674. https://doi.org/10.1080/00223891.2018.1476360

Kamphuis, J. H., Noordhof, A., & Hopwood, C. J. (2021). When and how assessment matters: An update on the Treatment Utility of Clinical Assessment (TUCA). *Psychological Assessment, 33*(2), 122-132. https://doi.org/10.1037/pas0000966

Kasen, S., Cohen, P., Skodol, A. E., First, M. B., Johnson, J. G., Brooks, J. S., & Oldham, J. M. (2007). Comorbid personality disorder and treatment use in a community sample of youths: A 20-year follow-up. *Acta Psychiatrica Scandinavica, 115*(1), 56-65. https://doi.org/10.1111/j.1600-0447.2006.00842.x

Kazis, L. E., Anderson, J. J., & Meenan, R. F. (1989). Effect sizes for interpreting changes in health status. *Medical care*, *27*(Supplement), S178-S189. https://doi.org/10.1097/00005650-198903001-00015

Keefe, J. R., McMain, S. F., McCarthy, K. S., Zilcha-Mano, S., Dinger, U., Sahin, Z., & Barber, J. P. (2020). A meta-analysis of psychodynamic treatments for borderline and cluster C personality disorders. *Personality Disorders: Theory, Research, and Treatment*, *11*(3), 157–169. https://doi.org/10.1037/per0000382

Keeley, J. W., Chmielewski, M. S., & Bagby, R. M. (2015). Interaction effects in comorbid psychopathology. *Comprehensive Psychiatry, 60*, 35-39. https://doi.org/10.1016/j.comppsych.2015.04.008

Keeley, J. W., DeLao, C. S., & Kirk, C. L. (2013). The commutative property in comorbid diagnosis: Does A + B = B + A? *Clinical Psychological Science, 1*(1), 16-29. https://doi.org/10.1177/2167702612455742

Keeley, J. W., Reed, G. M., Roberts, M. C., Evans, S. C., Medina-Mora, M. E., Robles, R., & Saxena, S. (2016). Developing a science of clinical utility in diagnostic classification systems field study strategies for ICD-11 mental and behavioral disorders. *American Psychologist*, *71*(1), 3–16. https://doi.org/10.1037/a0039972

Kim, N. S., & Ahn, W. (2002). Clinical psychologists' theory-based representations of mental disorders predict their diagnostic reasoning and memory. *Journal of Experimental Psychology: General, 131*(4), 451-476. https://doi.org/10.1037/0096-3445.131.4.451

Kiviruusu, O., Strandholm, T., Karlsson, L., & Marttunen, M. (2020). Outcome of depressive mood disorder among adolescent outpatients in an eight-year followup. *Journal of Affective Disorders, 266,* 520–527. https://doi.org/10.1016/j.jad.2020.01.174

Koster, N., Laceulle, O. M., van der Heijden, P. T., Klimstra, T., de Clercq, B., Verbeke, L., & van Aken, M. A. G. (2019). A psychometric evaluation of a reduced version of the PID-5 in clinical and non-clinical adolescents. *European Journal of Psychological Assessment, 36*(5), 758-766. https://doi.org/10.1027/1015-5759/a000552

Kotelnikova, Y., & Clark, L. A. (n.d.). *Clinical Utility Rating Form*.

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, M., Zimmerman, M. (2017). The hierarchical taxonomy of psychopathology: A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology, 126*(4), 454–477. doi:10.1037/ abn0000258

Kraemer, H. C. (2014). The reliability of clinical diagnoses: State of the art. *Annual Reviews, of Clinical Psychology, 10*(1), 111-130. https://doi.org/10.1146/annurev-clinpsy-032813-153739

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*(9), 1879–1890. doi:10.1017/S0033291711002674.

Krueger, R. F., Hopwood, C. J., Wright, A. G. C., Markon, K. E. (2014). Challenges and strategies in helping the DSM become more dimensional and empirically based. *Current Psychiatry Reports, 16*(12), 1-6. https://doi.org/10.1007/s11920-014-0515-3

Łakuta, P., Cieciuch, J., Strus, W., & Hutsebaut, J. (2022). Level of Personality Functioning Scale-Brief Form 2.0: Validity and reliability of the Polish adaptation. *Psychiatria Polska, 57*(2), 247-260. https://doi.org/10.12740/pp/onlinefirst/145912

Lambert, M. (2007). Presidential address: What we have learned from a decade of research aimed at improving psychotherapy outcome in routine care. *Psychotherapy research, 17*(1), 1-14. https://doi.org/10.1080/10503300601032506

Lambert, M. J., Harmon, C., Slade, K., Whipple, J. L., & Hawkins, E. J. (2005). Providing feedback to psychotherapists on their patients' progress: Clinical results and practice suggestions. *Journal of Clinical Psychology, 61*(2), 165-174. https://doi.org/10.1002/jclp.20113

Landelijke Stuurgroep Multidisciplinaire Richtlijnontwikkeling in de GGZ. (2008). *Multidisciplinaire richtlijn persoonlijkheidsstoornissen [Multidisciplinary guideline for personality disorders]*.

Laporte L., & Guttman, H. (1996). Traumatic childhood experiences as risk factors for borderline and other personality disorders. *Journal of Personality Disorders, 10*(3), 247-259. https://doi.org/10.1521/pedi.1996.10.3.247

Larsen, D. L., Attkisson, C. C., Hargreaves, W. A., & Nguyen, T. D. (1979). Assessment of client/patient satisfaction: Development of a general scale. *Evaluation and Program Planning, 2*(3), 197-207. https://doi.org/10.1016/0149-7189(79)90094-6

Lavan, H., & Johnson, J. G. (2002). The association between axis I and II psychiatric symptoms and high-risk sexual behavior during adolescence. *Journal of Personality Disorders, 16*(1), 73-94. https://doi.org/10.1521/pedi.16.1.73.22559

Le Corff, Y., Aluja, A., Rossi, G., Lapalme, M., Forget, K., García, L. F., & Rolland, J. P. (2022). Construct Validity of the Dutch, English, French, and Spanish LPFS-BF 2.0: Measurement Invariance Across Language and Gender and Criterion Validity. *Journal of Personality Disorders, 36*(6), 662-679. https://doi.org/10.1521/pedi.2022.36.6.662

Lengel, G. J., & Mullins-Sweatt, S. N. (2017). The importance and acceptability of general and maladaptive personality trait computerized assessment feedback. *Psychological Assessment, 29*(1), 1-12. https://doi.org/10.1037/pas0000321

Lenzenweger, M. F., Johnson, M. D., & Willett, J. B. (2004). Individual growth curve analysis illuminates stability and change in personality disorder features: the longitundinal study of personality. *Archives of General Psychiatry, 61*(10), 1015-1024. https://doi.org/10.1001/archpsyc.61.10.1015

Levy, K. N. (2012). Subtypes, dimensions, levels, and mental states in narcissism and narcissistic personality disorder. *Journal of Clinical Psychology*, *68*(8), 886–897. https://doi.org/10.1002/jclp.21893

Liang, M. H., Fossel, A. H., & Larson, M. G. (1990). Comparisons of five health status instruments for orthopedic evaluation. *Medical care, 28*(7), 632-642. https://doi.org/10.1097/00005650-199007000-00008

Liggett, J., & Sellbom, M. (2018). Examining the DSM-5 alternative model of personality disorders operationalization of obsessive-compulsive personality disorder in a mental health sample. *Personality Disorders: Theory, Research, and Treatment, 9*(5), 397– 407. https://doi.org/10.1037/per0000285

Liggett, J., Sellbom, M., & Carmichael, K. L. C. (2017). Examining the DSM-5 section III criteria for obsessive-compulsive personality disorder in a community sample. *Journal of Personality Disorders*, *31*(6), 790–809. https://doi.org/10.1521/pedi_2017_31_281

Lilienfeld, S. O., Watts, A. L., Murphy, B., Costello, T. H., Bowes, S. M., Smith, S. F., & Tabb, K. (2019). Personality disorders as emergent interpersonal syndromes: Psychopathic personality as a case example. *Journal of Personality Disorders, 33*(5), 577–622. https://doi.org/10.1521/pedi.2019.33.5.577

Linehan, M. (2021). *Building a life worth living: A memoir.* Random House Trade Paperbacks.

Little, T. D. (2013). *Longitudinal structural equation modeling*. Guilford Press.

Livesley, W. J. (1998). Suggestions for a framework for an empirically based classification of personality disorder. *Canadian Journal of Psychiatry/Revue Canadienne De Psychiatrie*, *43*(2), 137–147. https://doi.org/10.1177/070674379804300202

Livesley, W. J. (2012). Disorder in the proposed DSM–5 classification of personality disorders. *Clinical Psychology and Psychotherapy, 19*(5), 364–368. https://doi.org/10.1002/cpp.1808

Lobbestael, J., Leurgans, M., & Arntz, A. (2011). Inter-rater reliability of the structured clinical interview for DSM-IV Axis I disorders and Axis II disorders. *Clinical Psychology and Psychotherapy, 18*(1), 75-79. https://doi.org/10.1002/cpp.693

Lowe, J. R., & Widiger, T. A. (2009). Clinicians' judgments of clinical utility: A comparison of the DSM-IV with dimensional models of general personality. *Journal of Personality Disorders*, *23*(3), 211–229. https://doi.org/10.1521/pedi.2009.23.3.211

Maffei, C., Fossati, A., Agostoni, I., Barraco, A., Bagnato, M., Deborah, D., & Petrachi, M. (1997). Interrater reliability and internal consistency of the structured clinical interview for DSM-IV Axis II personality disorders (SCID-II), Version 2.0. *Journal of Personality Disorders, 11*(3), 279-284. https://doi.org/10.1521/pedi.1997.11.3.279

Markon, K. E., Chmielewski, M., & Miller, C. J. (2011). The reliability and validity of discrete and continuous measures of psychopathology: A quantitative review. *Psychological bulletin, 137*(5), 856-879. https://doi.apa.org/doi/10.1037/a0023678

Martin, D. J., Garske, J. P., & Davis, M. K. (2000). Relation of the therapeutic alliance with outcome and other variables: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 68*(3), 438-450. https://doi.org/10.1037/0022-006x.68.3.438

Martin, J. K., Pescosolido, B. A., Olafsdottir, S., & McLeod, J. D. (2007). The construction of fear: Americans' preferences for social distance from children and adolescents with mental health problems. *Journal of Health and Social Behavior, 48*(1), 50-67. https://doi.org/10.1177/002214650704800104

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1*(1), 30–46. https://doi.org/10.1037/1082-989x.1.1.30

Mehlum, L., Schmahl, C., Berens, A., Doering, S., Hutsebaut, J., Kaera, A., & di Giacomo, E. (2020). Euthanasia and assisted suicide in patients with personality disorders: A review of current practice and challenges. *Borderline Personality Disorder and Emotion Dysregulation, 7*(1), 1-7. https://doi.org/10.1186/s40479-020-00131-9

Milinkovic, M. S., & Tiliopoulos, N. (2020). A systematic review of the clinical utility of the DSM-5 Section III Alternative Model of Personality Disorder. *Personality Disorders: Theory, Research, and Treatment*, *11*(6), 377–397. https://doi.org/10.1037/per0000408

Minarčíková, N., Barcaj, M., & Preiss, M. (2019). Úroveň funkčních schopností osobnosti- pilotní studie české verze krátké dotazníkové metody LPFS-BF [Level of Personality Functioning-Czech Pilot Study of a Brief Questionnaire LPFS-BF]. *Aplikovaná Psychologie, 6*, 356-367.

Morey, L. C. (2017). Development and initial evaluation of a self-report form of the DSM-5 Level of Personality Functioning Scale. *Psychological Assessment*, *29*(10), 1302– 1308. https://doi.org/10.1037/pas0000450

Morey, L. C. (2019). Interdiagnostician reliability of the DSM-5 Section II and Section III alternative model criteria for borderline personality disorder. *Journal of Personality Disorders, 33*(6), 721-735. https://doi.org/10.1521/pedi_2019_33_362

Morey, L. C., Bender, D. S., & Skodol, A. E. (2013). Validating the proposed Diagnostic and Statistical Manual of Mental Disorders, 5th Edition , severity indicator for personality disorder. *The Journal of Nervous and Mental Disease, 201*(9), 729–735. https://doi.org/10.1097/NMD.0b013e3182a20ea8

Morey, L. C., & Benson, K. T. (2016). Relating DSM-5 section II and section III personality disorder diagnostic classification systems to treatment planning. *Comprehensive Psychiatry. 68*, 48– 55. https://doi.org/10.1016/j.comppsych.2016.03.010

Morey, L. C., Benson, K. T., Busch, A. J. & Skodol, A. E. (2015). Personality Disorders in DSM-5: Emerging research on the Alternative Model. *Current Psychiatry reports, 17*(4), 558-566. doi:10.1007/s11920-015-0558-0

Morey, L. C., Berghuis, H., Bender, D. S., Verheul, R., Krueger, R. F., & Skodol, A. E. (2011). Toward a model for assessing level of personality functioning in DSM–5, part II: Empirical articulation of a core dimension of personality pathology. *Journal of Personality Assessment, 93*(4), 347-353. https://doi.org/10.1080/00223891.2011.577853

Morey, L. C., & Skodol, A. E. (2013). Convergence between DSM-IV-TR and DSM-5 diagnostic models for personality disorder: Evaluation of strategies for establishing diagnostic thresholds. *Journal of Psychiatric Practice*, *19*(3), 179–193. https://doi.org/10.1097/01.pra.0000430502.78833.06

Morey, L. C., Skodol, A. E., & Oldham, J. M. (2014). Clinician judgments of clinical utility: A comparison of DSM-IV-TR personality disorders and the alternative model for DSM-5 personality disorders. *Journal of Abnormal Psychology, 123*(2), 398-405. https://doi.org/10.1037/a0036481

Morey, L. C., & Stagner, B. H. (2012). Narcissistic pathology as core personality dysfunction: Comparing the DSM-IV and the DSM-5 proposal for narcissistic personality disorder. *Journal of Clinical Psychology, 68*(8), 908-921. https://doi.org/10.1002/jclp.21895

Mościcki, E. K., Clarke, D. E., Kuramoto, S. J., Kraemer, H. C., Narrow, W. E., Kupfer, D. J., & Regier, D. A. (2013). Testing DSM-5 in routine clinical practice settings: Feasibility and clinical utility. *Psychiatric Services*, *64*(10), 952–960. https://doi.org/10.1176/appi.ps.201300098

Mukolo, A., & Heflinger, C. A. (2011). Factors associated with attributions about child health conditions and social distance preference. *Community Mental Health Journal, 47*(3), 286-299. https://doi.org/10.1007/s10597-010-9325-1

Mulder, C. L., & Kortrijk, H. E. (2012). De invloed van de duur van behandeling op het interpreteren van ROM-metingen bij ACT. *Tijdschrift voor Psychiatrie*, 191-196.

Mulder, R., & Tyrer, P. (2019). Diagnosis and classification of personality disorders: Novel approaches. *Current Opinion in Psychiatry, 32*(1), 27-31. https://doi.org/10.1097/yco.0000000000000461

Mullins-Sweatt, S. N., & Widiger, T. A. (2009). Clinical utility and DSM-V. *Psychological Assessment*, *21*(3), 302-312. https://doi.org/10.1037/a0016607

Mullins-Sweatt, S. N., & Widiger, T. A. (2011). Clinician's judgments of the utility of the DSM-IV and five-factor models for personality disordered patients. *Journal of Personality Disorders*, *25*(4), 463–477. https://doi.org/10.1521/pedi.2011.25.4.463

Muthén, L. K., & Muthén, B. O. (2010). *MPLUS user's guide (6th ed.).* Author.

Natoli, A. P. (2019). The DSM's reconnection to psychoanalytic theory through the alternative model for personality disorders. *Journal of the American Psychoanalytic Association, 67*(6), 1023-1045. https://doi.org/10.1177/0003065120903060

Natoli, A. P., Bach, B., Behn, A., Cottin, M., Gritti, E. S., Hutsebaut, J., & Lapalme, M. (2022). Multinational evaluation of the measurement invariance of the Level of Personality Functioning Scale–brief form 2.0: Comparison of student and community samples across seven countries. *Psychological Assessment, 34*(12), 1112– 1125. https://doi.org/10.1037/pas0001176

Nelson, S. M., Huprich, S. K., Meehan, K. B., Siefert, C., Haggerty, G., Sexton, J., & Jackson, J. (2018). Convergent and discriminant validity and utility of the DSM-5 Levels of Personality Functioning Questionnaire (DLOPFQ): Associations with medical health care provider ratings and measures of physical health. *Journal of Personality Assessment*, *100*(6), 671–679. https://doi.org/10.1080/00223891.2018.1492415

Nelson-Gray, R. O. (2003). Treatment utility of psychological assessment. *Psychological Assessment*, *15*(4), 521-531. https://doi.org/10.1037/1040-3590.15.4.521

Norcross, J. C., & Lambert, M. J. (2018). Psychotherapy relationships that work III. *Psychotherapy, 55*(4), 303-315. https://doi.org/10.1037/pst0000193

Nugter, M. A., & Buwalda, V. J. (2012). Background and possible use of ROM in mental health care. *Tijdschrift voor Psychiatrie, 54*(2), 111-120.

Ohse, L., Zimmermann, J., Kerber, A., Kampe, L., Mohr, J., Kendlbacher, J., & Hörz-Sagstetter, S. (2022). Reliability, structure, and validity of module I (personality functioning) of the Structured Clinical Interview for the alternative DSM–5 model for personality disorders (SCID-5-AMPD-I). *Personality Disorders: Theory, Research, and Treatment*, *14*(3), 287-299. https://doi.org/10.1037/per0000576

Paap, M. C., Pedersen, G., Kvarstein, E., & Hummelen, B. (2023). Evaluating the construct validity of the Norwegian version of the level of personality functioning scale–brief form 2.0 in a large clinical sample. *Journal of Personality Assessment,* 1-11. https://doi.org/10.1080/00223891.2023.2182694

Perry, B. L., Pescosolido, B. A., Martin, J. K., McLeod, J. D., & Jensen, P. S. (2007). Comparison of public attributions, attitudes, and stigma in regard to depression among children and adults. *Psychiatric Services, 58*(5), 632-635. https://doi.org/10.1176/ps.2007.58.5.632

Pilkonis, P. A., Hallquist, M. N., Morse, J. Q., & Stepp, S. D. (2011). Striking the (im) proper balance between scientific advances and clinical utility: Commentary on the DSM–5 proposal for personality disorders. *Personality Disorders: Theory, Research, and Treatment, 2*(1), 68–82. https://doi.org/10.1037/a0022226

Pilkonis, P. A., Heape, C. L., Proietti, J. M., Clark, S. W., McDavid, J. D., & Pitts, T. E. (1995). The reliability and validity of two structured diagnostic interviews for personality disorders. *Archives of General Psychiatry, 52*(12), 1025–1033. https://doi.org/10.1001/archpsyc.1995 .03950240043009

Pincus, A. L. (2011). Some comments on nomology, diagnostic process, and narcissistic personality disorder in the *DSM–5* proposal for personality and personality disorders. *Personality Disorders: Theory, Research, and Treatment, 2*(1), 41–53. https://doi.org/10.1037/a0021191

Pincus, A. L., Dowgwillo, E. A., & Greenberg, L. S. (2016). Three cases of narcissistic personality disorder through the lens of the DSM-5 alternative model for personality disorders. *Practice Innovations, 1*(3), 164-177. https://doi.org/10.1037/pri0000025

Pincus, A. L., Roche, M. J., & Good, E. W. (2015). Pathological narcissism and narcissistic personality disorder. In P. H. Blaney, R. F. Krueger, & T. Millon (Eds.), *Oxford textbook of psychopathology* (Vol. 6, pp. 791–813). Oxford University Press.

Poston, J. M., & Hanson, W. E. (2010). Meta-analysis of psychological assessment as a therapeutic intervention. *Psychological Assessment, 22*(2), 203-212. https://doi.org/10.1037/a0018679

Preti, E., Di Pierro, R., Costantini, G., Benzi, I. M. A., Panfilis, C. de, & Madeddu, F. (2018). Using the Structured Interview of Personality Organization for DSM-5 Level of Personality Functioning rating performed by inexperienced raters. *Journal of Personality Assessment*, *100*(6), 621–629. https://doi.org/10.1080/00223891.2018.1448985

Pretzer, J. (1990). Borderline personality disorder. In T. A. Beck, A. Freeman (Eds). *Cognitive therapy of personality disorders* (pp. 176-207). The Guilford Press, New York.

Ranøyen, I., Lydersen, S., Larose, T. L., Weidle, B., Skokauskas, N., Thomsen, P. H., & Indredavik, M. S. (2018). Developmental course of anxiety and depression from adolescence to young adulthood in a prospective Norwegian clinical cohort. *European Child and Adolescent Psychiatry, 27*(11), 1413-1423. https://doi.org/10.1007/s00787-018-1139-7

Reed, G. M. (2010). Toward ICD-11: Improving the clinical utility of WHO's InternationalClassification of Mental Disorders. *Professional Psychology: Research and Practice*, *41*(6), 457–464. https://doi.org/10.1037/a0021701

Regier, D. A., Narrow, W. E., Clarke, D. E., Kraemer, H. C., Kuramoto, S. J., Kuhl, E. A., & Kupfer, D. J. (2013). DSM-5 field trials in the United States and Canada, Part II: Test- retest reliability of selected categorical diagnoses. *American Journal of Psychiatry, 170*(1), 59-70. https://doi.org/10.1176/appi.ajp.2012.12070999

Ring, D., & Lawn, S. (2019). Stigma perpetuation at the interface of mental health care: a review to compare patient and clinician perspectives of stigma and borderline personality disorder. *Journal of Mental Health*, 1-21. https://doi.org/10.1080/09638237.2019.1581337

Roche, M. J., Jacobson, N. C., & Phillips, J. J. (2018). Expanding the validity of the Level of Personality Functioning Scale observer report and self-report versions across psychodynamic and interpersonal paradigms. *Journal of Personality Assessment, 100*(6), 571-580. https://doi.org/10.1080/00223891.2018.1475394

Roche, M. J., & Jaweed, S. (2021). Comparing measures of Criterion A to better understand incremental validity in the alternative model of personality disorders. *Assessment*, *30*(3), 689-705. https://doi.org/10.1177/10731911211059763

Rodriguez-Seijas, C., Ruggero, C., Eaton, N. R., & Krueger, R. F. (2019). The DSM-5 alternative model for personality disorders and clinical treatment: A review. *Current Treatment Options in Psychiatry, 6*(4), 284-298. https://doi.org/10.1007/s40501-019-00187-7

Rossi, G., & Diaz-Batanero, C. (2023). Differentiation of Self and Interpersonal Functioning with the Level of Personality Functioning Scale–Brief Form 2.0. *Journal of Personality Assessment*, 1-12. https://doi.org/10.1080/00223891.2023.2218931

Samuel, D. B. (2015). A review of the agreement between clinicians' personality disorder diagnoses and those from other methods and sources. *Clinical Psychology: Science and Practice, 22*(1), 1-19. https://doi.org/10.1111/cpsp.12088

Samuel, D. B., & Widiger, T. A. (2006). Clinicians' judgments of clinical utility: A comparison of the DSM-IV and five-factor models. *Journal of Abnormal Psychology*, *115*(2), 298–308. https://doi.org/10.1037/0021-843X.115.2.298

Scheepers, F. (2021). *Mensen zijn ingewikkeld: een pleidooi voor acceptatie van de werkelijkheid en het loslaten van model denken.* Singel Uitgeverijen.

Schmeck, K., Schlüter-Müller, S., Foelsch, P. A., & Doering, S. (2013). The role of identity in the DSM-5 classification of personality disorders. *Child and Adolescent Psychiatry and Mental Health, 7*(1), 1-11. https://doi.org/10.1186/1753-2000-7-27

Scott, J., & Henry, C. (2017). Clinical staging models: From general medicine to mental disorders. *BJPsych Advances, 23*(5), 292-299. https://doi.org/10.1192/apt.bp.116.016436

Sellbom, M., Carmichael, K. L. C., & Liggett, J. (2017). Examination of DSM-5 Section III avoidant personality disorder in a community sample. *Personality and Mental Health, 11*(4), 299–313. https://doi.org/10.1002/pmh

Sellbom, M., Sansone, R. A., Songer, D. A., & Anderson, J. L. (2014). Convergence between DSM-5 Section II and Section III diagnostic criteria for borderline personality disorder. *Australian and New Zealand Journal of Psychiatry, 48*(4), 325–332. https://doi.org/10.1177/0004867413511997

Serman, N., Johnson, J. G., Geller, P. A., Kanost, R. E., & Zacharapoulou, H. (2002). Personality disorders associated with substance abuse among American and Greek adolescents. *Adolescence, 37*(148), 841-854.

Shankman S. A., Funkhouser, C. J., Klein, D. N., Davila, J., Lerner, D., & Hee, D. (2018) Reliability and validity of severity dimensions of psychopathology assessed using the Structured Clinical Interview for DSM-5 (SCID). *International Journal of Methods in Psychiatric Research, 27*(1), e1590. https://doi.org/10.1002/mpr.1590

Sharp, C. (2016). Current trends in BPD research as indicative of a broader sea-change in psychiatric nosology. *Personality Disorders: Theory, Research, and Treatment, 7*(4), 334-343. https://doi.org/10.1037/per0000199

Sharp, C., & Wall, K. (2021). DSM-5 level of personality functioning: Refocusing personality disorder on what it means to be human. *Annual review of clinical psychology, 17*(1), 313-337. https://doi.org/10.1146/annurev-clinpsy-081219-105402

Sharp, C., Wright, A. G., Fowler, J. C., Frueh, B. C., Allen, J. G., Oldham, J., & Clark, L. A. (2015). The structure of personality pathology: Both general ('g') and specific ('s') factors?. *Journal of abnormal psychology, 124*(2), 387-398. https://doi.org/10.1037/abn0000033.supp

Shea, M. T., Stout, R., Gunderson, J., Morey, L. C., Grilo, C. M., McGlashan, T., & Keller, M. B. (2002). Short-term diagnostic stability of schizotypal, borderline, avoidant, and obsessive-compulsive personality disorders. *American Journal of Psychiatry, 159*(12), 2036-2041. https://doi.org/10.1176/appi.ajp.159.12.2036

Shedler, J., Beck, A., Fonagy, P., Gabbard, G. O., Gunderson, J., Kernberg, O., & Westen, D. (2010). Personality disorders in DSM-5. *American Journal of Psychiatry, 167*(9), 1026-1028. https://doi.org/10.1176/appi.ajp.2010.10050746

Sheehan, L., Nieweglowski, K., & Corrigan, P. (2016). The stigma of personality disorders. *Current Psychiatry Reports, 18*(1), 1-7. https://doi.org/10.1007/s11920-015-0654-1

Shiner, R. L., & Allen, T. A. (2013). Assessing personality disorders in adolescents: seven guiding principles. *Clinical Psychology Science and Practice 20*(4), 361-377. https://doi.org/10.1111/cpsp.12047

Sieswerda, S., Arntz, A., Mertens, I., & Vertommen, S. (2007). Hypervigilance in patients with borderline personality disorder: Specificity, automaticity, and predictors. *Behaviour Research and Therapy, 45*(5), 1011-1024. https://doi.org/10.1016/j.brat.2006.07.012

Simms, L. J., & Calabrese, W. R. (2016). Incremental validity of the DSM-5 Section III personality disorder traits with respect to psychosocial impairment. *Journal of Personality Disorders, 30*(1), 95–111. https://doi.org/10.1521/pedi_2015_29_185

Skodol, A. E. (2014) Personality disorder classification: Stuck in neutral, how to move forward? *Current Psychiatry Reports, 16*(10), 480-490. https://doi.org/10.1007/s11920-014-0480-x

Skodol, A., & Bender, D. S. (2009). The Future of Personality Disorders in DSM-V? *American Journal of Psychiatry*, *166*(4), 388–391. https://doi.org/10.1176/appi.ajp.2009.09010090

Skodol, A. E., Bender, D. S., & Morey, L. C. (2014). Narcissistic personality disorder in DSM-5. *Personality Disorders: Theory, Research, and Treatment*, *5*(4), 422–427. https://doi.org/10.1037/per0000023

Skodol, A. E., Bender, D. S., Oldham, J. M., Clark, L. A., Morey, L. C., Verheul, R., &

Siever, L. J. (2011). Proposed changes in personality and personality disorder assessment and diagnosis for DSM–5 Part II: Clinical application. *Personality Disorders: Theory, Research, and Treatment, 2*(1), 23-40. http://dx.doi.org/10.1037/a0021892

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

Skodol, A. E., Morey, L. C., Bender, D. S., & Oldham, J. M. (2015). The alternative DSM-5 model for personality disorders: A clinical application. *American Journal of Psychiatry, 172*(7), 606-613. https://doi.org/10.1176/appi.ajp.2015.14101220

Skodol, A. E., Oldham, J. M., Bender, D. S., Dyck, I. R., Stout, R. L., Morey, L. C., & Gunderson, J. G. (2005). Dimensional representations of DSM-IV personality disorders: Relationships to functional impairment. *American Journal of Psychiatry, 162*(10), 1919-1925. doi:10.1176/appi.ajp.162.10.1919

Sleep, C. E., Lynam, D. R., Widiger, T. A., Crowe, M. L., & Miller, J. D. (2019). An evaluation of DSM-5 Section III personality disorder Criterion A (impairment) in accounting for psychopathology. *Psychological Assessment*, *31*(10), 1181–1191. https://doi.org/10.1037/pas0000620

Sleep, C. E., Weiss, B., Lynam, D. R., & Miller, J. D. (2020). The DSM-5 section III personality disorder criterion a in relation to both pathological and general personality traits. *Personality Disorders: Theory, Research, and Treatment*, *11*(3), 202–212. https://doi.org/10.1037/per0000383

Soeteman, D. I., Verheul, R., & Busschbach, J. J. (2008). The burden of disease in personality disorders: Diagnosis-specific quality of life. *Journal of Personality Disorders, 22*(3), 259-268. https://doi.org/10.1521/pedi.2008.22.3.259

Somma, A., Borroni, S., Gialdi, G., Carlotta, D., Emanuela Giarolli, L., Barranca, M., & Fossati, A. (2020). The inter-rater reliability and validity of the Italian translation of the Structured Clinical Interview for DSM-5 Alternative Model for Personality Disorders module I and module II: A preliminary report on consecutively admitted psychotherapy outpatients. *Journal of Personality Disorders*, *34*(Supplement C), 95– 123. https://doi.org/10.1521/pedi_2020_34_511

Somma, A., Krueger, R. F., Markon, K. E., Alajmo, V. B. M., Arlotta, E., Beretta, S., & Fossati, A. (2019). DSM-5 Alternative Model of Personality Disorder dysfunctional personality traits as predictors of self-reported aggression in an Italian sample of consecutively admitted, personality-disordered psychotherapy patients. *Journal of Personality Disorders, 34*(Supplement C), 5-24. https://doi.org/10.1521/pedi_2019_33_430

Somma, A., Krueger, R. F., Markon, K. E., Borroni, S., & Fossati, A. (2019). Schizotypy from the perspective of the DSM-5 alternative model of personality traits: A study on a sample of 1056 Italian adult university students. *Journal of Psychopathology and Behavioral Assessment, 41*(4), 560–573. https://doi.org/10.1007/s10862-019-09718-1

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language, 25*(4), 359-393. https://doi.org/10.1111/j.1468-0017.2010.01394.x

Spitzer, C., Mueller, S., Kerber, A., Hutsebaut, J., Braehler, E., & Zimmermann, J. (2021). The German version of the level of personality functioning scale-brief form 2.0 (LPFS-BF): latent structure, convergent validity and norm values in the general population. *Psychotherapie, Psychosomatik, Medizinische Psychologie, 71*(7), 284-293. https://doi.org/10.1055/a-1343-2396

Stoffers, J. M., Vollm, B. A., Rucker, G., Timmer, A., Huband, N., & Lieb, K. (2012). Psychological therapies for people with borderline personality disorder. *Cochrane Database Systematic Review, 8*, CD005652.

Stone, L. E., Segal, D. L., & Noel, O. R. (2021). Psychometric evaluation of the Levels of Personality Functioning Scale—Brief Form 2.0 among older adults. *Personality Disorders: Theory, Research, and Treatment, 12*(6), 526-533. https://doi.org/10.1037/per0000413

Storebø, O. J., Stoffers-Winterling, J. M., Völlm, B. A., Kongerslev, M. T., Mattivi, J. T., Jørgensen, M. S., & Simonsen, E. (2020). Psychological therapies for people with borderline personality disorder. *Cochrane Database of Systematic Reviews*, 2020(5). https://doi.org/10.1002/14651858.CD012955.pub2

Thylstrup, B., Simonsen, S., Nemery, C., Simonsen, E., Noll, J. F., Myatt, M. W., & Hesse, M. (2016). Assessment of personality-related levels of functioning: A pilot study of clinical assessment of the DSM-5 level of personality functioning based on a semi-structured interview. *BMC Psychiatry*, *16*(1), 1-8. https://doi.org/10.1186/s12888-016-1011-6

Tyrer, P. (2005). The problem of severity in the classification of personality disorder. *Journal of Personality Disorders, 19*(3), 309-314. https://doi.org/10.1521/pedi.2005.19.3.309

Tyrer, P. (2012). Diagnostic and Statistical Manual of Mental Disorders: A classification of personality disorders that has had its day. *Clinical Psychology and Psychotherapy, 19*(5), 372–374. https://doi.org/10.1002/cpp.1810

Ustun, T. B., Kostanjesek, N., Chatterji, S., Rehm, J. & World Health Organization. (2010). Measuring health and disability: Manual for WHO Disability Assessment Schedule (WHODAS 2.0) / edited by T.B. Üstün, N. Kostanjsek, S. Chatterji, J.Rehm. World Health Organization. https://apps.who.int/iris/handle/10665/43974

Van Beek, N., & Verheul, R. (2008). Motivation for treatment in patients with personality disorders. *Journal of Personality Disorders, 22*(1), 89-100. https://doi.org/10.1521/pedi.2008.22.1.89

Van Buuren, S., & C. G. M. Groothuis-Oudshoorn. (2011). "Mice": Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software 45*(3): 1–67. https://doi.org/10.18637/jss.v045.i03

Van der Heijden, P., Ingenhoven, T., Berghuis, H., & Rossi, G. (2014). *DSM-5 persoonlijkheidsvragenlijst verkorte vorm[DSM-5 personality inventory brief form]: PID-5—BF-NL. Dutch translation of The Personality Inventory for DSM-5® - Brief Form (PID-5-BF) Adult, 2011* (American Psyhciatric Association). Boom.

Van Os, J., Kahn, R., Denys, D., Schoevers, R. A., Beekman, A. T. F., Hoogendijk, W. J. G., & Leentjens, A. F. G. (2012). ROM: gedragsnorm of dwangmaatregel? Overwegingen bij het themanummer over routine outcome monitoring. *Tijdschrift voor Psychiatrie, 54*(3), 245-253.

Ventura, J., Liberman, R. P., Green, M. F., Shaner, A., & Mintz, J. (1998). Training and quality assurance with the Structured Clinical Interview for DSM-IV (SCID-I/P). *Psychiatry Research, 79*(2), 163-173. https://doi.org/10.1016/s0165-1781(98)00038-9

Verheul, R., Andrea, H., Berghout, C. C., Dolan, C., Busschbach, J. J. V., van der Kroft, P. J. & Fonagy, P. (2008). Severity Indices of Personality Problems (SIPP-118): Development, factor structure, reliability, and validity. *Psychological Assessment, 20*(1), 23–34. https://doi.org/10.1037/1040-3590.20.1.23

Verheul, R., Bartak, A., & Widiger, T. (2007). Prevalence and construct validity of personality disorder not otherwise specified (PDNOS*). Journal of Personality Disorders, 21*(4), 359-370. https://doi.org/10.1521/pedi.2007.21.4.359

Verheul, R., & Widiger, T. A. (2004). A meta-analysis of the prevalence and usage of the personality disorder not otherwise specified (PDNOS) diagnosis. *Journal of Personality Disorders, 18*(4), 309-319. https://doi.org/10.1521/pedi.2004.18.4.309

Videler, A. C., Hutsebaut, J. Schulkens, J. E. M., Sobczak, S., & van Alphen, S. P. J. (2019). A life span perspective on borderline personality disorders. *Current Psychiatry Opinions, 21,* 1-8. https://doi.org/10.1007/s11920-019-1040-1

Wampold, B. E. (2015). How important are the common factors in psychotherapy? An update. *World Psychiatry, 14*(3), 270-277. https://doi.org/10.1002/wps.20238

Watters, C. A., Bagby, R. M., & Sellbom, M. (2019). Meta-analysis to derive an empirically based set of personality facet criteria for the alternative DSM-5 model for personality disorders. *Personality Disorders: Theory, Research, and Treatment, 10*(2), 97–104. https://doi.org/10.1037/per0000307

Waugh, M. H. (2019). Clinical utility and application of the AMPD. In C. J. Hopwood, A. L. Mulay, & M. H. Waugh (Eds.), *The DSM-5 Alternative Model for Personality Disorders: Integrating multiple paradigms of personality assessment* (pp 96-140). Routledge

Waugh, M. H., Bishop, & Schmidt. (2019). Clinical cases. In C. J. Hopwood, A. L. Mulay, & M. H. Waugh (Eds.), *The DSM-5 Alternative Model for Personality Disorders: Integrating multiple paradigms of personality assessment* (pp 96-140). Routledge.

Waugh, M. H., Hopwood, C. J., Krueger, R. F., Morey, L. C., Pincus, A. L., & Wright, A. G. (2017). Psychological assessment with the DSM–5 Alternative Model for Personality Disorders: Tradition and innovation. *Professional Psychology: Research and Practice, 48*(2), 79-89. https://doi.org/10.1037/pro0000071

Weekers, L. C., Hutsebaut, J., Bach, B., & Kamphuis, J. H. (2020). Scripting the DSM-5 Alternative Model for Personality Disorders assessment procedure: A clinically feasible multi-informant multi-method approach. *Personality and Mental Health, 14*(3), 304-318. https://doi.org/10.1002/pmh.1481

Weekers, L. C., Hutsebaut, J., & Kamphuis, J. H. (2019). The Level of Personality Functioning Scale-Brief Form 2.0: Update of a brief instrument for assessing level of personality functioning. *Personality and Mental Health, 13*(1), 3-14. https://doi.org/10.1002/pmh.1434

Weekers, L. C., Hutsebaut, J., Kamphuis, J. H. (2021). Client and clinical utility of the assessment of personality disorders. *Journal of Nervous and Mental Disorders, 209*(11), 846-850. https://doi.org/10.1097/nmd.0000000000001398

Weekers, L. C., Hutsebaut, J., Zimmermann, J., & Kamphuis, J. H. (2021). Changes in the classification of personality disorders: Comparing the DSM–5 Section II personality disorder model to the alternative model for personality disorders using structured clinical interviews. *Personality Disorders: Theory, Research, and Treatment, 13*(5), 527-535. https://doi.org/10.1037/per0000512

Weekers, L. C., Sellbom, M., Hutsebaut, J., Simonsen, S., & Bach, B. (2022). Normative data for the LPFS-BF 2.0 derived from the Danish general population and relationship with psychosocial impairment. *Personality and Mental Health, 17*(2), 157-164. https://doi.org/10.1002/pmh.1570

Weekers, L. C., Verhoeff, S. C. E., Kamphuis, J. H., & Hutsebaut, J. (2020). Assessing Criterion A in adolescents using the Semistructured Interview for Personality Functioning DSM–5. *Personality Disorders: Theory, Research, and Treatment, 12*(4), 312–319. https://doi.org/10.1037/per0000454

Weertman, A., Arntz, A., Dreessen, L., van Velzen, C., & Vertommen, S. (2003). Short-interval test-retest interrater reliability of the Dutch version of the structured clinical interview for DSM-IV personality disorders (SCID-II). *Journal of Personality Disorders, 17*(6), 562-567. https://doi.org/10.1521/pedi.17.6.562.25359

Weertman, A., Arntz, A., & Kerkhofs, M. (1996). *Gestructureerd klinisch interview voor DSM-IV as II persoonlijkheidsstoornissen [Structured clinical interview for DSM-IV Axis II personality disorders].* Harcourt Test Publishers.

Wertz, J., Caspi, A., Ambler, A., Arseneault, L., Belsky, D. W., Danese, A., & Moffitt, T. E. (2020). Borderline symptoms at age 12 signal risk for poor outcomes during the transition to adulthood: Findings from a genetically sensitive longitudinal cohort study. *Journal of the American Academy of Child and Adolescent Psychiatry, 59*(10), 1165–1177.e2. https://doi.org/10.1016/j.jaac.2019.07.005

Westen, D., Shedler, J., Durett, C., Glass, S., & Martens, A. (2003). Personality diagnoses in adolescence: DSM-IV axis II diagnoses and an empirically derived alternative. *American Journal of Psychiatry, 160*(5), 952-966. https://doi.org/10.1176/appi.ajp.160.5.952

Widiger, T. A. (2013). A postmortem and future look at the personality disorders in DSM–5. *Personality Disorders: Theory, Research, and Treatment, 4*(4), 382–387. https://doi.org/10.1037/per0000030

Widiger, T. A., & Samuel, D. B. (2005). Diagnostic categories or dimensions? A question for the Diagnostic and Statistical Manual for Mental Disorders – Fifth Edition. *Journal of Abnormal Psychology, 114*(4), 494-504.

Widiger, T. A., & Trull, T. J. (2007). Plate tectonics in the classification of personality disorder: Shifting to a dimensional model. *American Psychologist, 62*(2), 71–83. https://doi.org/10.1037/0003-066X.62.2.71

World Health Organization. (2019). *ICD-11: International classification of diseases* (11th revision). https://icd.who.int/

Wright, A. G., Calabrese, W. R., Rudick, M. M., Yam, W. H., Zelazny, K., Williams, T. F., & Simms, L. J. (2015). Stability of the DSM-5 Section III pathological personality traits and their longitudinal associations with psychosocial functioning in personality disordered individuals. *Journal of Abnormal Psychology, 124*(1), 199.

Wright, A. G. C., Pincus, A. L., Thomas, K. M., Hopwood, C. J., Markon, K. E., & Krueger, R. F. (2013). Conceptions of narcissism and the DSM-5 pathological personality traits. *Assessment, 20*(3), 339–352. https://doi.org/10.1177/1073191113486692

Wright, A. G. C., 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*(4), 951–957. https://doi.org/10.1037/a0027669

Wygant, D. B., Sellbom, M., Sleep, C. E., Wall, T. D., Applegate, K. C., Krueger, R. F., & Patrick, C. J. (2016). Examining the DSM-5 alternative personality disorder model operationalization of antisocial personality disorder and psychopathy in a male correctional sample. *Personality Disorders: Theory, Research, and Treatment, 7*(3), 229–239. https://doi.org/10.1037/per0000179

Young, S., & Beazley, P. (2023). Interrater reliability of criterion A of the alternative model for personality disorders (DSM-5 – Section III): A meta-analysis. *Personality disorders: Theory, Research, and Treatment.* https://doi.org/10.1037/per0000631

Zachar, P., Krueger, R. F., & Kendler, K. S. (2015). Personality disorder in DSM-5: an oral history. *Psychological Medicine, 46*(1), 1–10. https://doi.org/10.1017/s0033291715001543

Zanarini, M. C., Frankenburg, F. R., Reich, D. B., & Fitzmaurice, G. (2010). Time to attainment of recovery from borderline personality disorder and stability of recovery: A 10-year prospective follow-up study. *American Journal of Psychiatry, 167*(6), 663- 667. https://doi.org/10.1176/appi.ajp.2009.09081130

Zanarini, M. C., Frankenburg, F. R., Reich, D. B., & Fitzmaurice, G. (2012). Attainment and stability of sustained symptomatic remission and recovery among patients with borderline personality disorder and axis II comparison subjects: a 16-year prospective follow-up study. *American Journal of Psychiatry, 169*(5), 476-483. doi: 10.1176/appi.ajp.2011.11101550

Zettl, M., Taubner, S., Hutsebaut, J., & Volkert, J. (2019). Psychometrische evaluation der deutschen version des Semistrukturierten Interviews zur Erfassung der DSM-5 Persönlichkeitsfunkionen (STiP-5.1). *Psychotherapie Psychosomatik Medizinische Psychologie, 69*(12), 499-504. https://doi.org/10.1055/a-1010-6887

Zimmerman, M. (1994). Diagnosing personality disorders: A review of issues and research methods. *Archives of General Psychiatry, 51*(3), 225–245. https://doi.org/10.1001/archpsyc.1994.03950030061006

Zimmerman, M., Rothschild, L., & Chelminski, I. (2005). The prevalence of DSM-IV personality disorders in psychiatric outpatients. *American Journal of Psychiatry, 162*(10), 1911-1918. https://doi.org/10.1176/appi.ajp.162.10.1911

Zimmermann, J., Benecke, C., Bender, D. S., Skodol, A. E., Schauenburg, H., Cierpka, M., & Leising, D. (2014). Assessing DSM-5 Level of Personality Functioning from videotaped clinical interviews: A pilot study with untrained and clinically inexperienced students. *Journal of Personality Assessment*, *96*(4), 397–409. https://doi.org/10.1080/00223891.2013.852563

Zimmermann, J., Böhnke, J. R., Eschstruth, R., Mathews, A., Wenzel, K., & Leising, D. (2015). The latent structure of personality functioning: Investigating criterion a from the alternative model for personality disorders in DSM–5. *Journal of Abnormal Psychology, 124*(3), 532-548. https://doi.org/10.1037/abn0000059

Zimmermann, J., Hopwood, C. J., & Krueger, R. F. (2022). The DSM-5 Level of Personality Functioning Scale. In R. F. Krueger & P. H. Blaney (Eds.), *Oxford Textbook of Psychopathology (4th ed.).* Oxford University Press.

Zimmermann, J., Kerber, A., Rek, K., Hopwood, C. J., & Krueger, R. F. (2019). A brief but comprehensive review of research on the alternative DSM 5 model for personality disorders. *Current Psychiatry Reports, 21*(9), 1-19. https://doi.org/10.1007/s11920-019-1079-z

Zimmermann, J., Müller, S., Bach, B., Hutsebaut, J., Hummelen, B., & Fischer, F. (2020). A common metric for self-reported severity of personality disorder. *Psychopathology, 53*(3), 168-178. https://doi.org/10.1159/000507377

Curriculum Vitae

Laura Weekers was born on April 5, 1991 in Hellevoetsluis, the Netherlands. In 2010 she started her Clinical Psychology study at the Erasmus University Rotterdam. She completed a clinical internship at a mental health care center for mood disorders (PsyQ Rotterdam). In 2014-2015, she wrote her master thesis on the effectiveness of Mentalization-Based Treatment for adolescents at de Viersprong, a mental health care facility for the assessment and treatment of personality disorders. After graduating in 2015, she started working parttime as a research assistant at de Viersprong and as a social worker at a psychiatric unit (Delta psychiatrisch centrum, now called Antes). As of 2016, she started working for the Viersprong fulltime: half of her time as a junior researcher at de Viersprong Institute for Studies on Personality Disorders and half of her time as a psychologist in the intake and assessment unit. Assessment and personality disorders have since been a specific interest of her, both as a researcher and as a clinician. In 2024 she started a training to become a health care psychologist (Gezondheidszorgpsycholoog).

PhD Portfolio

Name PhD student:	Laura Weekers
PhD period:	2017-2024
Promotor(s):	Jan Henk Kamphuis
	Joost Hutsebaut

Publications in this thesis

Weekers, L. C., Hutsebaut, J., Bach, B., & Kamphuis, J. H. (2020). Scripting the DSM-5 Alternative Model for Personality Disorders assessment procedure: A clinically feasible multi-informant multimethod approach. *Personality and Mental Health, 14*(3), 304-318.

Weekers, L. C., Hutsebaut, J., De Saeger, H., & Kamphuis, J. H. (2024). Comparing the clinical utility of the Alternative Model for Personality Disorders to the Section II personality disorder model: A Randomized Controlled Trial. *Personality Disorders: Theory, Research, and Treatment.*

Weekers, L. C., Hutsebaut, J. & Kamphuis, J. H. (2019). The Level of Personality Functioning Scale – Brief Form 2.0 (LPFS-BF 2.0): Update of a brief instrument for assessing level of personality functioning. *Personality and Mental Health, 13*(1), 3-14.

Weekers, L. C., Hutsebaut, J., Kamphuis, J. H. (2021). Client and clinical utility of the assessment of personality disorders. *Journal of Nervous and Mental Disorders, 209*(11), 846-850.

Weekers, L. C., Hutsebaut, J., Rovers, J. M. C., & Kamphuis, J. H. (2023). Head to head comparison of the Alternative Model for Personality Disorders and Section II personality disorder model in terms of predicting patient outcomes one year later. *Personality Disorders: Theory, Research, and Treatment.*

Weekers, L. C., Hutsebaut, J., Zimmermann, J., & Kamphuis, J. H. (2021). Changes in the classification of personality disorders: Comparing the DSM–5 Section II personality disorder model to the Alternative Model for Personality Disorders using structured clinical interviews. *Personality Disorders: Theory, Research, and Treatment, 13*(5), 527-535.

Weekers, L. C., Verhoeff, S. C. E., Kamphuis, J. H., & Hutsebaut, J. (2020). Assessing Criterion A in adolescents using the Semistructured Interview for Personality Functioning DSM – 5. *Personality Disorders: Theory, Research, and Treatment, 12*(4), 312-319.

Publications not in this thesis

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

Hutsebaut, J., Weekers, L. C., Tuin, N., Apeldoorn, J. S., & Bulten, E. (2021). Assessment of ICD-11 personality disorder severity in forensic patients using the Semi-Structured Interview for Personality Functioning DSM-5 (STiP5.1): Preliminary findings. *Frontiers in Psychiatry, 12*, 617702.

Rossi, G., Weekers, L. C., & Hutsebaut, J. (2021). Resilient, undercontrolled, and overcontrolled personality types based upon DSM-5 maladaptive personality traits. *Heliyon, 7*(5), e06938.

Weekers, L. C., Sellbom, M., Hutsebaut, J., Simonsen, S., & Bach, B. (2022). Normative data for the LPFS-BF 2.0 derived from the Danish general population and relationship with psychosocial impairment. *Personality and Mental Health, 17*(2), 157-164.

Presentations		Year
20th International Congress of the European Society for Child and Adolescent Psychiatry	Symposium: Incremental validity of the Alternative Model for Personality Disorders: Predicting disability one year after assessment	2023
Vlaams GGZ Congres, Antwerpen	Symposium: Vroege detectie van beperkingen in persoonlijkheidsfunctioneren	2022
Society for Personality Assessment Convention	Symposium: Clinical utility and implementation of the AMPD	2022
Webinar de Viersprong	Het alternatieve model voor persoonlijkheidsstoornissen - van instrumentontwikkeling tot implementatie	2022
18th congress of the International Society for the Study of Personality Disorders	Symposium: Changes in the classification of personality disorders	2021
17th congress of the International Society for the Study of Personality Disorders	Poster: Clinical utility of the AMPD	2019
5th International Congress on BPD and allied disorders	Symposium: Semistructured interview for Personality Functioning DSM-5 (STiP 5.1) in an adolescent sample, preliminary findings	2018
16th congress of the International Society for the Study of Personality Disorders	Poster: The Level of Personality Functioning Scale Brief Form 2.0: a brief instrument for assessing the level of personality functioning	2017

Research skills	Year
European Society for Studies on Personality Disorders Summer School	2019

Teaching activities	Year
Opleiding tot gezondheidszorgpsycholoog, blok Methodische diagnostiek, RINO Zuid. Course: Alternatief model voor persoonlijkheidsstoornissen	2022-2023
Deskundigheidsbevordering persoonlijkheidsstoornissen, Akwa GGZ	2022
Training Alternatief model voor persoonlijkheidsstoornissen: State of the art en klinische toepassing, Viersprong Academy	2020-heden
Gastcollege: The Alternative Model for Personality Disorders, University of Chile	2020

Clinical activities	Year
Crisi Wartegg System level 1, 2 and 3	2020-2021
Live Therapeutic Assessment with an adolescent and his family	2020
Training Semi-gestructureerd interview voor persoonlijkheidsfunctioneren DSM-5 (STiP 5.1)	2017
Basiscursus Rorschach R-PAS	2017
Training DSM-5	2017
Basistraining Therapeutisch Psychologisch Onderzoek	2016
Live Therapeutic Assessment of an adult client	2016
Training MMPI	2015
Training SCID-I & SCID-II	2014-2015

Author contributions

Chapter 2

Weekers, L. C., Hutsebaut, J. & Kamphuis, J. H. (2019). The Level of Personality Functioning Scale – Brief Form 2.0 (LPFS-BF 2.0): Update of a brief instrument for assessing level of personality functioning. *Personality and Mental Health, 13*(1), 3-14.

L.C.W., J.H. and J.H.K. contributed to developing the study concept. L.C.W. performed the data analysis and interpretation. All authors participated in drafting the manuscript.

Chapter 3

Weekers, L. C., Verhoeff, S. C. E., Kamphuis, J. H., & Hutsebaut, J. (2020). Assessing Criterion A in adolescents using the Semistructured Interview for Personality Functioning DSM – 5. *Personality Disorders: Theory, Research, and Treatment, 12*(4), 312-319.

L.C.W., J.H. and J.H.K. contributed to developing the study concept. S.C.E.V and L.C.W. collected the data. L.C.W. performed the data analysis and interpretation. All authors participated in drafting the manuscript.

Chapter 4

Weekers, L. C., Hutsebaut, J., Zimmermann, J., & Kamphuis, J. H. (2021). Changes in the classification of personality disorders: Comparing the DSM–5 Section II personality disorder model to the Alternative Model for Personality Disorders using structured clinical interviews. *Personality Disorders: Theory, Research, and Treatment, 13*(5), 527-535.

L.C.W., J.H. and J.H.K. contributed to developing the study concept. L.C.W. collected the data. L.C.W. and J.Z. performed the data analysis and interpretation. All authors participated in drafting the manuscript.

Chapter 5

Weekers, L. C., Hutsebaut, J., Bach, B., & Kamphuis, J. H. (2020). Scripting the DSM-5 Alternative Model for Personality Disorders assessment procedure: A clinically feasible multi-informant multimethod approach. *Personality and Mental Health*, *14*(3), 304-318.

All authors participated in drafting the manuscript.

Chapter 6

Weekers, L. C., Hutsebaut, J., Kamphuis, J. H. (2021). Client and clinical utility of the assessment of personality disorders. *Journal of Nervous and Mental Disorders, 209*(11), 846-850.

L.C.W., J.H. and J.H.K. contributed to developing the study concept. L.C.W. and J.H. performed the data collection and interpretation. All authors participated in drafting the manuscript.

Chapter 7

Weekers, L. C., Hutsebaut, J., Rovers, J. M. C., & Kamphuis, J. H. (2023). Head to head comparison of the Alternative Model for Personality Disorders and Section II personality disorder model in terms of predicting patient outcomes one year later. *Personality Disorders: Theory, Research, and Treatment*.

L.C.W., J.H. and J.H.K. contributed to developing the study concept. J.M.C.R. and L.C.W. collected the data. L.C.W. performed the data analysis and interpretation. All authors participated in drafting the manuscript.

Chapter 8

Weekers, L. C., Hutsebaut, J., De Saeger, H., & Kamphuis, J. H. (2024). Comparing the clinical utility of the Alternative Model for Personality Disorders to the Section II personality disorder model: A Randomized Controlled Trial. *Personality Disorders: Theory, Research, and Treatment*.

L.C.W., J.H, H.D.S. and J.H.K. contributed to developing the study concept and design. L.C.W. collected the data. L.C.W. performed the data analysis and interpretation. All authors participated in drafting the manuscript.