

Emilia Soroko¹ , Lidia Wanda Cierpiałkowska¹ , Łukasz Mech²,¹Adama Mickiewicza University, Poznań, Poland²Wojewódzki Szpital dla Nerwowo i Psychicznie Chorych „Dziekanka” im. Aleksandra Piotrowskiego, Gniezno, Poland

CORRESPONDING AUTHOR – Emilia Soroko, soroko@amu.edu.pl

Self-report versus clinical ratings using the SWAP-200 in the assessment of personality disorders

ABSTRACT: The relationship between self- and informant reports of personality using psychometric instruments is constantly the focus of attention for researchers in the field of clinical assessment in psychology. The research shows weak agreement between clinicians and patients' assessments of personality disorders (PDs). The current study aimed at the convergence of measurement of PDs using the Shedler-Westen Assessment Procedure (SWAP-200), the self-report Character Styles Questionnaire-R (CSQ-R) and Borderline Personality Inventory (BPI). Paper-pencil questionnaires were administered to 102 inpatients (88.2% female, aged 18-64, M = 38.4) in a voivodeship hospital and outpatient health care centre. The SWAP-200 allowed us to gather expert (clinician) personality ratings basing on the intensive contact with patients. Results show that only a few SWAP-200 PD scales showed low positive correlations with corresponding self-reported PD scales from the CSQ-R. With the canonical correlation analysis, we identified two functions (borderline and internalising) that described similarities between the SWAP-200 and CSQ-R. SWAP-200 Obsessive-Compulsive PD correlated negatively with BPI scales. Consistent with previous studies, the self-report and the clinical assessment were only marginally convergent. Furthermore, OCPD stands out from other disorders in that it correlates positively with health indicators and negatively with some of the other personality disorders. The highest agreement was observed in the description of Borderline PD.

Keywords: *assessment, personality disorders, self-report, borderline, SWAP-200*

INTRODUCTION

The assessment of personality disorders (PDs) has been undergoing a significant transformation in recent years, both regarding the development of dimensional models (Hopwood, Zimmermann, Pincus, & Krueger, 2015; Zimmermann, Kerber, Rek, Hopwood, & Krueger, 2019) and incorporation into the assessment the fact that PDs are dynamic and context-dependent (Hopwood, 2018; Clarkin, Meehan, & Lenzenweger, 2015).

The first point is evident in the maintenance of the categories of PD types in Diagnostic and Statistical Manual of Mental Disorders 5 (DSM-5) Section II, while at the same time, very intently exploring the dimensional

approach within the Alternative Model of PDs (AMPD) Section III. The AMPD model, for example, criterion A offers a diagnostically new (although familiar from clinical practice) way of thinking about the severity of PD, listing the self and interpersonal areas in which disruption may occur at different levels of depth. Changes in a similar direction have been undertaken in the assessment of personality disorders according to the International Classification of Diseases 11th Revision (WHO, 2022). The new approach abandons personality types in favor of three elements, with only the diagnosis of the first being mandatory: personality disorder severity, trait domain qualifiers, and borderline pattern (Bach et al., 2022). Researchers and clinicians, at the current stage of



developing dimensional models, are focusing on creating crosswalks between categorical and dimensional thinking (Simon et al., 2023).

The second point is primarily the observation that PDs and behaviour are not as stable over time as previously thought, and that the expression of personality traits occurs differently depending on external and internal situations. Experiencing this diversity can be one of the complicating factors for self-description in self-report measures (like questionnaires), especially in PDs that are organised at the borderline level (due to active splitting and significant instability of functioning, e.g., Kernberg, 2004). Given that 1) both competing diagnostic models in DSM-5 consider self-report as an important source of information when building the empirical knowledge, and that 2) the type of distortions that may appear in self-report may be related to the situational specificity of PDs, it is worth assessing the convergence of clinicians' ratings and self-report in the context of dimensionally understood PD types and underlying dimensions of psychopathology.

(DIS)AGREEMENT BETWEEN SELF-REPORT AND OTHER-INFORMANT REPORTS

Convergence between self-report and assessment by other informants (both professionals and others who know the patient) is a subject of scientific attention and has an influence on practice. There is empirical evidence along with a theoretical explanation that data from different sources often do not converge (Bornstein, 2017; Bornstein, 2002). Research mainly focuses on the relationship between self and other-informant reports. In the area of child and adolescent assessment, self-reports are compared with assessments by parents, peers, or teachers (Rettew, Lynch, Achenbach, Dumenci, & Ivanova, 2009). Moreover, there are studies in clinical psychology comparing self-report assessment and quantified judgments of clinically experienced observers, such as treating clinicians. Hence, clinicians may use different methods with varying diagnostic properties and consequences, e.g., semistructured diagnostic interviews, clinician-assessed PD diagnoses, and systematised clinician ratings (e.g., SWAP-200). All this complicates the assessment and understanding of research results.

According to Olino and Klein (2015), discrepancies between self- and informant reports can come from at least three sources: self-report biases, access to different information, and psychometric differences across informants. Research on personality traits shows that the convergence between self-report and informant rating (e.g., mothers, fathers, peers) is higher for extraversion than for neuroticism (Harkness, Tellegen, & Waller, 1995), and the highest for extraversion and the lowest for agreeableness (Connolly, Kavanagh, & Viswesvaran, 2007). Who evaluates these characteristics and what factors mediate this also play a role in achieving convergence. According to Vazire (2010), the self is the best judge of neuroticism-related traits, friends are the best judge of intellect-related traits, and people of all

perspectives are equally good at judging extraversion-related traits. Although there is a high degree of construct overlap, both self and observer ratings have a substantial unique variance, which has been partially explained by the duration of acquaintance (strangers vs. close relatives). Kelley, Edens, and Morey (2016) compared parallel versions (for the subject and his/her roommate) of the same screening tool for the assessment of PDs (*Personality Assessment Screener*). The informants only moderately converged (total $r = 0.45$; $p < .01$), with generally greater agreement between perspectives observed for externalising behaviours compared with internalising distress. Again, convergence was associated with characteristics of the dyadic relationship (e.g., length of the acquaintanceship). It is consistent with other notions that the longer the familiarity, the higher the correspondence between self- and informant reports (Connolly, Kavanagh, & Viswesvaran, 2007). The severity of self-reported psychological difficulties and positive impression management were also associated with the convergence.

CLINICIANS' RATINGS AND SELF-REPORT IN PD ASSESSMENT

The effect of divergence among data from different sources is particularly evident among PD measures. A review of the literature reveals that agreement between informant- and self-reports of PD is modest at best (e.g., Klonsky, Oltmanns, & Turkheimer, 2002; Gritti, Samuel, & Lang, 2015; Samuel, Suzuki, Bucher, & Griffin, 2018). From a general perspective, Gritti, Samuel, and Lang (2015) suggest that the discrepancy between self- and clinician ratings are due to the source of information and variability in convergence across the PDs, which in turn derives from the specificity of different PDs (e.g., different levels and types of distress).

The limitations of self-report measures referred to the source of information might be linked to the ego-syntonic nature of many forms of PDs, distortions of self-perception, or the meta-perception of personality traits (Oltmanns, Gleason, Klonsky, & Turkheimer, 2005). The inability to realistically self-assess may be related to difficulties in reality testing or to dissociated or repressed content that is experienced as features of the other but not the self (Davison, Obonsawin, Seils, & Patience, 2003). People with PDs do not recognise the impact that their behaviour has on others, and they have difficulty understanding how they are seen by others (Oltmanns, Gleason, Klonsky, & Turkheimer, 2005). Another explanation of inadequate self-reporting is the heightened need for social approval or deliberate distortion of information about oneself. Research shows that the Personality Inventory for DSM-5 (PID-5) is susceptible to intentional distortion like defensiveness (denying symptoms and psychological impairment) and social desirability (putting forth an exaggerated positive image) (Williams, Rogers, Sharf, & Ross, 2019), which may result in invalid self-reports. When a systematic measure of traits along with an adapted version of PID-5 was employed in therapist-client dyads

within an outpatient clinic, a median correlation of 0.41 across the PID-5 domains was discovered (Samuel, Suzuki, Bucher, & Griffin, 2018).

The juxtaposition of self-report and clinical assessment has so far been attempted in the assessment of types of PDs in several studies. In this area, we can observe the low correspondence between the clinical diagnosis and both standardised methods, and the somewhat higher correspondence between the standardised methods themselves (Tenney, Schotte, Denys, Megen, & Westenberg, 2003; Zimmerman & Coryell, 1990; Samuel et al. 2013). According to Klonsky, Oltmanns, & Turkheimer (2002), self-informant concordance appears to be higher for older subjects and for Cluster B traits (excluding narcissism), lower for Cluster A and Cluster C pathology, and lowest for traits related to narcissism. Agreement was higher for the few studies examining non-DSM domains of personality pathology ($r = 0.47$). In light of meta-perception, the self-reports tend to be most accurate for PD pathology focused on internalising symptoms of PD, whereas informant reports have more value for externalising pathology related to agreeableness and conscientiousness (Samuel, Suzuki & Griffin, 2015).

Samuel and colleagues (2013) studied the diagnoses of PDs with clinician reports, self-reports, and structured interviews. First, they studied agreement between treating clinicians' PD diagnoses and the semistructured diagnostic interview. For a categorical agreement, they found that *kappas* ranged from 0.21 (avoidant) to 0.42 (schizotypal) and for a dimensional agreement Pearson's correlations ranged from 0.30 (avoidant) to 0.44 (borderline). Second, agreement between clinicians' ratings and self-report questionnaires was lower than between clinicians' ratings and semistructured diagnostic interviews, with *kappas* ranging from 0.00 (OCPD) to 0.20 (borderline) and Pearson correlations ranging from 0.18 (schizotypal) to 0.28 (borderline). As a result of a meta-analytic review, Samuel (2015) also showed that agreement between clinicians' diagnoses and those from research methods (mainly self-report questionnaire) was modest (median dimensional agreement across 27 studies ranged from 0.05 to 0.36). He also found that clinicians' diagnoses agreed more with semi-structured interviews than self-report questionnaires, and convergence increased slightly when clinicians used more systematic diagnostic methods and when dimensional measurements of traits and PDs were used (Samuel, 2015).

Samuel, Suzuki & Griffin (2016) noticed that clinicians are biased to perhaps a similar degree as patients, so speaking about the limitations of self-report only should not be the case. Similarly to Bornstein (2017), they highlight the issue of the relative value of different sources in adult psychopathology. They indicate the need for the integration of different methods to refine the diagnosis and assessment process. Studying the diagnosis of Borderline PD, Hopwood and colleagues (2008) found that self-report yielded higher base rates of criteria endorsement. Moreover, they insist that their results did not support the common assumption that diagnostic

interviews are more valid than self-reports, but instead indicated that the combined use of these methods optimally identifies Borderline PD criteria. It is worth noting that some studies accentuate the convergence between self-report and clinicians' ratings. Individuals with pathological personality traits possess a reasonable degree of insight into their actual trait levels and associated impairment (Sleep, Lamkin, Lynam, Campbell, & Miller, 2019).

SWAP-200 AND SELF-REPORT IN PD ASSESSMENT

Westen & Weinberger (2004) emphasised that, in the diagnosis of personality disorders, the clinician relies on quite different information than the patient. The clinician sees the patient in the office or hospital ward and knows the patient's functioning from the stories the patient presents, from the biographical elements provided, and from his or her own in-depth experience situated in the anchorage of the transference and countertransference relationship. It is also oriented towards a holistic approach to assessment, situating them well to describe complex personality pathology. The patient, on the other hand, performs self-observation in much richer situations, but systematically may not have access to some of the content concerning his or her functioning due to various psychological mechanisms, mainly defensive. Most self-report assessments, thanks to the structure of the questionnaire, are easily subject to statistical aggregation, but most of the clinician's assessments are based on informal, non-systematic ways of assembling a different diagnostic information. However, there is an opportunity to make it more systematic in the form of the Shedler-Westen Assessment Procedure (SWAP-200; Westen, DeFife, Bradley & Hilsenroth, 2010; Shedler & Westen, 2007).

Some research on the convergence of self-report versus clinical assessment is conducted using the SWAP-200 expert clinical observation (the test is completed by the clinician, not the patient; Westen, DeFife, Bradley & Hilsenroth, 2010; Shedler & Westen, 2007). The SWAP-200 is used to assess personality and to formulate case studies, with use of a prototypical and dimensional approach to personality assessment. It is based on a psychoanalytic understanding of individuals' personalities, and includes ratings of unconscious processes, defences, interpersonal functioning, and affect regulation. The SWAP-200 contains 200 statements and is based on a Q-sort technique, i.e. the clinician assigns a score from 0 to 7 to each statement, depending on how well the statement describes the patient, while having a limited number of choices (e.g., a score of 0 can be assigned to 100 statements and a score of 7 can only be assigned 8 times). Many items in the SWAP-200 method are used to uncover subtle psychological processes, i.e. those that are not obvious and explicit (overt). Selecting statements that fit the patient uses clinical reasoning that goes beyond the facade information evident in the patient's behaviour and statements, and enables a deeper and clinically meaningful

understanding of the patient (Shedler & Westen, 2007). The SWAP-200 has been subjected to psychometric criticisms, including its reliance on a fixed skewed distribution, test-retest reliability, or validity across observers and situations (Wood et al., 2007). In response to these criticisms, Blagov and colleagues (2012) conducted a series of analyses and demonstrated that the tool is sufficiently reliable and accurate. Additionally, the introduction of normalised T scores helps to minimize a distorted impression of profiles of disorders with different base rates (such as Borderline vs. Schizotypal PD). The statistical aggregation of clinical scores in SWAP assists the professional in clinical judgment, but does not substitute it. Undoubtedly, it is better to assess patients systematically rather than simply intuitively. Therefore, methods such as the SWAP-200, with appropriate methodological awareness on the part of practitioners, are paramount in solving diagnostic problems.

Davidson, Obonsawin, Seils, and Patience (2003) assessed the agreement between a modified version of the SWAP (to make it suitable for use as a self-report questionnaire) and the original version of SWAP-200. A study was conducted on a small sample of outpatients and their treating clinicians. The results showed that PD prototype scales had very poor agreement. The limitations of this study primarily relate to the limited knowledge of the tailor-made self-report method, resulting in little clinical utility of the results. The median correlation was 0.28.

Bradley and colleagues (2007) studied self and informant convergences in Borderline PD, Antisocial PD and Obsessive-Compulsive PD. They discovered small to moderate correlations between personality as assessed by clinician reports using the SWAP-200 and personality assessed by self-report using the Personality Assessment Inventory (PAI; Morey, 2004). The SWAP-200 Borderline PD scale correlated moderately with the PAI, including 0.31 with borderline features, 0.40 with affective stability, and -0.07 with self-harm. Similarly, the SWAP-200 Antisocial PD scale correlated 0.35 with PAI antisocial features, 0.44 with antisocial behaviour, 0.21 with stimulation seeking, 0.45 with aggressive attitudes, and 0.46 with drug problems. They also found strong negative correlations between SWAP Obsessive-Compulsive PD scale scores and the PAI scores related to Borderline PD, Antisocial PD, aggression, and substance use and abuse. Median correlation was also calculated and was of 0.33. A limitation of this study is that a small number of disorder types were considered.

Exceeding these constraints, Gritti, Samuel, and Lang (2015) investigated the convergence between the self-reported Millon Clinical Multiaxial Inventory-III (MCMI-III; Millon, Davis, & Millon, 1997) and the PD scales from the SWAP-200, completed by treating therapists. To describe only the main results, the highest positive correlations (Pearson's r) they obtained were for Avoidant PD (0.45), Obsessive-Compulsive PD (0.43), Antisocial PD (0.41), and Borderline (0.4), Histrionic PD (0.36), and Schizoid PD (0.34). Paranoid and Schizotypal PDs were

not correlated and Narcissistic and Dependent PDs were not significantly correlated. The dimensional agreement across the 10 PDs ranged from -0.10 to 0.45, with a median of 0.35. As the authors point out, an important limitation of these studies was the temporal lag between the patients' completion of the MCMI-III and the clinicians' ratings on the SWAP-200. Furthermore, these studies mainly relied on categorical personality models, resulting in excessive comorbidity.

In sum, no prior study has examined the agreement between SWAP-200 ratings, assigned by a practicing clinician, and a self-report measure of the both the full complement of the DSM-5 PDs and borderline structural features (characteristic for borderline level of personality organisation), where both assessments were made at almost the same time. Although we do not exceed the limitation of using diagnostic categories, by introducing a level of personality organisation (analogous to AMPD Criterion A) and by treating personality disorders dimensionally (not as a dichotomous category, but a score indicating the quantitative score in a given type of disorder), we predominantly use a dimensional approach.

CURRENT STUDY

The general aim of the current study was to empirically address the convergence between clinical ratings (SWAP-200-PL) and self-report measures of character (personality) types (Character Styles Questionnaire – Revised, CSQ-R; Cierpiałkowska & Pasikowski, 2004) and borderline structural dimensions (The Borderline Personality Inventory, BPI, Leichsenring, 1999). Although similar questions have already been asked in other (still only a few) studies, this is the first attempt to show how the personality concept behind the SWAP-200-PL is related to other PD constructs used in a relatively large clinical sample with the tools mentioned above.

To examine the extent of convergence of clinical and self-report methods examining personality types, we chose the CSQ-R because it also addresses the main types of PDs introduced in the DSM-IV and maintained in the DSM-5. In this way, this tool is compatible with the SWAP-200. Moreover, the SWAP-200 and CSQ-R share not only the distinction between types of disorders, but also to co-occurring psychological mechanisms whose clinical meaning is derived from the psychodynamic approach. Similarly, the BPI, although it refers to borderline PD, contains subscales that allow it to obtain a picture of the severity of the pathology of the borderline personality structural dimensions (fear of fusion, primitive defences, identity diffusion, reality testing; see the Measurement section).

Firstly, we expected to identify low to moderate convergence when assessing individual PD types. Of these, those that are more externalising in nature (histrionic, borderline) and diagnosed largely on the basis of behavioural expressions will have higher diagnostic convergence between clinical ratings (SWAP-200-PL) and self-report measures (CSQ-R). Moreover, as an explora-

tory part of the study, we decided to identify the shared variance between SWAP-200 PDs and CSQ-R PDs treated as the two multivariate constructs. This would inform us of how the instruments as a whole measure similar constructs.

Secondly, we expected that the borderline self-report encompassing dimensions of personality structure (BPI) is correlated with the SWAP Borderline clinical assessment. At the same time, other PDs diagnosed with the use of clinical ratings (SWAP-200) placed on low or high levels of borderline personality organisation (Kernberg, 2004: schizoid, borderline, paranoid, narcissistic, antisocial and histrionic, passive-aggressive, dependent) will also be associated with the structural dimensions of the BPI. Based on Kernberg's concept (*ibidem*), obsessive-compulsive and depressive personality will not correlate positively with personality structure dimensions in the BPI. Additionally, the empirical research show that the obsessive-compulsive PD differs substantially from other PDs (e.g., DSM-5 cluster B), but also that it tends to be associated with relatively less impaired functioning compared with other disorders (see the conformation of a similar hypothesis in Bradley, Hilsenroth, Guarnaccia, Westen, 2007). Thus, we expected the obsessive-compulsive PD scores on the SWAP-200 to correlate negatively with the BPI measures.

METHOD

Measurement

Clinician-reported PDs. The Shedler-Westen Assessment Procedure-200 (SWAP-200, Shedler, Westen, 1999) was used in order to aggregate the information derived from an expert clinical assessment. The SWAP-200 is a Q-sort instrument with a fixed distribution for the assessment of personality pathology. It consists of 200 items dealing with both specific behaviours and more inferential processes. The clinician sorts items into eight categories, from 0 (not descriptive of the person) to 7 (most descriptive). This administration procedure corresponds to the prototypal thinking of the clinician, where he/she decides how well each statement matches the patient. Then, the SWAP scoring algorithms generate a dimensional score for each PD included in the DSM-5 (PD T-Scores were used here) that refers to the similarity between the patient and prototype (e.g., the 'ideal' form of paranoid PD). The SWAP also includes a dimensional measure of PD aspects (as e.g., narcissism, obsessiveness). An additional scale measuring psychological strengths and adaptive functioning (Psychological Health Index) that was used here as a divergent validity measure. In this study, the Polish version of the SWAP-200 was used. The Polish version was back-translated several times and finally accepted by the authors of the original version.

Self-reported types of PDs. Character Styles Questionnaire-Revised (CSQ-R) (Cierpiakowska & Pasikowski, 2004) was developed in 1996 on the basis of descriptions of functioning patterns of people with different types of PDs in the psychodynamic approach. In the present study, the latest revised version of the questionnaire was used. This version refers to the criteria

for PDs presented in DSM-IV-R and, consequently, also DSM-5. The questionnaire allows for the examination of ten types of PDs from bundles A, B and C, and its test items refer to four areas of fixed patterns of internal experiencing and behaviour, which to a significant extent differ from the expectations present in the culture in which the individual lives. These areas are: 1) patterns of cognitive functioning, which include the way of perceiving and interpreting oneself, other people and events in the surrounding reality; 2) patterns of affective responding, especially the strength, complexity, variability and adequacy of emotional reactions to stimuli; 3) patterns of social functioning, which refer to the way of establishing, maintaining and resolving conflict situations in relationships with loved ones; 4) patterns of aggressive and sexual impulse control. The psychometric study involved three stages. In the first, competent judges (therapists and clinicians) sorted the 280 test items into 11 groups, 10 classes (10 types of PDs and a class 'other'). In the second, tests were conducted on a group of 350 men and women, aged 21 to 64 years. In the third, on the basis of the discriminatory power of individual items, the final version of the method was established. The Character Styles Questionnaire-R consists of 130 items (with possible answers to choose from 1 - definitely not to 5 - definitely yes) and examines 10 types of PD in clusters A (paranoid, schizoid and schizotypal PD), B (narcissistic, borderline, histrionic, and antisocial PD), and C (dependent, avoidant/anxiety, and obsessive-compulsive personality) (Cierpiakowska & Pasikowski, 2004). The reliability of scales ranges from Cronbach's α Dependent PD - 0.682 to Avoidant PD - 0.854.

Self-reported borderline personality. The Borderline Personality Inventory (Leichsenring, 1999; see also: Cierpiakowska, 2001) is a highly reliable and valid self-reporting method, recommended for use in both borderline personality organisation screening and borderline PD. It consists of 53 true/false items on four scales: an identity diffusion (in our study Cronbach's α = 0.839), a primitive defence mechanisms (α = 0.766), an impaired reality testing (α = 0.829), and a fear of closeness (fusion) (α = 0.642). The BPI has good internal consistency, test-retest reliability, and satisfactory rates for sensitivity (0.85 to 0.89) and specificity (0.78 to 0.90) (Leichsenring, 1999). Higher scores on the scales indicate a higher level of structural personality aspects indicating borderline personality. A score of 20 points or more denotes borderline pathology. The most discriminately powerful items were selected to create the Cut-20 subscale (Cronbach's α = 0.849), where a score of 10 or more also provides an identification of borderline personality organisation.

Participants and procedure

A power analysis (using *g*power* software) was conducted before the study to determine the appropriate sample size. Assuming a first-order error rate of α = .05 and a power of $1-\beta$ = .90, with an expected explained variance (r^2) of 0.10, the resulting sample size from the

power analysis should be 101 individuals for the two-sided point biserial model. Participants in this study were both patients consecutively admitted for inpatient treatment at a voivodeship hospital in Greater Poland and patients who remained in treatment for at least several months at the Mental Health Outpatient Clinic. The invitation for participation in the study was addressed to 306 patients. The decision to complete the questionnaires depended on the informed consent of the study participant (both on self-report part and clinician-assessment part) and on clinical assessment of his/her mental state of the patients. The study took place after stabilisation of the psychological functioning of the patient.

The final sample consisted of 102 Caucasian patients, including 90 women (88.2%) and 12 men (11.8%). Their mean age was 38.41 years (SD = 12.78, range = 18-64). Inclusion criteria were broad, and both single and comorbid disorders were included (see more in Table 1). The exclusion criteria encompassed severe substance addiction disorders and disorders with the possibility of active psychotic states (e.g., schizophrenia, schizo-affective disorders). However, if there was a history of these disorders, the person was included in the study. Nosological diagnoses were assigned by applying the ICD-10 criteria to data gathered in the systematic clinical interview by a psychiatrist. Thus, the sample was heterogeneous in terms of the nosological diagnoses, but similar in terms of not including voluntary patients with the severe problems with reality testing at the moment of the study.

The SWAP-200 scores in our study were obtained by an assessment made by a qualified psychiatrist and psychologist who was trained in psychodynamic approach. Before using the SWAP-200 on patients, the clinician performed several trial assessments under the supervision of the authors of this article (LC, ES), who had adapted the SWAP-200 into Polish. During the data collection phase, the clinical ratings were prepared after an intake interview and after a minimum of 3 hours of individual contact with a patient (sometimes this was much longer), and a minimum of two weeks of observation on the ward during hospitalisation. As a result of the natural setting, there was not a rigidly predetermined interval between the moment when the patient filled in the self-report questionnaire and the rating of the SWAP-200 by the clinician; however, it was not longer than 1 month. SWAP-200 preparation time averaged about 45 minutes for each patient. It corresponds with requirements for assessment with this method (Shedler & Westen, 2007; Marmarosh, Bieri, Stuber, Gunnia, Nwigwe, Sevilla, & Rice 2010). For a demographic and diagnostic characteristics of this sample, please see Table 1.

RESULTS

Statistical analysis employed the program jamovi (2.2.2.0) and R (4.1.2). The descriptive statistics for the 12 PDs from the SWAP-200 (PD T-Scores and Psychological Health Index), 10 PDs from the CSQ-R, and dimensions of borderline personality structure (BPI) are provided in

Table 1. Demographic and nosological characteristics of the sample (N = 102)

Variable	N	%
Female	90	88.2
Male	12	11.8
Age: mean (SD), min-max	38.41 (12.78), 18-64	
Education:		
Primary	3	2.9
Vocational	20	19.6
Secondary	51	50
Higher	28	27.5
Marital status		
Single	46	45.4
In a relationship	45	44.6
Divorced	10	9.9
Psychotherapy	33	32.7
Number of hospitalisations: M (SD), min-max	2.13 (1.5), 1-6	
Nosological diagnosis		
<i>PDs</i>		
PD, unspecified (F60.9)	20	19.6
Borderline PD (F60.3)	10	9.8
Other specific PDs (F60.8)	4	3.9
Paranoid PD F60.0	2	2.0
Antisocial PD (F60.2)	2	2.0
Dependent PD (F60.7)	1	1.0
Schizoid PD (F60.1)	1	1.0
Obsessive-compulsive PD (F60.5)	1	1.0
<i>Other disorders</i>		
Depressive disorders (F32, F33, F34)	41	40.0
Reaction to severe stress, and adjustment disorders (F43)	28	27.5
Bipolar disorder (F31)	16	15.7
Mental and behavioural disorders due to psychoactive substance use (F10-F19)	13	12.7
Anxiety disorders (F40, F41)	10	9.8
Unspecified mood disorder (F39)	7	6.9
Obsessive-compulsive disorder (F42)	5	4.9
Eating disorders (F50)	1	1.0
Schizophrenia (F20)	1	1.0
Schizoaffective disorders (F25)	1	1.0
Other anxiety disorders (F41)	1	1.0

Table 2. As the study was conducted in a natural setting on a clinical group, the distributions of the variables were predominantly non-normal; however, the parametric tests were used because deviations do not significantly bias the results (Mardia's multivariate normality test for CSQ-R indicates that the assumptions are met and for SWAP-200, the distribution problems concern kurtosis in Histrionic, Narcissistic, Avoidant, and Depressive (Table 2). In assessing the interpretation of correlation coefficients, Cohen's (1988) guidelines were used.

Correspondence between self-report (CSQ-R) and clinician ratings (SWAP-200) of particular PD types. We

calculated an overall correlation matrix to identify the degree of dimensional agreement between scores on the SWAP-200 and CSQ-R PD in order to establish convergent validity. Table 3 shows the full matrix of the correlations between SWAP-200 PD scores and CSQ-R scores for all 10 corresponding PDs. We took the clinical assessment (SWAP-200) as a point of reference when discussing the results. The magnitude of the correlation between the self- and clinician-reported scores varied considerably across the PDs, ranging from a low and insignificant of 0.05 (Obsessive-Compulsive) to higher and significant 0.42 (Borderline). In addition to the low

Table 2. Descriptive statistics for study variables (N = 102)

	M	SD	Range	Skewness	Kurtosis	Shapiro-Wilk W	Shapiro-Wilk p
Shedler-Westen Assessment Procedure-200 (SWAP-200)							
Paranoid	48.17	7.54	30.30-67.70	0.09	-0.27	0.99	0.865
Schizoid	49.31	8.82	34.70-68.60	0.31	-1.00	0.96	0.002
Schizotypal	52.31	7.50	35.70-72.00	0.10	-0.31	0.99	0.927
Antisocial	50.18	9.71	35.30-74.40	0.33	-0.74	0.96	0.003
Borderline	51.16	9.70	26.20-69.60	-0.16	-0.70	0.98	0.125
Histrionic	53.74	10.59	28.90-71.40	-0.10	-1.10	0.95	0.002
Narcissistic	49.87	10.74	31.80-71.30	0.14	-1.10	0.96	0.003
Avoidant	46.82	9.84	29.30-65.50	0.02	-1.10	0.96	0.006
Dependent	49.74	8.57	28.60-66.10	-0.08	-0.66	0.98	0.281
Obsessive Compulsive	45.20	9.13	28.60-69.10	0.01	-0.70	0.97	0.048
Depressive	47.39	9.03	29.50-64.20	-0.01	-0.97	0.97	0.036
Passive-Aggressive	47.03	9.05	26.40-64.50	-0.09	-0.79	0.98	0.069
Psych.Health.Index	46.51	10.43	32.30-75.10	0.75	-0.36	0.93	<.001
Character Styles Questionnaire-R (CSQ-R)							
Paranoid	5.57	3.32	0-12	0.21	-0.82	0.96	0.006
Schizoid	5.93	3.13	0-12	-0.22	-0.93	0.96	0.002
Schizotypal	4.63	3.33	0-13	0.63	-0.39	0.94	<.001
Antisocial	4.19	2.66	0-13	0.76	0.54	0.94	<.001
Borderline	6.51	3.47	0-13	0.00	-0.78	0.97	0.024
Histrionic	6.44	2.78	0-13	-0.12	-0.35	0.98	0.197
Narcissistic	4.88	2.68	0-12	0.53	-0.10	0.96	0.005
Avoidant	7.54	3.56	0-13	-0.38	-0.83	0.95	<.001
Dependent	6.30	2.68	1-13	0.15	-0.62	0.97	0.042
Obsessive Compulsive	7.51	3.00	1-13	-0.18	-0.60	0.97	0.029
Borderline Personality Inventory (BPI)							
BPI	18.34	10.87	0-46	0.58	-0.30	0.96	0.004
BPI_identity	3.53	2.86	0-10	0.62	-0.69	0.91	<.001
BPI_defences	3.78	2.32	0-8	-0.03	-1.02	0.95	<.001
BPI_realitytesting	0.68	1.31	0-5	1.93	2.46	0.58	<.001
BPI_fearoffusion	2.96	2.00	0-8	0.56	-0.57	0.93	<.001
BPI-Cut-20	9.55	5.38	0-22	0.22	-0.73	0.98	0.058

value for Obsessive-Compulsive PD, the convergences for Paranoid, Narcissistic, and Dependent were also minimal and insignificant ($r = 0.05$), whereas Schizotypal and Avoidant were of 0.20 but still insignificant. Schizoid, Antisocial and Histrionic were around 0.3 and became statistically significant.

It should be noted, however, that for some disorders, the higher correlations were not between their counterparts (e.g., SWAP Paranoid and CSQ-R Paranoid) but other PD types. For example, Paranoid SWAP was significantly correlated with CSQ-R Antisocial (0.27); Schizotypal SWAP with CSQ-R Schizoid (0.30); Histrionic SWAP with CSQ-R Antisocial (0.32); Narcissistic SWAP with CSQ-R Antisocial (0.31). These correlation coefficients could be treated as discriminant values that indicate lower than the convergent value for those PDs. In sum, 7 of the 10 SWAP-200 scales (Paranoid, Schizotypal, Narcissistic, Histrionic, and Obsessive-compulsive) obtained at least one discriminant correlation that was as high as, or higher than, the convergent correlation. In this sense, the Schizoid, Histrionic, Antisocial and Borderline SWAP PDs seem to converge the most, while there is evidence of some lack of specificity in the associations between disorder types measured by self-report and clinician ratings.

What is also interesting, the SWAP Obsessive-compulsive PD was significantly negatively correlated with CSQ-R Antisocial and CSQ-R Histrionic PDs ($r = -0.33$). At the same time, on the contrary, SWAP Borderline PD was correlated positively and significantly with CSQ-R Antisocial PD. SWAP Health index was significantly negatively correlated with three CSQ-R PDs

(the strongest correlation -0.43 with CSQ-R Borderline PD); however, CSQ-R Dependent and Anancastic was not significantly correlated with Health Index. OCPD and BPD may seem to be opposites in terms of health indices.

Correspondence between self-report (CSQ-R) and clinician ratings (SWAP-200) treated as generalised PD constructs. To identify shared variance between the two multivariate constructs, a canonical (Sherry & Henson, 2005) correlation analysis was conducted for 12 variables from the SWAP scale as predictors of 10 variables from the PD scale (Table 4). The analysis contributed to the identification of two significant canonical explanatory functions with $R_c^2 = 45.12\%$ ($\lambda = 0.14$, $F(120, 636) = 1.55$, $p < 0.001$) and 36.04% ($\lambda = 0.25$, $F(99, 583) = 1.28$, $p = 0.044$), respectively. Since Wilks' λ represents the variance unexplained by the model, $1 - \lambda$ represents the effect size for the full model expressed as r^2 . For the model and all functions tested, this effect was 86.3%. Table 4 presents the standardised canonical coefficients for the first two dimensions across both sets of variables. The first canonical dimension was most strongly influenced by CSQ-R Borderline (-0.546), CSQ-R Schizotypal (-0.364), and SWAP Borderline (-0.469). Function 1 can be described as a dimension of 'severe personality disorder' because CSQ-R Anancastic and SWAP Avoidant, SWAP Dependent, and SWAP Obsessive Compulsive were inversely related to Function 1.

Function 2 was not as strong as the previous one, and was mostly influenced by CSQ-R Schizoid (0.32), SWAP Schizoid (0.46), SWAP Schizotypal (0.38), SWAP Avoidant (0.45), SWAP Dependent (0.33), SWAP Obsessive Compulsive (0.38), and SWAP Depressive (0.39). Look-

Table 3. Pearson's correlation coefficient matrix of SWAP-200 and CSQ-R PD Scales ($N = 102$).

	CSQ-R Paranoid	CSQ-R Schizoid	CSQ-R Schizotypal	CSQ-R Antisocial	CSQ-R Borderline	CSQ-R Histrionic	CSQ-R Narcissistic	CSQ-R Avoidant	CSQ-R Dependent	CSQ-R Anancastic
SWAP Paranoid	0.10	0.07	0.16	0.27*	0.19	0.05	0.05	-0.05	-0.04	0.05
SWAP Schizoid	0.17	0.30*	0.07	-0.19	0.11	-0.19	-0.04	0.25	0.10	-0.11
SWAP Schizotypal	0.31*	0.36**	0.21	-0.05	0.27*	-0.03	0.09	0.30	0.14	-0.08
SWAP Antisocial	0.06	-0.02	0.20	0.33*	0.23	0.22	0.14	-0.09	-0.05	-0.01
SWAP Borderline	0.18	0.07	0.25	0.33*	0.42***	0.23	0.20	0.15	0.09	-0.11
SWAP Histrionic	0.11	-0.02	0.16	0.32*	0.24	0.29*	0.18	0.05	0.06	-0.02
SWAP Narcissistic	0.02	-0.04	0.15	0.31*	0.15	0.20	0.12	-0.12	-0.04	0.03
SWAP Avoidant	0.09	0.23	-0.06	-0.23	0.022	-0.24	-0.11	0.22	0.10	-0.10
SWAP Dependent	0.07	0.20	-0.10	-0.19	-0.01	-0.15	-0.08	0.22	0.16	-0.11
SWAP Obsessive Compulsive	-0.09	0.03	-0.23	-0.33*	-0.26	-0.33*	-0.22	-0.00	-0.01	0.05
SWAP Depressive	0.08	0.19	-0.05	-0.21	0.06	-0.23	-0.1	0.21	0.1	-0.14
SWAP Passive-Aggressive	0.16	0.23	0.18	0.19	0.26	-0.03	0.07	0.08	0.081	-0.03
Psych. Health Index	-0.25	-0.19	-0.32*	-0.30*	-0.43***	-0.19	-0.23	-0.19	-0.06	0.13

Note: Grey cells refer to the matching PD type. Significance level with False Discovery Rate (FDR) correction: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 4. Canonical solution for personality types based on SWAP and CSQ-R for Functions 1 and 2

	Function 1			Function 2			
	Coeff.	r_c	r_c^2 (%)	Coeff.	r_c	r_c^2 (%)	h^2 (%)
CSQ-R Paranoid	-0.316	-0.212	4.49	0.219	0.131	1.72	6.21
CSQ-R Schizoid	-0.269	-0.181	3.28	0.534	0.321	10.30	13.58
CSQ-R Schizotypal	-0.542	-0.364	13.25	0.042	0.025	0.06	13.31
CSQ-R Antisocial	-0.389	-0.262	6.86	-0.288	-0.173	2.99	9.86
CSQ-R Borderline	-0.813	-0.546	29.81	0.090	0.054	0.29	30.10
CSQ-R Histrionic	-0.204	-0.137	1.88	-0.478	-0.287	8.24	10.11
CSQ-R Narcissistic	-0.293	-0.197	3.88	-0.079	-0.047	0.22	4.10
CSQ-R Avoidant	-0.283	-0.190	3.61	0.286	0.171	2.92	6.53
CSQ-R Dependent	-0.093	-0.062	0.38	0.204	0.122	1.49	1.87
CSQ-R Anancastic	0.411	0.276	7.62	0.098	0.059	0.35	7.97
R_c^2		0.672	45.12		0.600	36.04	
SWAP Paranoid	-0.293	-0.197	3.88	0.014	0.008	0.01	3.89
SWAP Schizoid	-0.110	-0.074	0.55	0.763	0.458	20.98	21.52
SWAP Schizotypal	-0.291	-0.196	3.84	0.629	0.378	14.29	18.13
SWAP Antisocial	-0.412	-0.277	7.67	-0.411	-0.247	6.10	13.77
SWAP Borderline	-0.699	-0.469	22.00	-0.270	-0.162	2.62	24.62
SWAP Histrionic	-0.364	-0.244	5.95	-0.512	-0.307	9.42	15.38
SWAP Narcissistic	-0.279	-0.188	3.53	-0.414	-0.248	6.15	9.68
SWAP Avoidant	0.003	0.002	0.00	0.742	0.446	19.89	19.89
SWAP Dependent	0.060	0.040	0.16	0.550	0.330	10.89	11.05
SWAP Obsessive Compulsive	0.434	0.291	8.47	0.624	0.375	14.06	22.53
SWAP Depressive	-0.086	-0.058	0.34	0.657	0.394	15.52	15.86
SWAP Passive-Aggressive	-0.414	-0.278	7.73	0.375	0.225	5.06	12.79

Note: *Coeff.* = standardised canonical function coefficient; *rs* = structure coefficient; rs^2 = squared structure coefficient; h^2 = communality coefficient; R_c^2 = squared canonical correlations

ing at the structure coefficients for the entire function, we see at the same time that CSQ-R Histrionic, CSQ-R, CSQ-R Narcissistic, CSQ-R Antisocial and SWAP Antisocial, SWAP Borderline, SWAP Histrionic, and SWAP Narcissistic were negatively related to the relevant synthetic variables. Given the nature of these variables, we labelled this function 'internalised anxiety'.

Correspondence between clinician ratings (SWAP-200) and self-reported indices of borderline personality structure (BPI). In order to test the hypothesis that self-report in terms of borderline indicators, including dimensions of personality structure (BPI), were correlated with clinical ratings of Borderline PD, we calculated the Pearson's *r* correlation coefficients (Table 5). SWAP Borderline was positively correlated with the general score on the BPI ($r = 0.31$) and BPI-Cut-20 ($r = 0.41$). It was also correlated with the primitive defences ($r = 0.33$), fear of closeness ($r = 0.28$), but no significant correlation was found with the reality testing dimension and identity diffusion. Moreover, the hypothesis that BPI would also

correlate positively with such PD types that were treated by Kernberg (2004) as organised on the low (severe) borderline level of personality pathology was only partially confirmed. The SWAP Antisocial and Histrionic PDs were positively correlated with general and BPI-Cut-20 score; however, only one structural dimension (primitive defences) was detected as positively correlated with the group of PDs located on the borderline level of personality organisation. The SWAP Schizoid, Schizotypal, and Paranoid PDs were not correlated with the BPI general score and BPI dimensions. As expected, the SWAP Obsessive-compulsive, Depressive (those PDs that were classified by Kernberg as operating under neurotic personality organization), and the BPI general score (and structural dimensions) were neither negative nor statistically significant. Specifically, while SWAP Depressive was not correlated at all, the SWAP Obsessive-compulsive displayed negative correlations in terms of BPI general score ($r = -0.32$), BPI-Cut-20 ($r = -0.34$), and structural dimensions such as primitive defences and fear of fusion

Table 5. Correlation matrix of SWAP-200 and CSQ with BPI general score and sub-scales ($N = 102$)

	BPI Total	BPI Identity	BPI Defences	BPI Reality testing	BPI Fear of closeness	BPI-CUT20
SWAP Paranoid	0.19	0.19	0.19	0.15	0.09	0.24*
SWAP Schizoid	0	0.07	0.01	0.04	0.05	-0.03
SWAP Schizotypal	0.19	0.18	0.18	0.17	0.19	0.17
SWAP Antisocial	0.27*	0.17	0.21	0.16	0.17	0.31**
SWAP Borderline	0.31**	0.18	0.33**	0.06	0.28**	0.41***
SWAP Histrionic	0.25*	0.12	0.26*	0.06	0.16	0.30**
SWAP Narcissistic	0.18	0.13	0.15	0.09	0.08	0.23*
SWAP Avoidant	-0.11	-0.02	-0.04	-0.08	-0.05	-0.12
SWAP Dependent	-0.11	-0.05	-0.02	-0.1	-0.08	-0.10
SWAP Obsessive Compulsive	-0.32**	-0.17	-0.22	-0.17	-0.25*	-0.34***
SWAP Depressive	-0.09	-0.02	-0.02	-0.09	-0.02	-0.10
SWAP Passive Aggressive	0.22*	0.24*	0.21	0.14	0.15	0.27*
SWAP Psych. Health Index	-0.35***	-0.24*	-0.35***	-0.16	-0.33**	-0.40***
CSQ-R Paranoid	0.73***	0.67***	0.73***	0.43***	0.64***	0.72***
CSQ-R Schizoid	0.59***	0.58***	0.58***	0.36***	0.49***	0.62***
CSQ-R Schizotypal	0.83***	0.78***	0.65***	0.54***	0.67***	0.76***
CSQ-R Antisocial	0.54***	0.43***	0.40***	0.26*	0.44***	0.47***
CSQ-R Borderline	0.72***	0.6***	0.73***	0.3**	0.68***	0.79***
CSQ-R Histrionic	0.55***	0.44***	0.45***	0.29**	0.39***	0.54***
CSQ-R Narcissistic	0.56***	0.46***	0.41***	0.30**	0.48***	0.53***
CSQ-R Avoidant	0.47***	0.38***	0.57***	0.16	0.56***	0.53***
CSQ-R Dependent	0.52***	0.43***	0.53***	0.31**	0.45***	0.51***
CSQ-R Anancastic	0.10	0.09	0.10	0.02	0.08	0.10

Note: Significance level with False Discovery Rate (FDR) correction: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

($r = -0.25$). SWAP Obsessive-compulsive appears as a PD that is not similar to other PDs when structural dimensions are considered.

For exploratory purposes, we conducted a correlation analysis between the dimensions of personality pathology structure in the BPI and the intensity of the personality types according to the CSQ. The analysis showed higher correlation coefficients between self-reported BPI and CSQ-R than between BPI and SWAP-200. Particularly noteworthy were the results regarding the correlation between BPI and CSQ-R. The Identity BPI correlated highly with disorders such as the CSQ-R Schizotypal (0.76), Paranoid (0.66), and Borderline (0.64), while it did not correlate with the CSQ-R Anancastic. Similar correlations were observed for the BPI Defences scale. For the BPI Reality testing subscale, the highest significant correlations were observed with the CSQ-R Schizotypal (0.50), Paranoid (0.4), and Schizoid (0.35), and again no relationship was found with the CSQ-R Anancastic. The BPI Fear of closeness scale correlated highly with the CSQ-R Borderline (0.67), Paranoid (0.62), and Schizotypal (0.61). The BPI total score, showing the level of

personality organization, correlated highly with the CSQ-R Schizotypal (0.81), Borderline (0.75), and Paranoid (0.73), while the CUT-20 scale, indicating a diagnosis of Borderline PD with this particular type of disorder (0.79). It is worth noting that the CSQ-R Anancastic, as the only one, did not correlate with any dimension of the BPI.

DISCUSSION

Considering the review of the results of different studies, concerning both general knowledge about discrepancies in assessment and specific issues for PDs, we could not have expected high correlations between self-report and clinical assessment in our study. However, we should have enriched the picture of diagnosing personality disorders using different methods, looking for the meaning of the direction of the observed correlations. It is also possible to contribute to the clarification of the clinical utility of the different tools (here SWAP-200, CSQ-R, BPI). Our study has shown how the relationships between self- and clinical assessment look and in the discussion we share our understanding of them.

Correspondence between self-report (CSQ-R) and clinician ratings (SWAP-200) of particular PD types

We identified low convergence when assessing particular PD types with a self-report and clinical assessment with SWAP-200. The highest convergences (although still at the medium and low level) were obtained in assessing three types of externalizing personality disorder: borderline, antisocial, and histrionic (cluster B, DSM). In contrast, the lowest convergences were observed for obsessive-compulsive, paranoid, dependent, and narcissistic personalities (from a different clusters). These results are consistent with the work of Klonsky, Oltmanns, & Turkheimer (2002), who also noted the lowest self-informant concordances for traits related to narcissism and the highest for other Cluster B traits. It may be noted that higher concordance scores occurred for externalizing personality disorders (see Kelley, Edens & Morey, 2016). Higher concordance in externalization may be due to the dominance of some pathological personality traits, such as disinhibition, and antagonism (cf. AMPD DSM-5, criterion B: pathological personality traits), which manifest even with brief social contact (Kotov et al., 2018). In addition, narcissistic individuals may be motivated to idealize self-presentation and positive overt self-esteem that does not go along with the clinician's assessment of narcissistic struggles (e.g., narcissistic 'scar'). The clinician is more likely to see self-esteem problems like instability, discrepancy, or contingency (e.g., Kuchynka & Bosson, 2018). The low convergence between self-report (CSQ-R) and clinician ratings (SWAP-200) in the case of obsessive-compulsive and paranoid PD is more challenging to interpret. The most plausible hypothesis seems to be that the discrepancy may result from an incongruity between the self-presentation during contact with the clinician and the patients' self-disclosure and, with it, the experience self-evident in the CSQ-R. Various clinically inspired explanations can be sought here. For example, people with obsessive-compulsive PD may be strongly ego-syntonic and value their symptoms positively and thus not report them or, by contrast, be very meticulous in listing them. When paranoid PD is concerned, patients may either sound convincing to the clinician or hide their possible difficulties in self-description.

In some cases, the study showed higher correlations between PDs' counterparts in both methods (types according to SWAP-200 and types according to CSQ-R), but between different types of personality disorders. The number of convergent correlations (between the same PDs in both the SWAP and CSQ-R methods) and discriminant correlations (between different PD types in each SWAP and CSQ-R method) indicated high similarity between the manifestations of borderline, antisocial, and schizoid personality assessed based on the SWAP. Although one rarely points to similarities between them, individuals with borderline and antisocial PD may be similarly experienced by the clinician during the assessment process, particularly by the severity of their irritability, aggression, and combativeness, as well as their lack of empathy (Bender, 2005). It is worth noting that these significant cross-PD

correlations mentioned here reflect symptom overlap among the DSM-5 PD diagnostic categories (e.g., Paranoid, Schizotypal, Borderline, Antisocial). The traditional PD (types) approach may at least partially explain the small effect sizes and the lack of discriminant validity. The limited utility of the PD type category and the fuzzy boundaries and high comorbidity between disorders are evident. With such difficulties, the retreat from categorical models of personality disorders in diagnosis is justified.

The overall configuration of correlations indicates the opposition of Borderline PD and Obsessive Compulsive PD in terms of Health Index. For example, negative correlations were noted between SWAP Obsessive-Compulsive PD and CSQ-R Histrionic and Antisocial PDs scores. In addition, SWAP Borderline and SWAP Obsessive-Compulsive have different patterns of correlations with the BPI scales. As predicted, SWAP Obsessive-Compulsive PD also correlated negatively with fear of closeness and personality organization (BPI Total), showing a similar correlation pattern to the SWAP Health Index. Assuming after Leichsenring (1999) that a high BPI Total score shows the level of organization of the borderline personality, it is worth noting that the Obsessive-Compulsive SWAP-200 correlates negatively with the BPI Total and the CSQ-R Anankastic does not correlate at all. Given the low representation of individuals with severe Obsessive-Compulsive PDs (see Limitations section), this result may not be entirely conclusive. However, it is worth further research looking at this disorder in the context of similarity to other personality disorders and properties of mental structure (e.g., the prevalence of displacement over splitting). Some studies already provide similar results. Bradley and colleagues (2007) found strong negative correlations between SWAP Obsessive-Compulsive PD scale scores and the PAI scores related to Borderline PD. The similarities and differences between Obsessive-Compulsive PD and Borderline PD were also studied. Although Obsessive-Compulsive PD may be characterized by notable difficulties in several emotional domains (e.g., negative affectivity) there is an evidence that the similarity with Borderline PD is rather low (Steenkamp, Suvak, Dickstein, Shea & Litz, 2015). As Aleknavičute and colleagues (2016) observed, Obsessive-Compulsive PDs and Borderline PDs may have a similar burden of subjective mood disturbance, but they substantially differ in physiological stress reactivity.

The canonical correlation analysis showed that both methods (clinician assessment with SWAP and self-report with CSQ-R) at an overall level do not strongly converge. Function 1 shows 45% convergence and Function 2 shows 36%. Furthermore, we can see that the highest concordance between the self-report method and the clinical assessment is in the description of borderline PD. Function 1 shows us that, in the clinical assessment when evaluating 'severe PDs', the borderline type is distinguished, and a slightly more differentiated self-description (with addition of the Schizotypal type) corresponds to it. Function 2, referring to 'internalised anxiety' in the clinicians' assessment, is more differentiated in terms of PD type,

while in the self-description mainly schizoid symptoms are reported. Perhaps the expertise of the recognition of the BPD has increased in the last decade (as Gunderson advised in 2009). At least the externalising pattern has been sufficiently unified between clinicians (their professional view) and patients (their ability to articulate symptoms). Of course, we must keep in mind that two specific diagnostic tools (SWAP & CSQ-R) are the basis of this suggestion, but their emergence is a sign of their times and reflects the current level of knowledge about PDs.

Correspondence between clinician ratings (SWAP-200) and self-reported indices of borderline personality structure (BPI).

We expected that the borderline self-report (encompassing dimensions of personality structure with BPI) is correlated with the SWAP Borderline clinical assessment. Results supported the hypothesis of associations between the BPI and PD types with lower levels of personality integration. SWAP Borderline, Antisocial, and Histrionic PD correlated positively with the total score and the BPI Cut-20 score, indicating a pathology of borderline personality organization (Kernberg, 2004). At the same time, it should be emphasized that positive correlations in these personality disorders were observed only with the primitive defence mechanisms scale. SWAP schizoid, schizotypal, and paranoid PDs were not significantly correlated with the overall and BPI dimensions. Perhaps these PDs are characterized by a lower than borderline pathology of personality organization, which is not examined by the BPI (Leischenring, 1999). Similarly, there was also a negative correlation between obsessive-compulsive PD scores on the SWAP-200 and BPI measures (for defence, the anxiety of fusion, and the total BPI and CUT-20).

The differences in the magnitudes of the correlations between self-reported BPI and CSQ-R, compared to between BPI and SWAP-200, favor higher scores between the self-report methods, suggesting a discussion on the role of informants' perspectives. It may be that a different source of information contributes more to the discrepancies between the results of different tools than the distinct theoretical underpinnings of the diagnostic tools. After all, in our study, we only had instruments derived from a psychodynamic perspective at our disposal, which referred to subjective experiences, conflicting desires, internal conflicts, and defense mechanisms.

LIMITATIONS

To our knowledge, this is the first application of the SWAP-200 for adults in a Polish inpatient sample. This highlights the uniqueness of the study but also underscores the necessity to view it as preliminary and investigate the issues of validity and reliability of the SWAP-200 measure in its Polish version, particularly as a diagnostic tool (J. Shedler and D. Westen are the owners of the Polish version for diagnostic purposes). The present study

represents one of the initial steps in establishing the validity of the SWAP-200. In our study, one clinician used the SWAP-200. He was an intended user of the SWAP-200, with a high level of expertise in psychiatry, psychology and psychodynamic approach and after intensive training with the method. This situation lacks more detailed information on coding agreement when compared with other raters and there is the potential risk of susceptibility of the rater to idiosyncrasies. The SWAP-200 results may also have been affected by the variable related to patient contact hours (as an indicator of familiarity), and it should be controlled in further studies. However, this article is crucial as it may trigger research using the SWAP-200 in Poland, and be the starting point for further work on this clinically useful tool.

It should be agreed that, in order to determine the true validity of therapist ratings relative to self-report methods, they must be collected via equivalent scales to remove measurement confounders and isolate source effects (Samuel, Suzuki & Griffin, 2015). In our study, we wanted to avoid a method parallel to the SWAP-200, which would have been created only for this study, but instead we reached for a tool that shares the basic assumptions, namely psychoanalytic and psychodynamic theoretical background and is used in a case-conceptualization for the purpose of treatment. It turned out, however, that the shared provenance was not enough to obtain greater correspondence between the self-report and the SWAP-200 than in previous studies.

Some limitations also apply to the research sample. The recruited participants were relatively high functioning, with a somewhat restricted range of personality pathology in terms of PD types. Less than half of patients met criteria for a ICD-10 diagnosis. The overall level of personality psychopathology with respect to nosological diagnoses appears to be low (with the exception of BPD). It is possible that the strongest results were observed for BPD, because there is sufficient variability in those features. Although involving a clinical sample, the results obtained in this study varied in terms of personality functioning. There are people in the sample with severe personality disorders and those who have not received this diagnosis; there is also the possible over-representation of the internalizing type of psychopathology (depression was often diagnosed on admission to hospital). The significant range restriction concerning other PD types may have contributed to the weak correlations; however, when the BPI as an indicator of personality organization level is considered, the possible ranges and the observed ranges overlap. It suggests that the sample is diversified in terms of the severity of personality pathology. The sample characteristics should also influence caution in formulating conclusions about Obsessive-Compulsive PD. Although the distribution of the Obsessive-Compulsive PD scores is not distorted, there are no individuals with this nosological diagnosis in our sample. Also, the results of the canonical correlation analysis may be biased due to the limited sample size and the partial non-normality of the distribution in the subscales of CSQ. In addition, the majority were

female, which is associated with a greater propensity to seek psychological help. Moreover, the sample bias may influence the results, given the association between the prevalence of various types of disorders and gender. Although the tools used here allow for the assessment of personality pathology by the severity of a type or prototype on a dimension, a sample with individuals with higher (moderate and severe) severity of personality pathology could better highlight diagnostic problems. Clearly, the next step in this research is to collect data on a larger, broader sample of patients, mainly with a PD diagnosis. It is still an open question as to which personality disorders (which types or which trait dimensions) account for the greatest discrepancies between self-report and clinical assessment.

CONCLUSION

In conclusion, in our study, we did not observe high convergence between self-report and clinician assessment. Instead, we observed areas of convergence, among which it is worth pointing out the clinical description of Borderline PD symptoms and the non-manifestation of borderline indicators in disorders from the so-called higher level of personality organization (Kernberg, 2004). The notion of a low level of correspondence between self-report and clinical assessment (here the SWAP-200-PL and CSQ-R) also leads to further reflection on PD assessment in general, especially the twilight of traditional PD approach. We highlight the need to establish rules to take into account different sources of information and the integration of different methods to refine PD assessment.

REFERENCES

- American Psychological Association (2013). *Diagnostic and statistical manual of mental disorders (DSM-5)*. American Psychiatric Publishing.
- Aleknaviciute, J., Tulen, J. H., Kamperman, A. M., de Rijke, Y. B., Kooiman, C. G., & Kushner, S. A. (2016). Borderline and cluster C personality disorders manifest distinct physiological responses to psychosocial stress. *Psychoneuroendocrinology*, *72*, 131-138. <https://doi.org/10.1016/j.psyneuen.2016.06.010>
- Bach, B., Kramer, U., Doering, S., di Giacomo, E., Hutsebaut, J., Kaera, A., De Panfilis, C., Schmahl, C., Swales, M., Taubner, S., & Renneberg, B. (2022). The ICD-11 classification of personality disorders: A European perspective on challenges and opportunities. *Borderline Personality Disorder and Emotion Dysregulation*, *9*(1), 12. <https://doi.org/10.1186/s40479-022-00182-0>
- Bender, D. S. (2005). The therapeutic Alliance in the treatment of personality disorders. *Journal of psychiatry practice*, *11*, 2, 73-87.
- Blagov, P. S., Bi, W., Shedler, J., & Westen, D. (2012). The Shedler-Westen Assessment Procedure (SWAP): Evaluating Psychometric Questions About Its Reliability, Validity, and Impact of Its Fixed Score Distribution. *Assessment*, *19*(3), 370-382. <https://doi.org/10.1177/1073191112436667>
- Bornstein, R. F. (2002). A process dissociation approach to objective-projective test score interrelationships. *Journal of personality assessment*, *78*, 47-68. https://doi.org/10.1207/S15327752JPA7801_04
- Bornstein, R. F. (2017). Evidence-based psychological assessment. *Journal of personality assessment*, *99*(4), 435-445. <https://doi.org/10.1080/00223891.2016.1236343>
- Bradley, R., Hilsenroth, M., Guarnaccia, C., & Westen, D. (2007). Relationship between clinician assessment and self-assessment of personality disorders using the SWAP-200 and PAI. *Psychological assessment*, *19*(2), 225-229. <https://doi.org/10.1037/1040-3590.19.2.225>
- Cierpiałkowska, L. (2001). *Adaptacja Kwestionariusza Osobowości Borderline F. Leichsenringa [Polish adaptation of the Borderline Personality Inventory F. Leichsenring's]*. Poznań: Adam Mickiewicz University, Institute of Psychology (unpublished materials).
- Cierpiałkowska, L., Pasikowski, T. (2004). *Kwestionariusz stylów charakteru. Wersja poprawiona [Character Styles Questionnaire-Revised (CSQ-R)]*. Poznań: Adam Mickiewicz University, Institute of Psychology (unpublished materials).
- Clarkin, J. F., Meehan, K. B., & Lenzenweger, M. F. (2015). Emerging approaches to the conceptualization and treatment of personality disorder. *Canadian Psychology/Psychologie Canadienne*, *56*(2), 155-167. <https://doi.org/10.1037/a0038744>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. L. Erlbaum Associates.
- Connolly, J. J., Kavanagh, E. J., & Viswesvaran, C. (2007). The convergent validity between self and observer ratings of personality: A meta-analytic review. *International journal of selection and assessment*, *15*(1), 110-117. <https://doi.org/10.1111/j.1468-2389.2007.00371.x>
- Davidson, K. M., Obonsawin, M. C., Seils, M., & Patience, L. (2003). Patient and clinician agreement on personality using the SWAP-200. *Journal of personality disorders*, *17*(3), 208-218. <https://doi.org/10.1521/pedi.17.3.208.22148>
- Gritti, E. S., Samuel, D. B., Lang, M. (2015). Diagnostic Agreement between Clinicians and Clients: The Convergent and Discriminant Validity of the SWAP- 200 and MCMI-III Personality Disorder scales. *Journal of personality disorders*, *29*, 231-248. https://doi.org/10.1521/pedi_2015_29_231
- Gunderson, J. G. (2009). Borderline personality disorder: ontogeny of a diagnosis. *American Journal of psychiatry*, *166*(5), 530-539. <https://doi.org/10.1176/foc.8.2.foc230>
- Hopwood, C. J., Morey, L. C., Edelen, M. O., Shea, M. T., Grilo, C. M., Sanislow, C. A., McGlashan, T. H., Daversa, M. T., Gunderson, J. G., Zanarini, M. C., Markowitz, J. C., & Skodol, A. E. (2008). A comparison of interview and self-report methods for the assessment of borderline personality disorder criteria. *Psychological assessment*, *20*(1), 81-85. <https://doi.org/10.1037/1040-3590.20.1.81>
- Hopwood, C. J. (2018). Interpersonal dynamics in personality and personality disorders. *European journal of personality*, *32*(5), 499-524. <https://doi.org/10.1002/per.2155>
- Hopwood, C. J., Zimmermann, J., Pincus, A. L., & Krueger, R. F. (2015). Connecting personality structure and dynamics: Towards a more evidence-based and clinically useful diagnostic scheme. *Journal of personality disorders*, *29*(4), 431-448. <https://doi.org/10.1521/pedi.2015.29.4.431>
- Kelley, S. E., Edens, J. F., & Morey, L. C. (2017). Convergence of self-reports and informant reports on the personality assessment screener. *Assessment*, *24*(8), 999-1007. <https://doi.org/10.1177/10731911166636>
- Kernberg, O. F. (2004). *Contemporary controversies in psychoanalytic theory, techniques, and their applications*. New Haven, CT: Yale University Press.
- Klonsky, E. D., Oltmanns, T. F., & Turkheimer, E. (2002). Informant-reports of personality disorder: Relation to self-reports and future research directions. *Clinical psychology: Science and practice*, *9*(3), 300-311. <https://doi.org/10.1093/clipso.9.3.300>
- Kotov, R., Krueger, R.F., Watson, D., Achenbach, T.M., Althoff, R.R. I in., (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A Dimensional Alternative to Traditional Nosologies. *Journal of abnormal psychology*, *126*, 4, 454-477. <https://doi.org/10.1037/abn0000258>
- Kuchynka S. L., Bosson J. K. (2018). The Psychodynamic Mask Model of Narcissism: Where Is It Now?. In: Hermann A., Brunell A., Foster J. (eds.), *Handbook of trait narcissism*. Springer, Cham.
- Leichsenring, F. (1999). Development and first results of the Borderline Personality Inventory: A self-report instrument for assessing borderline personality organization. *Journal of personality assessment*, *73*(1), 45-63. <https://doi.org/10.1207/S15327752JPA730104>

- Marmarosh, C. L., Bieri, K., Stuber, E., Gunnia, K., Nwigwe, M., Sevilla, C., & Rice, K. (2010). Using the SWAP-200 with novice psychodynamic clinicians: Is it reliable, useful, and non-biased?. *Psychoanalytic psychology, 27*(4), 442-459. <https://doi.org/10.1037/a0020846>
- Morey, L. C. (2004). *The Personality Assessment Inventory (PAI)*. Lawrence Erlbaum Associates Publishers.
- Olino, T. M., & Klein, D. N. (2015). Psychometric comparison of self- and informant-reports of personality. *Assessment, 22*(6), 655-664. <https://doi.org/10.1177/1073191114567>
- Oltmanns, T. F., Gleason, M. E., Klonsky, E. D., & Turkheimer, E. (2005). Meta-perception for pathological personality traits: Do we know when others think that we are difficult?. *Consciousness and cognition, 14*(4), 739-751. <https://doi.org/10.1016/j.concog.2005.07.001>
- Rettew, D. C., Lynch, A. D., Achenbach, T. M., Dumenci, L., & Ivanova, M. Y. (2009). Meta-analyses of agreement between diagnoses made from clinical evaluations and standardized diagnostic interviews. *International journal of methods in psychiatric research, 18*(3), 169-184. <https://doi.org/10.1002/mpr.289>
- Samuel, D. B., Sanislow, C. A., Hopwood, C. J., Shea, M. T., Skodol, A. E., Morey, L. C., Ansell, E. B., Markowitz, J. C., Zhanarini, M. C., & Grilo, C. M. (2013). Convergent and incremental predictive validity of clinician, self-report, and structured interview diagnoses for personality disorders over 5 years. *Journal of consulting and clinical psychology, 81*(4), 650-659. <https://doi.org/10.1037/a0032813>
- Samuel, D. B., Suzuki, T., Bucher, M. A., & Griffin, S. A. (2018). The agreement between clients' and their therapists' ratings of personality disorder traits. *Journal of consulting and clinical psychology, 86*(6), 546-555. <https://doi.org/10.1037/ccp0000304>
- 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>
- Shedler, J., & Westen, D. (2007). The Shedler-Westen assessment procedure (SWAP): making personality diagnosis clinically meaningful. *Journal of personality assessment, 89*(1), 41-55. <https://doi.org/10.1080/00223890701357092>
- Sherry, A., & Henson, R. K. (2005). Conducting and interpreting canonical correlation analysis in personality research: A user-friendly primer. *Journal of personality assessment, 84*(1), 37-48. https://doi.org/10.1207/s15327752jpa8401_09
- Simon, J., Lambrecht, B., & Bach, B. (2023). Cross-walking personality disorder types to ICD-11 trait domains: An overview of current findings. *Frontiers in Psychiatry, 14*, 1175425. <https://doi.org/10.3389/fpsy.2023.1175425>
- Sleep, C. E., Lamkin, J., Lynam, D. R., Campbell, W. K., & Miller, J. D. (2019). Personality disorder traits: Testing insight regarding presence of traits, impairment, and desire for change. *Personality disorders: Theory, research, and treatment, 10*(2), 123-131. <https://doi.org/10.1037/per0000305>
- Steenkamp, M. M., Suvak, M. K., Dickstein, B. D., Shea, M. T., & Litz, B. T. (2015). Emotional functioning in obsessive-compulsive personality disorder: Comparison to borderline personality disorder and healthy controls. *Journal of personality disorders, 29*(6), 794-808. https://doi.org/10.1521/pedi_2014_28_174
- Tenney, N. H., Schotte, C. K., Denys, D. A., Megen, H. J. V. V., & Westenberg, H. G. (2003). Assessment of DSM-IV personality disorders in obsessive-compulsive disorder: comparison of clinical diagnosis, self-report questionnaire, and semi-structured interview. *Journal of personality disorders, 17*(6), 550-561. <https://doi.org/10.1521/pedi.17.6.550.25352>
- Williams, M. M., Rogers, R., Sharf, A. J., & Ross, C. A. (2019). Faking good: An investigation of social desirability and defensiveness in an inpatient sample with personality disorder traits. *Journal of personality assessment, 101*(3), 253-263. <https://doi.org/10.1080/00223891.2018.1455691>
- Westen, D., DeFife, J. A., Bradley, B., & Hilsenroth, M. J. (2010). Prototype personality diagnosis in clinical practice: A viable alternative for DSM-5 and ICD-11. *Professional psychology: Research and practice, 41*(6), 482-487. <https://doi.org/10.1037/a0021555>
- Westen, D., & Weinberger, J. (2004). When clinical description becomes statistical prediction. *American psychologist, 59*(7), 595-613. doi: 10.1037/0003-066X.59.7.595
- Wood, J. M., Garb, H. N., Nezworski, M. T., & Koren, D. (2007). The Shedler-Westen Assessment Procedure-200 as a basis for modifying DSM personality disorder categories. *Journal of Abnormal Psychology, 116*(4), 823-836. <https://doi.org/10.1037/0021-843X.116.4.823>
- World Health Organization. (2022). *ICD-11 clinical descriptions and diagnostic requirements for mental and behavioural disorders*. Available: <https://icd.who.int/browse11/l-m/en>
- Vazire, S. (2010). Who knows what about a person? The self-other knowledge asymmetry (SOKA) model. *Journal of personality and social psychology, 98*(2), 281. <https://doi.org/10.1037/a0017908>
- Zimmerman, M., & Coryell, W. H. (1990). Diagnosing personality disorders in the community: A comparison of self-report and interview measures. *Archives of general psychiatry, 47*(6), 527-531. 10.1001/archpsyc.1990.01810180027005
- 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>