Psychiatr. Pol. ONLINE FIRST Nr 388: 1–13
Published ahead of print 28 November 2025
www.psychiatriapolska.pl
ISSN 0033-2674 (PRINT), ISSN 2391-5854 (ONLINE)
DOI: https://doi.org/10.12740/PP/OnlineFirst/207497

Polish adaptation of the Structured Interview of Personality Organization (STIPO-R-PL)

Emilia Soroko[1], Monika Jańczak[1], Dominika Górska[1], Lidia Cierpiałkowska[1], Dorota Frydecka[2], Dorota Szcześniak[2], Małgorzata Kaczmarek-Fojtar[3], Piotr Musiał[2]

[1] Faculty of Psychology and Cognitive Sciences, Adam Mickiewicz University, Poznań
[2] Department of Psychiatry, Wroclaw Medical University
[3] Lower Silesian Centre for Mental Health for Children and Adolescents in Lubin

Summary

Aim. The aim of the study was to develop the Polish version of the Structured Interview for Personality Organization (STIPO-R-PL), conduct a preliminary assessment of its psychometric properties (interrater reliability, internal consistency, and validity), and disseminate knowledge about the tool among professionals working with patients with personality disorders.

Method. Trained psychotherapists conducted STIPO-R-PL interviews with 104 patients aged 18–64 years (68.3% women). Additionally, data were collected using the Structured Clinical Interview for DSM-5 Personality Disorders (SCID-5-PD) and self-report questionnaires, including the Symptom Checklist (SCL-90-R) and the Borderline Personality Inventory (BPI).

Results. The findings indicate that STIPO-R-PL allows for consistent evaluations by independent, trained raters. The tool demonstrates acceptable reliability (with the exception of the mature defense mechanisms subscale). STIPO-R-PL scores correlate positively with both self-reported dimensions of personality pathology and the severity of psychopathological symptoms. Furthermore, patients meeting the diagnostic criteria for personality disorders (based on the clinical interview) obtain statistically significantly higher scores across all STIPO-R-PL dimensions compared to those who do not meet the criteria, with effect sizes in the medium range. These results support the validity of the interview.

Conclusions. We recommend STIPO-R-PL for further testing in clinical practice (psychotherapy and supervision) as well as for future research aimed at a more comprehensive evaluation of its psychometric properties and diagnostic utility.

Key words: personality disorders, structured clinical interview, validity

Introduction

In recent years, there has been a growing interest among researchers and clinicians in dimensional conceptualizations of personality pathology. The shift away from the categorical classification system stems from its lack of diagnostic reliability, primarily due to the overlap of criteria across different disorders, which leads to comorbid diagnoses. Empirical research has confirmed the criticism of the categorical model, resulting in a gradual transition toward a more dimensional approach to personality pathology [1]. Significant changes have been introduced in classification systems, including the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) and its Alternative Model for Personality Disorders (AMPD) [2], as well as the International Classification of Diseases, 11th Revision (ICD-11) [3]

The new dimensional understanding aligns with psychodynamic conceptualizations of personality pathology, particularly with the model of personality organization proposed by O. Kernberg [4]. This model is based on object relations theory and integrates findings from developmental psychology, neurobiology, attachment theory, and classical psychoanalysis [5]. Similar to DSM-5 AMPD and ICD-11, Kernberg's model includes (1) identity integration (sense of self, sense of others, and the ability to pursue long-term personal goals) and (2) object relations (internal relational models that organize interpersonal functioning) as the main determinants of personality pathology. However, in contrast to DSM-5 and ICD-11, the object relations model also incorporates (3) defense mechanisms, (4) aggression, and (5) moral values, which, together with identity and object relations, are considered structural criteria [6]. Their assessment (diagnosis) provides a dimensional profile of personality structure functioning and allows for the determination of the level of personality organization. The value of this model lies in its ability to integrate Kernberg's proposed diagnostic classification with prognosis and treatment planning [7, 8]. The clinical utility of diagnosing a patient's level of personality organization based on this model has been well documented [5, 9].

The Structured Interview of Personality Organization (STIPO-R)

A diagnostic method theoretically grounded in Kernberg's model is the Structured Interview of Personality Organization (STIPO), with its revised version, STIPO-R, which is a shortened and psychometrically improved version [10, 11]. Their original prototype was Structural Interview (SI) [12].

The STIPO-R is a 55-item, semi-structured interview designed to assess five domains of personality functioning: (1) level of identity integration, (2) quality of object relations, (3) defense mechanisms, (4) level of aggression, and (5) moral values. Additionally, the STIPO-R version includes a narcissism scale, which consists of selected items from other domains that are particularly diagnostic for this form of psychopathology (see Table 1).

Table 1. **Domains and subdomains of Personality Structure According to STIPO-R**

Domain	Subdomain
Identity 15 items	Capacity to Invest Sense of self – Coherence and Continuity Representation of Others
Object Relations 14 items	Interpersonal Relationships Intimate Relationships and Sexuality Investment in Others
Defense Mechanisms 10 items	Lower-level, Primitive Defenses Higher-level Defenses
Aggression 9 items	Self-directed Aggression Other-directed Aggression and Hostility
Moral Values 6 items	-

Diagnosing with STIPO-R requires two competencies: (1) conducting the interview according to a structured protocol and (2) a two-step evaluation of the subject's responses. During the interview, the interviewer asks mandatory questions for each subdomain (see Table 2) and, if necessary, additional follow-up questions for clarification. Responses are rated on a 0–2 scale, where 2 indicates significant impairment and 0 reflects no impairment. The interviewer then sums the scores, calculates the mean or percentage score, and conducts a second clinical assessment on a 1–5 scale (see Table 3). The arithmetic score is based on interview responses, whereas the five-point rating incorporates clinical judgment, allowing for the use of transference-countertransference processes and relational patterns in diagnosis. Including this element aligns with contemporary conceptualizations of personality disorders as interpersonal and with diagnostic recommendations emphasizing the importance of relational quality [13].

Based on these assessments, the interviewer can construct a personality organization profile for the respondent across five domains. The dimensional ratings allow for classification of the individual's results within a specific level of personality organization. In STIPO-R-PL, three levels of personality organization are distinguished: normal, neurotic, and borderline. For research purposes, an overall score can also be calculated by averaging the ratings across domains, providing an indicator of the severity of personality disorder (STIPO-R total score).

Respondents at a normal and neurotic level have a consolidated identity, do not exhibit the use of primitive defense mechanisms, and demonstrate intact reality testing. At the neurotic level, individuals may show a certain degree of superficiality in their sense of self and/or others and might use some of immature defense mechanisms. At the borderline level, as the severity of the disorder increases – from high to moderate to low – there is a greater identity diffusion, increased reliance on primitive defense

mechanisms (such as splitting), overt aggression, disturbed object relations, and weaker internal moral standards.

Table 2. **Sample Question (11) for the Subdomain "Sense of Self – Coherence and Continuity" within the Identity Domain, along with Arithmetic Score (0, 1, 2).**

11. *Tożsamość* *Poczucie ja – spójność i ciągłość* *Samoocena* *Narcyzm 3*	**Czy powiedział(a)by Pan(i), że Pana(-i) samoocena zmienia się tak, że czasami widzi się Pan(i) jako jednostkę szczególną lub wyjątkową, a kiedy indziej jak osobę małą, nudną i wybrakowaną?** Jeśli tak: Czy mógłby/mogłaby Pan(i) powiedzieć, że zmiany w Pana (-i) samoocenie są dość poważne, że zdarzają się często lub że Pana(-nią) martwią? Czy może Pan(i) podać przykład tego, w jaki sposób przydarzają się Panu(-i) te zmiany w poczuciu siebie?	
	0 =	*Stałe, pozytywne i realistyczne poczucie samooceny.*
	1 =	*Nieco niestałe poczucie samooceny; samoocena podatna na wydarzenia zewnętrze, np. w reakcji na czynniki stresowe samoocena może na dłuższy czas odwrócić się z pozytywnej na negatywną.*
	2 =	*Poczucie samooceny znacznie lub poważnie niestałe, nieustannie negatywne lub wyraźnie wielkościowe czy nierealistycznie pozytywne.*

Table 3. **Overall clinical rating of domain with descriptions (Example for the Identity Domain)**

TOŻSAMOŚĆ – OGÓLNY wynik_____

1. **Tożsamość skonsolidowana** – badany(-a) zarówno ma poczucie swojego ja i inych dobrze zintegrowane, jak i angażuje się głęboko w pracę i rekreację.
2. **Tożsamość skonsolidowana**, lecz w niektórych obszarach występują drobne deficyty – poczucie ja i innych* w większości jest dobrze zintegrowane, lecz z łagodną powierzchownością, niestałością lub zniekształceniem I/LUB dostrzegalna jest pewna trudność z zaangażowaniem w pracę lub rekreację.
3. **Łagodna patologia tożsamości** – poczucie ja i/lub innych jest nieco źle zintegrowane (wyraźna powierzchowność lub niespójność i niestałość, niekiedy sprzeczności i zniekształcenia)* z wyraźnym zaburzeniem zdolności do zaangażowania w pracę/szkołę i/lub rekreację, ewentualnie badany(-a) angażuje się w większości dla zaspokojenia narcystycznych potrzeb.
4. **Umiarkowana patologia tożsamości** – poczucie ja i inych jest źle zintegrowane (Znaczna powierzchowność, niespójność, wyraźna niestałość, sprzeczności i zniekształcenia)* przy niewielkiej zdolności do zaangażowania w pracę/szkołę lub rekreację.
5. Głęboka patologia tożsamości – poczucie ja i innych jest niezintegrowane (ogromnie powierzchowne, niespójne, chaotyczne, rażąco sprzeczne i niezmiernie zniekształcone)* przy braku znaczącego zaangażowania w pracę lub rekreację.

* Uwaga: Jeśli występuje wyraźna rozbieżność między niestałością lub powierzchownością w poczuciu ja a w poczuciu innych, większa niestałość czy powierzchowność w poczuciu innych w odróżnieniu do poczucia siebie wskazuje na patologię narcystyczną.

Study aims

This study aimed to develop the Polish version of the STIPO-R interview and evaluate its psychometric properties (reliability and validity). An additional goal was to promote the use of STIPO-R among researchers and practitioners (diagnosticians, psychotherapists, and supervisors) who apply Kernberg's theory. We hypothesized that STIPO-R-PL would demonstrate psychometric properties comparable to the original version and other adaptations[5, 14–19]. We also expected that STIPO-R scores would positively correlate with self-report measures of structural characteristics of personality pathology and the severity of psychopathological symptoms. Additionally, individuals with a categorical diagnosis of personality disorders were anticipated to obtain higher scores across STIPO-R dimensions.

Materials and methods

Development of the Polish translation

The adaptation process of STIPO-R-PL followed WHO standards [20] and included translation by two independent experts, a detailed review by the authors, back-translation, and comparison with the original to eliminate discrepancies. Cultural adaptation and stylistic refinements were also carried out to align the protocol with the Polish context, including the introduction of formal address forms. The Polish version is freely available online [www].

Training of interviewers

The STIPO-R-PL interviews were conducted by 12 clinicians, psychodynamic psychotherapists or those familiar with Kernberg's theory, who had previously completed specialized training. The training process included analyzing recorded interviews conducted by experienced researchers, scoring assessments, receiving detailed feedback, and final qualification through a case analysis based on a video recording. Each interviewer conducted between 1 and 43 interviews, with an average of 12 per clinician.

Study procedure

Participants with suspected or diagnosed personality disorders were recruited by psychodynamic psychotherapists. After a detailed explanation of the study's purpose, they voluntarily provided informed consent for participation, interview recording, and the use of their data for research purposes. The procedure included self-report questionnaires, a clinical interview, and an in-depth STIPO-R interview, typically conducted over two sessions (range: 1–3). The study received approval from the Bioethics Committee of the Wroclaw Medical University (Nr 344/2018).

Participants in the validation study

The study was conducted in Poland between 2018 and 2023. Inclusion criteria encompassed patients receiving treatment for mental disorders (except for those meeting exclusion criteria) in either inpatient psychiatric wards or outpatient settings, with no intellectual disability. Exclusion criteria included: 1) cognitive impairments related to various mental disorders, 2) a diagnosis of schizophrenia or other psychotic disorders, 3) a current diagnosis of mental and behavioral disorders due to psychoactive substance use (excluding caffeine and nicotine), 4) a diagnosis of other mental disorders resulting from brain damage or dysfunction, and somatic diseases, 5) significant visual or hearing impairments preventing questionnaire completion and interview participation. Patients currently undergoing therapy with the recruiting clinician were not invited to participate in an interview conducted by that clinician.

The sample consisted of 104 patients (aged 18–65; $M = 31.8$, $SD = 9.95$; 68.3% women), most of whom had completed secondary ($n = 50$) or higher education ($n = 44$), with a smaller number having vocational or primary education ($n = 8$). The majority were employed ($n = 54$), while 13 were unemployed, 16 were students, and 13 were receiving social benefits. Slightly more than half were in a romantic relationship ($n = 54$), 41 were single, and 10 described their relationship status as unspecified. Forty-two individuals had been hospitalized, while 19 had received treatment in a day-care unit. A significant portion of the sample was receiving pharmacotherapy ($n = 68$), with 33 individuals taking antidepressants and 29 using other medications or a combination of other medications and antidepressants. Thirty-three participants were engaged in psychotherapy.

Personality disorder diagnoses assessed with the SCID-5-PD interview revealed that 29 individuals did not meet the criteria for any personality disorder. The most common diagnoses were borderline personality disorder ($n = 42$) and avoidant personality disorder ($n = 30$), followed by obsessive-compulsive ($n = 21$), dependent ($n = 17$), paranoid ($n = 14$), narcissistic ($n = 9$), antisocial ($n = 7$), schizoid ($n = 6$), histrionic ($n = 5$), and schizotypal personality disorder ($n = 1$). This overview indicates that, excluding the 29 individuals without a personality disorder, the average participant met the criteria for two types of personality disorders.

Measurement of the validation variables

Borderline Personality Inventory (BPI; [21–23]) is a self-report measure based on Kernberg's theory, recommended for both the screening assessment of borderline personality organization and the diagnosis of borderline personality disorder. It consists of 53 true/false items divided into four scales: identity diffusion, primitive defense mechanisms, impaired reality testing, and fear of fusion (closeness). The BPI demonstrates good internal consistency, high test-retest reliability, and satisfactory sensitivity (0.85–0.89) and specificity (0.78–0.90) [21], and it is also gender-neutral [23]. Higher scores on the scales indicate a greater presence of structural personality features characteristic of borderline organization. A score of 20 or higher suggests the presence of

borderline pathology understood as a borderline personality organization. The Cut-20 subscale was created by selecting the most diagnostically relevant items for borderline personality type, with a score of 10 or more indicating this type of personality disorder.

Symptom Checklist-90-Revised (SCL-90-R; [24, 25]) consists of 90 items and is used to assess the presence of problems and symptoms of psychopathology. Its nine subscales cover somatization, obsessive-compulsive symptoms, interpersonal sensibility, depression, anxiety, anger-hostility, phobic anxiety, paranoid ideation, and psychoticism. The overall score, Global Severity Index (GSI), reflects the severity of psychological distress. Each item is rated on a five-point scale, based on how much an individual was bothered by each item in the past four weeks (from 0 – "not at all" to 4 – "extremely"). The questionnaire demonstrates good reliability, and its convergent and discriminant validity has been confirmed in numerous studies [25].

The Structured Clinical Interview for DSM-5 Personality Disorders (SCID-5-PD; [26]) is used to assess ten categorical personality disorders (from clusters A, B, and C) included in Section II of DSM-5. Each disorder is evaluated based on seven to nine criteria, rated on a three-point scale (0 = criterion not met; 1 = criterion met subclinically; 2 = criterion met). SCID-5-PD is widely used in clinician-administered interviews and demonstrates good psychometric properties.

Results

Inter-rater reliability of the STIPO-R-PL interview

A large group of psychotherapists (n = 23) participated in the coding reliability assessment. After completing training, they evaluated a video-recorded case. To determine inter-rater reliability, the intraclass correlation coefficient (ICC) was calculated, following the approach used in previous studies on STIPO-R [18]. For all subdomains treated as quantitative variables (clinical 1 to 5 scale), an ICC value of 0.74 was obtained. Increasing the level of generalization to the five main domains (Identity, Object Relations, Defense Mechanisms, Aggression, and Moral Values) raised the reliability coefficient to 0.84. For the mean values derived from grouping subdimensions assessed on a 0–2 scale according to the STIPO-R subscales, the overall ICC reached 0.99. These results indicate a high level of rating agreement.

Descriptive statistics and reliability

Descriptive statistics for the key dimensions of STIPO-R are presented in Table 4. To evaluate internal consistency, a reliability analysis (Cronbach's α) was conducted for the six subscales, based on arithmetic scores ranging from 0 to 2. All subscales demonstrated adequate to high internal consistency, with the exception of the Higher-Level Defense Mechanisms subscale, where the lower reliability score may be due to the small number of items in this category.

Table 4. **Descriptive statistics for key STIPO-R dimensions based on clinical ratings and their reliability analysis (Cronbach's α)**

Key STIPO-R dimensions	M	Md	SD	Min	Max	Skew	Kurt	W	p	α
STIPO-R Total Score	2.51	2.50	0.737	1.00	4.17	0.315	-0.585	0.969	0.015	-
Identity	2.69	3	0.882	1	5	0.217	-0.200	0.887	<0.001	0.757
Object Relations	2.48	2	0.914	1	4	0.068	-0.775	0.879	<0.001	0.855
Lower-level Defenses	2.62	3	0.978	1	5	0.145	-0.560	0.901	<0.001	0.791
Higher-level Defenses	2.93	3	0.906	1	5	0.214	-0.476	0.887	<0.001	0.469
Agression	2.30	2	1.09	1	5	0.488	-0.691	0.876	<0.001	0.712
Moral Values	2.07	2	0.927	1	4	0.534	-0.540	0.850	<0.001	0.809
Narcissism	2.55	3	0.835	1	4	0.002	-0.534	0.870	<0.001	0.698

Validity

To assess convergent validity, Pearson's correlation coefficients (*r*) were calculated between STIPO-R scores and self-reported borderline structure dimensions measured by the BPI. As expected, most observed correlations were positive and statistically significant in our sample (Table 5). Excluding the total score and narcissism scale, the highest correlations were found between identity diffusion (BPI) and aggression (STIPO-R-PL), as well as between primitive defenses (BPI) and the corresponding STIPO-R-PL scale. The lowest correlations were observed between reality testing (BPI) and identity and object relations (STIPO-R-PL). The median correlation across dimensions was 0.342. Regarding narcissism scale in STIPO-R-PL, low but significant positive correlations were found with fear of fusion, primitive defenses, and the overall BPI score.

Table 5. **Correlations between STIPO-R-PL dimensions and self-reported Borderline Personality Inventory (BPI)**

STIPO-R	BPI Identity Diffusion	BPI Fear of Fusion	BPI Primitive Defenses	BPI Reality Testing	BPI Cut-20	BPI Total Score
STIPO-R Total Score	0.467***	0.485***	0.472***	0.367***	0.553***	0.580***
Identity	0.351***	0.332**	0.428***	0.194	0.419***	0.450***
Object Relations	0.204*	0.372***	0.385***	0.164	0.371***	0.361***
Lower-level DM	0.420***	0.389***	0.447***	0.284**	0.471***	0.516***
Higher-level DM	0.336***	0.384***	0.405***	0.353***	0.476***	0.468***
Aggression	0.477***	0.412***	0.271**	0.278**	0.485***	0.484***

table continued on the next page

Moral Values	0.284**	0.317**	0.314**	0.331**	0.365***	0.428***
Narcissism	0.187	0.321**	0.305**	0.122	0.338***	0.350***

*p <0.05; **p <0.01; ***p <0.001. DM – defense mechanisms.

Additionally, the relationships between the severity of personality pathology (STIPO-R-PL) and the severity of psychopathological symptoms were examined (Table 6). Overall, the results indicate that greater personality pathology (i.e., lower personality organization) is associated with increased severity of psychopathological symptoms, with correlations ranging from $r = 0.249$ for somatization to $r = 0.457$ for psychoticism. The highest median correlation was observed for the lower-level defense mechanisms dimension ($r = 0.309$), while the moral values dimension of STIPO-R-PL showed significant correlations only with hostility and psychoticism (SCL). Additionally, significant but small correlations were found between narcissism and paranoid ideation, depression, interpersonal sensibility, and obsessive-compulsive symptoms.

Table 6. **Correlations between STIPO-R-PL dimensions and self-reported Symptom Checklist (SCL-90-R)**

STIPO-R	SCL Somatization	SCL Obsessive-compulsive	SCL Interpersonal Sensibility	SCL Depression	SCL Anxiety	SCL Anger-Hostility	SCL Phobic Anxiety	SCL Paranoid Ideation	SCL Psychoticism
STIPO-R Total Score	0.249*	0.296**	0.318**	0.312**	0.326**	0.357***	0.349***	0.328**	0.457***
Identity	0.236*	0.221*	0.276**	0.256*	0.282**	0.307**	0.310**	0.292**	0.352***
Object Relations	0.177	0.299**	0.363***	0.325**	0.297**	0.232*	0.260*	0.252*	0.360***
Lower-level DM	0.174	0.274**	0.215*	0.172	0.221*	0.247*	0.278**	0.266**	0.315**
Higher-level DM	0.271**	0.300**	0.308**	0.335***	0.326**	0.309**	0.366***	0.261*	0.358***
Aggression	0.270**	0.218*	0.242*	0.285**	0.284**	0.315**	0.248*	0.254*	0.387***
Moral Values	0.015	0.061	0.079	0.075	0.098	0.239*	0.160	0.195	0.342***
Narcissism	0.120	0.231*	0.269**	0.294**	0.169	0.180	0.139	0.294**	0.297**

*p <0.05; **p <0.01; ***p <0.001. DM – defense mechanisms.

To examine whether the presence or absence of a categorical personality disorder diagnosis differentiates scores across STIPO-R-PL domains (dimensions), Student's t-tests and effect size estimation (Cohen's d) were conducted (Table 7). The results indicate that total score and individual STIPO-R-PL dimensions are significantly more

disturbed in individuals meeting the criteria for a personality disorder. The effect sizes range from small to moderate. The most strongly differentiated dimensions, apart from the overall score (d = 0.767), were identity (d = 0.652), aggression (d = 0.650), and object relations (d = 0.648).

Table 7. **Between-group comparisons of patients with and without SCID-5-PD personality disorders (n = 104)**

STIPO-R	Personality disorder n = 29		No personality disorder n = 75		Group comparisons		
	M	***SD***	***M***	***SD***	***t***	***df***	***d***
STIPO-R Total Score	2.13	0.670	2.66	0.712	3.55***	53.9	0.767
Identity	2.53	0.835	2.31	0.712	3.15**	64.1	0.652
Object Relations	2.07	0.884	2.64	0.880	2.96**	50.8	0.648
Lower-level DM	2.21	0.978	2.77	0.938	2.68**	49.1	0.591
Higher-level DM	2.66	0.897	3.04	0.892	1.96*	50.7	0.430
Aggression	1.83	0.889	2.48	1.107	3.12**	63.1	0.650
Moral Values	1.72	0.797	2.20	0.944	2.59**	60.0	0.545
Narcissism	2.28	0.841	2.65	0.814	2.07*	49.5	0.456

*p <0.05; **p <0.01; ***p <0.001. DM – defense mechanisms.

Discussion

In assessing coding consistency, evaluations made by psychotherapists after completing STIPO-R-PL administration and scoring training were considered. In this study, the ICC coefficient for clinical scales indicates that clinical ratings were satisfactory to good, while arithmetic scores demonstrated very good agreement. It is important to note that these values were maintained for generalized subscales, which enhances inter-rater agreement. This approach is justified, as determining personality organization level relies on generalized rather than partial observations. This contrasts with findings from studies on other versions of the instrument [16, 18, 19], where the ICC coefficient was consistently high for both scoring systems, ranging from 0.88 to 1.0 across different dimensions. The lower agreement rate observed in our study may be attributed to varying levels of familiarity with this method among psychotherapists.

For most subscales, the obtained reliability results are acceptable and slightly lower than those reported in similar studies [18]. An exception is the Higher-level Defense Mechanisms scale, which demonstrated low internal consistency, likely due to the small number of items (4). However, this inconsistency disappears when higher – and lower-level defense mechanisms are combined into a single scale. At the same time, the distinct correlation patterns of higher – and lower-level defense mechanisms

with psychopathological symptoms suggest that maintaining a separate subscale for higher-level defenses enhances the instrument's value.

To assess the construct validity of STIPO-R, we examined its correlations with a well-established self-report measure of borderline features based on Kernberg's model (BPI). As expected, we found low to moderate correlations between STIPO-R and all BPI dimensions, aligning with the shared theoretical foundation of these instruments and previous research findings [18, 19, 20]. These results further support the validity of STIPO-R in assessing personality organization and its structural dimensions.

Our study demonstrated that a lower level of personality organization is associated with a greater severity of various psychopathological symptoms, including depressive and anxiety symptoms, somatization, psychoticism, and paranoid ideation. This finding aligns with theoretical assumptions that structural personality pathology contributes to the development and persistence of psychological difficulties, which manifest as symptoms of different mental disorders. Moreover, a lower level of personality organization is linked to an increased likelihood of co-occurring psychotic symptoms, as individuals with severe personality pathology may experience episodes of psychotic decompensation under significant stress. In this study, correlations ranged from low to moderate, suggesting that, in addition to personality pathology, other significant factors also contribute to the severity of psychopathological symptoms.

The study also found that STIPO-R scores were higher among individuals who received a personality disorder diagnosis based on the SCID-5-PD interview compared to those who did not meet the diagnostic criteria. The effect sizes ranged from small to moderate, with the most strongly differentiated dimensions – identity, aggression, and object relations – representing core constructs in Kernberg's model and contemporary dimensional models to personality [7, 13].

Conclusions

The findings of this study indicate that STIPO-R-PL is a reliable and valid instrument, allowing for consistent assessments by independent raters. Given the growing interest in personality disorder diagnosis and the importance of a strong theoretical foundation for clinical interviews in its practical applications, these results support the recommendation of STIPO-R-PL as both a research and diagnostic tool. At the same time, we encourage further studies on its psychometric properties, including internal and external validity as well as clinical utility, to better define the place of the Polish version of STIPO-R among diagnostic and research instruments. Using STIPO-R-PL may also contribute to research on the effectiveness of psychodynamic psychotherapies, such as transference-focused psychotherapy.

It is important to emphasize that the diagnosis of personality disorders requires going beyond self-report data (questionnaires); it is based on integrating knowledge from various theoretical frameworks and on a deep understanding of the patient's functioning at both the explicit and implicit levels, as well as establishing an emotional connection with the patient [28]. Furthermore, it should be remembered that the outcome of the diagnostic process may change over time and can be modified based

on new data. Even the most advanced diagnostic tools merely provide support in this process and cannot replace the reflectiveness, experience, or insight of the clinician conducting the assessment.

Acknowledgements. *The authors express their gratitude to the interviewers and participants for their time and commitment, as well as to the Polish Society for Psychodynamic Psychotherapy for its support of the research.*

References

1. Hopwood CJ, Kotov R, Krueger RF, Watson D, Widiger TA, Althoff RR et al.*The time has come for dimensional personality disorder diagnosis.* Personal. Ment. Health 2018; 12(1): 82–86.
2. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5* 2013.
3. World Health Organization (WHO). *ICD-11: International classification of diseases (11th revision)* 2022.
4. Kernberg OF. *Borderline personality disorder and borderline personality organization: psychopathology and psychotherapy.* In*:* Magnavita JJ, editor. *Handbook of Personality Disorders. Theory and Practice.* New Jersey 2004; 92–119.
5. Biberdzic M, Clarkin JF. *Establishing Levels of Personality Functioning Using the Structured Interview of Personality Organization (STIPO-R). A Latent Profile Analysis.* J. Pers. Assess. 2024; 727-739.
6. Caligor E, Clarkin JF. *An object relations model of personality and personality pathology.* In: Clarkin JF, Fonagy P, Gabbard GO, editors. *Psychodynamic psychotherapy for personality disorders: A clinical handbook.* American Psychiatric Publishing 2010; 3–35.
7. Caligor E, Kernberg OF, Clarkin JF, Yeomans FE. *Psychoterapia psychodynamiczna patologii osobowości. Leczenie self i funkcjonowania interpersonalnego.* Polskie Towarzystwo Psychoterapii Psychodynamicznej 2019.
8. Bach B, Simonsen S. *How does level of personality functioning inform clinical management and treatment? Implications for ICD-11 classification of personality disorder severity.* Curr. Opin. Psychiatry 2021;. 34(1): 54–63.
9. Hörz-Sagstetter S, Volkert J, Rentrop, Benecke C, Gremaud-Heitz D, Unterrainer H-F et al. *A Bifactor Model of Personality Organization.* J. Pers. Assess. 2021; 103(2): 149–160.
10. Clarkin JF, Caligor E, Stern B, Kernberg OF. *Manual for the structured interview of personality organization-revised (STIPO-R).* Weill Cornell Medical College 2019.
11. Clarkin JF, Caligor E, Stern B, Kernberg OF. *Structured interview of personality organization (STIPO).* Weill Medical College of Cornell University 2004.
12. Kernberg OF. *Structural interviewing.* Psychiatric Cxlinics 1981; 4(1): 169-195.
13. Hopwood CJ, Wright AG, Ansell EB, Pincus AL. *The interpersonal core of personality pathology*. J. Pers. Disord. 2013; 27(3): 270-295.
14. Wang YLZ, Zhong J. *Psychometrics of the Chinese Version of the Structured Interview of Personality Organization (STIPO-CH).* BMC Psychiatry 2023; 568.
15. Di Pierro R, Gargiulo I, Poggi A, Madeddu F, Preti E. *The Level of Personality Functioning Scale Applied to Clinical Material From the Structured Interview of Personality Organization (STIPO):Utility in Detecting Personality Pathology.* J. Pers. Disord. 2020; 1–15.

16. Hörz-Sagstetter S, Caligor E, Preti E, Stern BL, De Panfilis C, Clarkin JF. *Clinician-Guided Assessment of Personality Using the Structural Interview and the Structured Interview of Personality Organization (STIPO).* J. Pers. Assess. 2018; 100(1): 30–42.
17. Nazari A, Huprich SK, Hemmati A, Rezaei F. *The Construct Validity of the ICD-11 Severity of Personality Dysfunction Under Scrutiny of Object-Relations Theory*. Frontiers in Psychiatry 2021; 12
18. Doering S, Burgmer M, Heuft G, Menke D, Bäumer B, Lübking M et al. *Reliability and validity of the German version of the Structured Interview of Personality Organization (STIPO).* BMC Psychiatry 2013; 13(1): 210.
19. Stern BL, Caligor E, Clarkin JF, Critchfield KL, Horz S, MacCornack V et al. *Structured Interview of Personality Organization (STIPO): Preliminary Psychometrics in a Clinical Sample.* J. Pers. Assess. 2010; 92(1): 35–44.
20. WHO guidelines on translation and adaptation of instruments-process of translation and adaptation of instruments. World Health Organization 2018; P4.
21. Leichsenring F. *Development and First Results of the Borderline Personality Inventory: A Self-Report Instrument for Assessing Borderline Personality Organization.* J. Pers. Assess. 1999; 73(1): 45–63.
22. Cierpiałkowska L. *Kwestionariusz zaburzenia osobowości borderline.* Uniwersytet im. Adama Mickeiwcza 2001. Unpublished materials.
23. Soroko E, Kleka P, Cierpiałkowska L, Leichsenring F. *Internal Structure, Reliability, and Gender Neutrality of the Borderline Personality Inventory* 2025. [In review].
24. Derogatis LR, Cleary PA. *Confirmation of the dimensional structure of the SCL-90-R: a study in construct validation.* J. Clin. Psychol. 1977; 33: 981–989.
25. Kuncewicz D, Dragan M, Hardt J. *Validation of the polish version of the symptom checklist-27-plus questionnaire.* Psychiatr. Pol. 2014; 48(2): 345-358.
26. Grande TL, Newmeyer MD, Underwood LA, Williams III CR. *Path analysis of the SCL-90-R: Exploring use in outpatient assessment.* Meas. Eval. Couns. Dev. 2014; 47(4): 271-290.
27. First MB, Williams JBW, Benjamin LS, Spitzer RL. *User's guide for the structured clinical interview for DSM-5® personality disorders (SCID-5-PD).* American Psychiatric Association Publishing 2016.
28. Hopwood CJ, Bornstein RF. *Conclusion: Toward a framework for integrating multimethod clinical assessing data.* In: Hopwood CJ, Bornstein RF, editors. Multimethod Clinical Assessment. New York: Guilford Publications 2014; 427–443.