A German version of the Shedler-Westen Assessment Procedure (SWAP-200) for the dimensional assessment of personality disorders

Die deutsche Version der Shedler-Westen Assessment Procedure (SWAP-200) zur dimensionalen Erfassung von Persönlichkeitsmerkmalen

Abstract

Theoretical background: The traditional diagnosis of personality disorders has been criticized for lack of empirical support, reliance on categorical classifications and low validity.

Objective: A German version of the Shedler-Westen Assessment Procedure (SWAP-200) is presented. This Q-Sort procedure is well-established in the USA and facilitates the dimensional diagnosis of personality pathology in accordance with DSM-IV. In addition, a taxonomy of personality has been derived using factor analyses.

Methods: The SWAP-200 was applied to 18 patients on the basis of semi-structured interviews. Interviewer ratings were correlated with the SWAP-200 completed by an independent observer and the patients' therapists.

Results: Good inter-rater reliability (between r=.69 and r=.76) was found. While average convergent validity coefficients between interviewer, observer and therapist were satisfactory (between r=.54 and r=.68), outliers were observed. Results suggest that the factor analytically derived taxonomy may be more valid.

Keywords: SWAP, personality disorders, dimensional personality diagnosis, Q-Sort

Zusammenfassung

Theoretischer Hintergrund: Die herkömmliche Diagnostik von Persönlichkeitsstörungen wird wegen ihrer fehlenden empirischen Absicherung, ihrer kategorialen Einteilung und geringer Validität kritisiert.

Zielsetzung: Vorgestellt wird die deutsche Version der Shedler-Westen Assessment Procedure (SWAP-200). Dieses im amerikanischen Sprachraum bewährte Q-Sort-Verfahren dient der dimensionalen Persönlichkeitsdiagnostik in Anlehnung an DSM-IV. Zusätzlich existiert eine Taxonomie der Persönlichkeit auf der Grundlage faktorenanalytischer Untersuchungen.

Methodik: In der vorliegenden Studie wurde mit 18 Patienten nach einem halbstrukturierten Interview die SWAP-200 durchgeführt. Die Ergebnisse des Interviewers wurden mit der SWAP-200 eines unabhängigen Beobachters und des behandelnden Therapeuten korreliert.

Ergebnisse: Es ergab sich eine gute Interrater-Übereinstimmung zwischen r=,69 und r=,76. Die Übereinstimmung zwischen Interviewer bzw. Beobachter und Bezugstherapeut (konvergente Validität) war mit r=,54 bis r=,68 im Mittel zufriedenstellend, es gab jedoch auch Ausreißerwerte. Es zeigten sich Hinweise darauf, dass die faktorenanalytisch gewonnene Taxonomie valider ist.

Schlüsselwörter: SWAP, Persönlichkeitsstörungen, dimensionale Persönlichkeitsdiagnostik, Q-Sort

Anke Höflich[1]
Marcus Rasting[2]
Jens Mach[2]
Silke Pless[2]
Simon Danckworth[1]
Christian Reimer[2]
Manfred E. Beutel[1]

1 Clinic and Policlinic for Psychosomatic Medicine and Psychotherapy, Johannes-Gutenberg-University-Clinic, Mainz, Germany

2 Clinic for Psychosomatics and Psychotherapy, Justus-Liebig-University-Clinic, Giessen, Germany

Introduction

Personality disorders are predominantly diagnosed according to the established classification systems ICD-10 and DSM-IV [7]. For diagnostic purposes, various self and observer-rating instruments are available. Examples of observer-rating instruments include the International Diagnostic Checklists for Personality Disorders according to ICD-10 and DSM-IV [3], the Structured Clinical Interview for DSM-IV Axis II [40] and the Personality Disorder Examination [19]. Examples of self-rating instruments include the Inventory of Personality Organization [9] and the Personality Style and Disorder Inventory [17].

The fundamental approach to diagnosing personality disorders is currently subject to much debate [33], [34], [35], [31], [14]:

1. It is criticized that empirical results played only a minor role in the development of the ICD and DSM diagnostic systems [41], [1], [27], [16].
2. The question as to whether personality disorders represent extreme variations of essentially normal personality traits or rather distinct entities remains open [39]. Defining the cut-off point at which the criteria for a personality disorder are met is therefore difficult [37], [39], [28], [32], [8]. It is possible for a patient to meet various criteria of various personality disorders without reaching the minimal number required for one diagnosis [30], [37]. In such a case, this information is discarded by the conventional diagnostic systems. In this context, Widiger [37] presented a striking example, in which 162 different combinations of borderline symptoms were identified without a single diagnosis of personality disorder being made.
3. It is also unclear which of the two diagnostic systems should be adopted. Although ICD-10 and DSM-IV differ hardly in their taxonomies, Sara et al. [26] found only a 29% agreement of personality disorder diagnoses between the two and Widiger et al. [38] found an even lower level of agreement of 7%. Differences are also to be found in prevalence estimates using the two diagnostic systems [25], [26].
4. It is diagnostically questionable whether patients are able to answer direct questions pertaining to personality pathology, given that personality disorders are considered to be ego-syntonic [28]. According to Westen [36], the majority of 1900 surveyed therapists preferred to draw upon the descriptions given by patients and the observations of patients' behavior during interviews or the therapy as a source of information in diagnosing personality disorders. The therapists did not consider direct questions regarding symptom patterns or self-rating instruments to be particularly helpful.
5. While the introduction of structured interviews, checklists and self-rating questionnaires has considerably improved the inter-rater reliability of personality diagnoses [6], temporal stability remains unsatisfactory [42], with test-retest reliabilities ranging from low to moderate [15], [11], [42], [12].
6. The validity of personality diagnoses (operationalized for example as agreement between clinical diagnoses and diagnoses on the basis of interviews) is rather moderate [6], [5], [21]. Decisions as to whether a personality disorder is present at all prove more reliable than those concerning which specific personality disorder(s) the patient has [4], [5], [29], [24], .[23], [21].
7. High comorbidity rates are to be found among personality disorders. At least one further personality disorder is diagnosed in between 50% and 100% of affected patients [20], [22], [32], [6], [18].

The question as to whether personality disorders should be classified using a categorical or a dimensional approach has been subject to long-standing discussion. Widiger [37] lists the ease of conceptualization and communication, as well as the familiarization with categorical diagnoses found in clinical practice as advantages of a categorical system. Advantages of a dimensional approach include a low loss of information and increased flexibility. Furthermore, many of the above-described problems associated with a categorical system can be avoided using a dimensional approach. On these grounds, a growing number of authors have begun to argue for a dimensional model of personality disorders [14], [37], [33], [20], [5], [7], [6], [29]. A consensus concerning the exact nature of the dimensional approach to be applied has, however, thus far not been reached.

With the development of the Shedler-Westen Assessment Procedure (SWAP-200), Westen and Shedler [34], [35] have provided a promising instrument for the dimensional evaluation of personality disorders. The SWAP-200 aims to delineate the personality structure of an individual patient in the form of a personality profile. The SWAP-200 is a Q-sort procedure [2]. The objective of such an approach is to avoid the various types of potential rater bias, for example the tendency to make ratings which are too high or too low based on a general impression. The items of the SWAP-200 were gradually collected and improved by Westen and Shedler [34], [35] over a period of seven years before being compiled into the final version. Items were drawn from a diverse range of sources: a) from DSM-III and DSM-IV criteria, b) from clinical and empirical literature on personality disorders, c) from suggestions made by clinically active physicians and therapists who had applied preliminary versions of the SWAP, d) from publications on 'normal' personality traits and e) from the clinical experiences of the authors. A more detailed account of the development of the SWAP-200 can be found in Westen and Shedler [33] and Shedler and Westen [28].

The SWAP-200 has already found wide-spread and successful application in the USA and shall now be presented for the first time in its German version. Reliability and validity of the German version of the SWAP-200 will be evaluated.

Methods

Participants

18 patients consecutively admitted to the in-patient crisis intervention unit at Giessen University Clinic for Psychosomatics and Psychotherapy voluntarily took part in the study.

Interviews were carried out and evaluated solely on the basis of patients' written consent. Inclusion criteria were as follows: (1) willingness to participate in the entire interview, (2) in-patient treatment of at least four weeks (extendable up to six weeks). Indications for crisis interventions included: stress reactions following acute life events, crises in the case of personality disorders and neuroses, crises within current therapies, social maladjustment and lack of social support, as well as preparation for long-term treatment. Exclusion criteria were: acute psychoses, acute suicidal tendency, addictive disorders and medical illnesses requiring intensive medical monitoring and supervision. Patients' mean age was 36 years (SD=8.6). The sample consisted of 11 women (61.1%) and seven men (38.9%). Six patients (33.3%) were married and five patients (27.8%) were single. Seven patients (38.9%) were separated, divorced or widowed. One third (N=6) each had an intermediate education or high school, five patients (27.8%) had a general school leaving certificate and one patient a school leaving certificate from a school for the disabled.

Depressive disorders (F32, 33, 34.1) constituted the most prevalent primary diagnoses (N=9, 50%) according to ICD-10, followed by adjustment disorders (F43, N=6, 33.3%) and anxiety disorders (F40, 41, N=3, 16.6%). A comorbid diagnosis of narcissistic personality disorder was made twice and that of emotionally unstable personality disorder once. The average illness duration was 23.9 months (SD=27.1).

Design

The experimental design was based on the procedure adopted by Shedler and Westen [28]. A semi-structured clinical interview [34], [35] was carried out with the patients in the clinic by a qualified psychologist. This 1½ to 2 hour interview was video-recorded. According to the procedure by Shedler and Westen covering various areas of the patient's life: reason for present stay in the clinic, important relationships in the past and the present, the patient's education and employment history, previous therapy experiences, reaction to stressful situations, moods, emotions, attitudes and ways of thinking. In responding to each of these topics, the patient was asked to provide specific examples. The patient's interaction with the interviewer was also observed. For the purpose of assessing inter-rater agreement, the video-tape was evaluated by a trained independent rater (doctoral student of medicine). Training of both the interviewer and

observer was carried out with the aid of two test interviews by a psychoanalyst of many years experience. Neither the interviewer nor evaluator of the video-tape possessed additional patient information.

Validity was tested using the SWAP-200 ratings provided by the respective therapist (three therapists with long-standing clinical experience) of each of the 18 patients. Therapists' ratings were based on their clinical impressions upon conclusion of in-patient treatment. In addition to three weekly sessions of group therapy, contact therapists had also conducted twice weekly individual therapy sessions. Patients were further subject to regular discussion and supervision within the entire therapeutic team (body therapists, art and music therapists, nursing staff). The 200 SWAP-200 statement cards were subsequently correlated with the personality disorder prototypes and Q-factor-profiles developed by Westen and Shedler, in order to determine PD and Q-factor scores (see below) for the respective patient. Three personality profiles (interviewer, observer and therapist) were thus compiled for each patient on the basis of PD and Q-factor scores. Approximately one hour was required for performance and analysis of the SWAP-200.

Shedler-Westen Assessment Procedure (SWAP-200)

The SWAP-200 comprises 200 statements that describe personality characteristics (e.g. "relationships tend to be unstable, chaotic, and rapidly changing", "tends to avoid social situations because of fear of embarrassment or humiliation"). The 200 items are classified into eight categories according to the Q-sort method. The first category (Category 0) contains all statements which do not apply to the patient, which are irrelevant or for which no information is available. This category is the largest and comprises 100 cards. The next category contains statements which apply somewhat to the patient, the following category statements which apply a little more etc. The eighth and final category contains those statements which directly apply to the patient. The cards are distributed across the categories represented in Table 1.

The description of the individual should pertain to stable traits and characteristics. On account of this, the previous two years are to be considered when making the ratings.

SWAP-200 prototypes

Westen and Shedler [34], [35] had 530 therapists rate a current patient with a personality disorder and further 267 therapists carried out the SWAP-200 for a fictitious prototypical patient with a specified personality disorder. Using these prototypes, the authors computed scores for each of the 200 SWAP cards for each of the 10 DSM-IV personality disorders (paranoid, schizoid, schizotypal, antisocial, borderline, histrionic, narcissistic, avoidant, dependent and obsessive-compulsive). Analogous to the "Global Assessment of Functioning" found in DSM-IV, scores for a "high level of functioning" were also com-

Table 1: Distribution of cards across categories

Category	0	1	2	3	4	5	6	7
Number of cards	100	22	18	16	14	12	10	8

puted. By correlating scores with patients' SWAP ratings, 10 personality disorder scores (PD scores) are yielded for each patient (more detailed information regarding the evaluation and corresponding files on analysis can be requested from beutel@psychosomatik.klinik.uni-mainz.de). These PD scores can, in turn, be plotted to form a personality profile.

Q-factors of the SWAP-200

In addition to the DSM-IV-based PD scores, Westen and Shedler [33] also developed a new taxonomy of personality disorders. To this end, 496 randomly selected clinically active psychologists and psychiatrists were asked to conduct the SWAP-200 for a current patient with a DSM-IV Axis-II diagnosis. A factor analysis was performed on the resulting data, with the aim of identifying clusters of patients with common personality characteristics and simultaneously distinguishing between patients with differing traits. The factor analysis yielded seven orthogonal clinically and theoretically coherent factors: dysphoric, schizoid, antisocial, obsessive, paranoid, histrionic and narcissistic. The dysphoric factor was further sub-divided into avoidant, high-functioning, emotionally dysregulated, dependent-masochistic and hostile-externalizing. A "high level of functioning" factor is also to be found here. As was the case for PD scores, one point is given for each of the SWAP-200 items per Q-factor score. By correlating the SWAP-200 items with these points, a patient's personality structure can be plotted as a Q-factor profile.

Questionnaires

Upon being admitted, patients responded to questions regarding sociodemographic characteristics and previous treatment according to Psy-BaDo [13]. Therapists completed medical documentation, including ICD-10 diagnoses, upon patient admittance and discharge.

Data analysis

Statistical analyses were carried out using SPSS (11.0 for Windows). Pearson correlation coefficients were calculated. The level of significance (5%, 1% or 0.1%) was stated for each significant result.
For the purpose of computing PD and Q-factor scores, correlation coefficients were, in accordance with Diehl and Staufenbiel [10], transformed into t-values using Fisher's z'-transformation.

Results

I. Inter-rater reliability

Table 2 and Table 3 present correlations of PD and Q-factor scores between interviewer, observer and therapist for all patients.
The mean correlation between interviewer and observer PD scores was r=.76 (N=18; SD=.12; min: .50; max: .92, see Table 2). As can be seen in Table 3, the mean correlation between interviewer and observer Q-factor scores was r=.69 (N=18; SD=.20; min: .25; max: .92). Good inter-rater agreement can thus be assumed for SWAP-200 ratings.

II. Agreement between interviewer/observer and therapist

In order to establish convergent validity, PD and Q-factor scores of the interviewer/observer were correlated with those of the patient's therapist. The mean correlation between therapist and interviewer PD scores was r=.54 (N=18; SD=.51; min: -.72; max: .91). A mean correlation of .59 was found between therapist and observer ratings (N=18; SD=.43; min: -.30; max: .98).
Correlations between interviewer/observer and therapist Q-factor scores were slightly higher than correlation results for PD scores. The mean correlation between therapist and interviewer Q-factor scores was r=.60 (N=18; SD=.33; min: -.28; max: .94), and r=.68 between therapist and observer (N=18; SD=.26; min: .07; max: .93). Agreement between interviewer/observer and therapist thus ranges from satisfactory to good. The extremely high differences among correlation coefficients are particularly striking. On account of the fact that it is less susceptible to outliers than the mean correlation, the median is also presented. As expected, the median was consistently higher than the mean and amounted to r=.81 for interviewer and therapist PD scores, and r=.79 for observer and therapist PD scores. The median correlation between interviewer and therapist Q-factor scores was r=.70 and r=.77 between observer and therapist (see Table 2 and Table 3).
Given that a number of outliers were observed, a case example (Patient 4) shall be presented in which negative correlations were found between interviewer/observer and therapist. Potential problems when carrying out the interview should become clear from this example. It is of

Table 2: Correlations of interviewer, observer and therapist PD scores for all patients

Patient	Correlation interviewer - observer	Correlation interviewer - therapist	Correlation observer - therapist
1	.69*	.87**	.56
2	.81**	.88***	.79**
3	.87***	.71*	.93***
4	.68*	-.72*	-.22
5	.84**	.89***	.72*
6	.84**	.28	.38
7	.92***	-.11	-.22
8	.81**	.88***	.79**
9	.86**	.64*	.89***
10	.91***	.86**	.80**
11	.54	.90***	.82**
12	.74**	.72*	.95***
13	.50	-.23	.23
14	.61*	.84**	.79**
15	.70*	.91***	.87***
16	.88***	.86**	.98***
17	.70*	.77**	.78**
18	.73*	-.28	-.30
mean (SD)	**.76 (.12)**	**.54 (.51)**	**.59 (.43)**
median	**.77**	**.81**	**.79**

* $p<.05$; ** $p<.01$; *** $p<.001$

great importance, not only here but also in other interview-based approaches, to pay particular attention to the relationship formed during the interview and to critically scrutinize countertransference arising in the context of the interview.

Figure 1 and Figure 2 depict profiles of the PD and Q-factor scores of the 28 year old female patient.

The patient suffered from post-traumatic stress disorder. During the interview, she focused on highly stressful details of a rape which had occurred some time ago. In contact with the female interviewer, she predominantly presented herself as a needy and helpless victim, as a consequence of which the topic of persistently maladaptive relationship patterns was inadequately broached.

During in-patient treatment, the patient appeared needy at the same time as being dismissive towards both staff and fellow patients. Despite her vulnerability, she asserted herself in an uncompromising way dominant and became involved in power struggles. She displayed strong mood swings. These patterns of behavior can be inter-preted as the patient's attempts to compensate for feelings of helplessness. Such behavior repeatedly resulted in conflict with both clinic staff and fellow patients. This in turn was reflected in the therapist's ratings, which pointed to a DSM-IV cluster B personality disorder (narcissistic, antisocial, borderline, histrionic).

In addition to outlier values, it is also apparent that differing results in inter-rater reliability and validity were found according to whether PD or Q-factor scores were employed.

Discussion and future prospects

In the present study, a Q-sort procedure for the dimensional assessment of personality disorders which is widely employed in the USA was presented as a German version for the very first time. A detailed and semi-structured clinical interview forms the basis of the assessment procedure. The SWAP-200 was developed in order to address the unresolved problems associated with traditional per-

Table 3: Correlations of interviewer, observer and therapist Q-factor scores for all patients

Patient	Correlation interviewer - observer	Correlation interviewer - therapist	Correlation observer - therapist
1	.68*	.83**	.68*
2	.70**	.92***	.68*
3	.83***	.70**	.92***
4	.72**	-.28	.07
5	.92***	.91***	.80**
6	.86***	.22	.20
7	.87***	.25	.18
8	.70**	.92***	.68*
9	.65*	.33	.81**
10	.74**	.57*	.61*
11	.38	.77**	.81**
12	.52	.36*	.86***
13	.25	.41	.74**
14	.33	.62*	.68*
15	.84***	.94***	.91***
16	.88***	.85***	.93***
17	.78**	.70***	.81**
18	.85***	.80**	.90***
mean (SD)	**.69 (.20)**	**.60 (.33)**	**.68 (.26)**
median	**.73**	**.70**	**.77**

* p<.05; ** p<.01; *** p<.001

sonality diagnostics. It facilitates clinical assessment by qualified staff on the basis of either an interview or observations over the course of therapy. High comorbidity rates generally observed among personality disorders are accommodated for by the representation of personality structure in the form of a profile. The taxonomy of Q-factor scores is founded on empirical data. The application of a Q-sort approach further solves the problem of determining cut-off values, in so far as decisions are made concerning individual behaviors and not criteria that are in part extremely complex. The critique expressed by Widiger [37] is also accounted for; even in the case that a patient fails to fulfill all necessary criteria for the diagnosis of a personality disorder, personality accentuations can be ascertained from the profile. The advantage of the SWAP-200 in therapeutic practice lies above all in the valuable clinical information provided by the plotting of a personality profile. With the aid of the SWAP-200, the facets of a patient's personality can be more distinctly recognized and specified. Lasting approximately one hour, the length of time required for performance and assessment can be considered acceptable.

The data of 18 patients yielded good agreement (r=.69 and r=.76) between the interviewer and an independent observer. Convergent validity was operationalized as the agreement between interviewer/observer and the patient's therapist, and proved to be satisfactory (between r=.54 and r=.68). Based on the median rather than the mean, validity was good (between r=.70 and r=.81). Values of inter-rater agreement and agreement with the therapist thus lie above those reported by Shedler and Westen [34], [35], who found a reliability of r=.61 and validity of r=.54 based on the data of eight patients.

Inter-rater agreement and validity results differed according to whether PD or Q-factor scores were employed. This is particularly noticeable in agreement between interviewer/observer and therapist. PD and Q-factor scores could represent distinctive personality taxonomies and thus be responsible for these differences. High levels of agreement in Q-factor scores between interviewer/observer
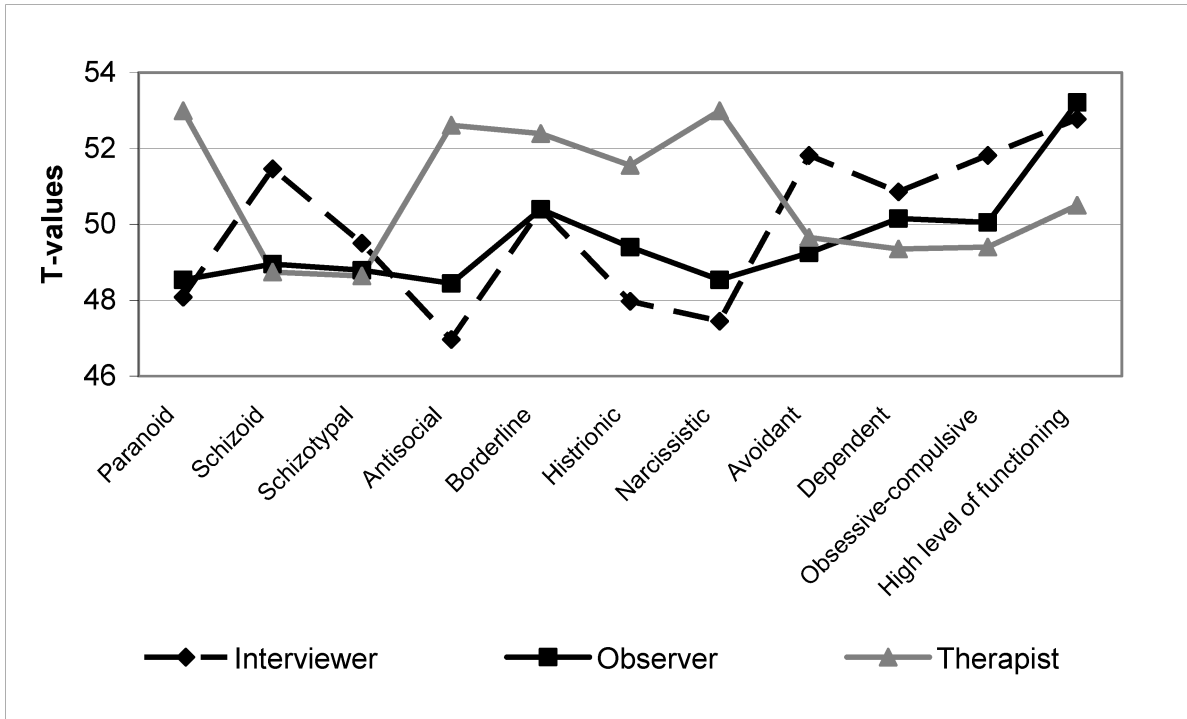
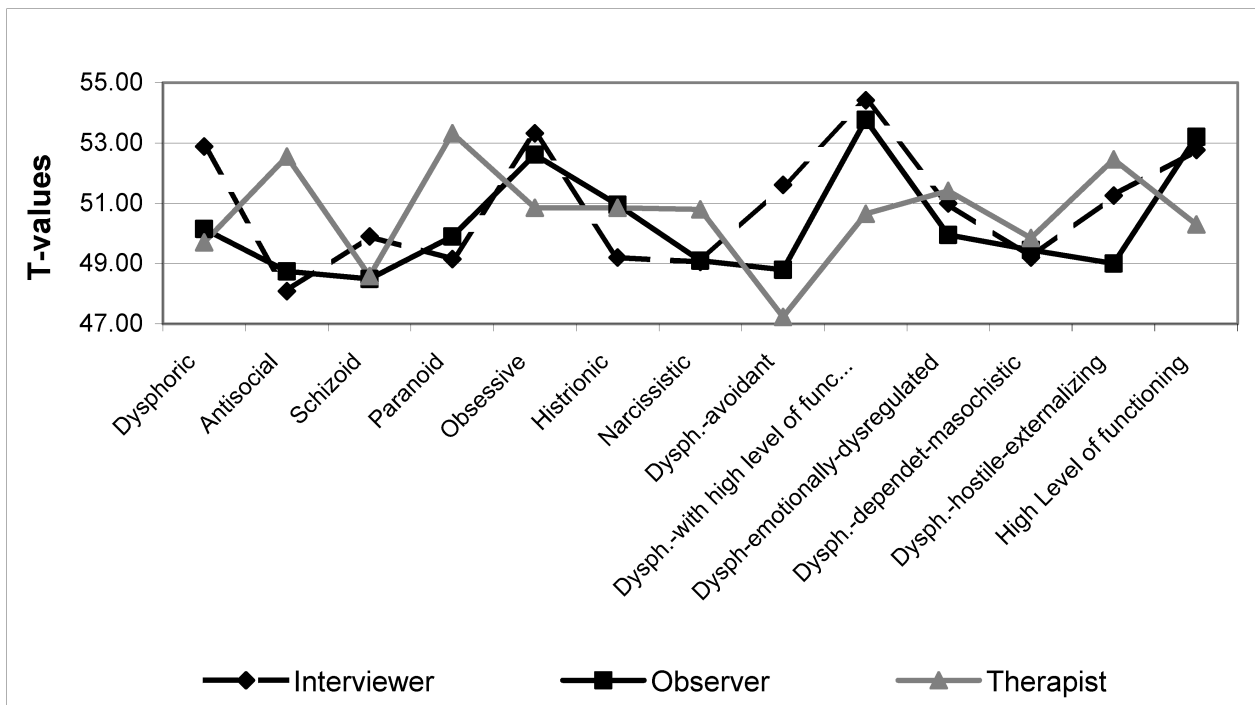**Figure 1: PD score profile for patient 4**

Figure 2: Q-factor profile for patient 4

and therapist could be indicative of a higher validity of Q-factor as compared with PD scores. This is turn could be explained by the fact that Q-factor scores were determined empirically, whereas PD scores were based on the DSM-IV taxonomy.

It should be noted that extreme outlier values were observed for individual patients and that these values have a particularly attenuating effect on the validity. In the studies of Shedler and Westen, validity also proved unsatisfactory for some patients. The authors claimed that this was the case for those patients who did not have a personality disorder. This explanation cannot be drawn upon to explain the outlier values of the present study. A lack of clinical experience on the part of the interviewer and the observer could be responsible for observed results. As is also true of other systems within personality dia-

gnostics, the interviewer is drawn into the patients' maladaptive patterns of interaction. This was clearly seen in the case example of Patient 4, in which the interviewer allowed herself to be influenced in her objectivity and exploration by the dramatic portrayal of a rape. It is of great importance in applying both this and other diagnostic approaches that specific examples are asked for and that particular attention is paid to aspects of the therapeutic relationship.

The question concerning which of the two taxonomies (PD or Q-factor scores) should be employed remains unanswered. While the PD scores adhere to the commonly used DSM-IV taxonomy, the Q-factor scores of Shedler and Westen represent a new taxonomy. This taxonomy could – based on the data of the current study - prove more valid, although also disadvantageous in so far as it entails the introduction of a new system of diagnostic labels. Given that a dimensional approach could solve many of the problems associated with traditional personality diagnostics, further research would appear necessary and promising.

Notes

Conflicts of interest

None declared.

References

1. Blais M, Norman D. A psychometric evaluation of the DSM-IV personality disorder criteria. J Personal Disord. 1997;11:168-76.

2. Block J. The Q-Sort Method in Personality Assessment and Psychiatric Research. Palo Alto, CA: Consulting Psychologists Press; 1978.

3. Bronisch T, Hiller W, Mombour W, Zaudig M. IDCL-P Internationale Diagnose Checkliste für Persönlichkeitsstörungen. Göttingen: Hogrefe; 1995.

4. Bronisch T, Mombour W. Comparison of a diagnostic checklist with a structured interview for the assessment of DSM-III-R and ICD-10 personality disorders. Psychopathology. 1994;27:312-20.

5. Bronisch T, Mombour W. The modern assessment of personality disorders part 2: reliability and validity of personality disorders. Psychopathology. 1998;31:293-301.

6. Bronisch T. Diagnostik von Persönlichkeitsstörungen nach den Kriterien aktueller internationaler Klassifikationssysteme. Verhaltenstherapie.1992;2(2):140-50.

7. Bronisch T. Diagnostik von Persönlichkeitsstörungen. Persönlichkeitsstörungen.1999;3:5-15.

8. Clark L. Resolving taxonomic issues in personality disorders: the value of larger scale analyses of symptom data. J Personal Disord. 1992;6:360-76.

9. Clarkin JF, Martius P, Dammann G, Smole S, Buchheim P. IPO - "Inventory of Personality Organization": Ein Selbstbeurteilunginstrument zur Erfassung der Persönlichkeitsorganisation. Persönlichkeitsstörungen - Theorie und Therapie. 1998;169-72.

10. Diehl JM, Staufenbiel T. Statistik mit SPSS Version 10.0. Frankfurt a.M.: Dietmar Klotz GmbH; 2001.

11. Ferro T, Klein DN, Schwartz JE, Kasch KL, Leader JB. 30-Month Stability of Personality Disorder Diagnoses in Depressed Outpatient. Am J Psychiatry. 1998;155 (5):653-9.

12. First M, Spitzer R, Gibbon M, Williams J, Davis JB, Howes M, Kane J, Pope H, Rounsaville B. Structured Clinical Interview for DSM-III-R Personality Disorders (SCID-II), part II: multi-site test-retest relibility study. Verhaltenstherapie. 1995;9:92-104.

13. Heuft G, Senf W. Praxis der Qualitätssicherung: Das Manual zur Psy-BaDo. Stuttgart: Thieme; 1998.

14. Jablensky A. The classification of personality disorders: critical review and need for rethinking. Psychopathology. 2002;35:112-6.

15. Johnson JG, Williams JBW, Goetz RR, Rabkin JG, Lipsitz JD, Remien RH. Stability and Change in Personality Disorder Symptomatology Findings From a Longitudinal Study of HIV + and HIV - Men. J Abnorm Psychol. 1997;106(1):154-8.

16. Kass F, Skodol AE, Charles E, Spitzer RL, Williams JBW. Scale ratings of DSM-III personality disorders. Am J Psychiatry. 1985;142:627-30.

17. Kuhl J, Kazen M. Persönlichkeits-Stil- und Störungs-Inventar (PSSI) Handanweisung. Göttingen: Hogrefe; 1997.

18. Loranger AW, Sartorius N, Berner W, Buchheim P, Channabasavanna SM, Coid B, Dahl A, Diekstra RFW, Mombour W, Ono Y, Regier DA, Tyrer P, von Cranach M. The international personality disorders examination. The WHO/ADAMHA international pilot study of personality disorders. Arch Gen Psychiatry. 1994;51:215-24.

19. Loranger AW, Susman V, Oldham J, Russakoff M. The Personality Disorder Examination. A preliminary report. J Personal Disord. 1987;1;1-13.

20. Marinangeli MG, Butti G, Scinto A, Di Cicco L, Petruzzi C, Daneluzzo E, Rossi A. Patterns of comorbidity among DSM-III-R personality disorders. Psychopathol. 2000;33:69-74.

21. O'Boyle M, Self D. A comparison of two interviews for DSM-III-R personality disorders. Psychiatry Res. 1990;32:85-92.

22. Oldham JM, Skodo AE, Kellmann, DH, Hyler SE, Doidge N, Rosnick L, Gallaher PE. Comorbidity of Axis I and Axis II Disorders. Am J Psychiatry. 1995;152(4):571-8.

23. Perry J. Problems and considerations in the valid assessment of personality disorders. Am J Psychiatry. 1992;149:1645-53.

24. Pilkonis PA, Heape CL, Proietti JM, Clark SW, McDavid JD, Pitts TE. The reliability and validity of two structured diagnostic interviews for personality disorders. Arch Gen Psychiatry. 1995;52(12):1025-33.

25. Samuels J, Eaton WW, Bienvenu III OJ, Brown CH, Costa PT, Nestadt G. Prevalence and correlates of personality disorders in a community sample. Br J Psychiatry. 2002;180:536-42.

26. Sara G, Raven P, Mann A. A comparison of DSM-III-R and ICD-10 personality disorder criteria in an outpatient population. Psychol Med. 1996;26:151-60.

27. Saß H, Steinmeyer EM, Ebel H, Herpertz S. Untersuchungen zur Kategoriesierung und Dimensionierung von Persönlichkeitsstörungen. Zeitschrift für Klinische Psychologie. 1995;24(3):239-51.

28. Shedler J, Westen D. Refining the measurement of Axis II: A Q-sort procedure for assessing personality pathology. Assessment. 1998;5:335-55.

29. Skodol A, Oldham J, Rosnick I, Kellmann D, Hyler S. Diagnosis of DSM-III-R personality disorders: a comparison of two structured interviews. Int J Methods in Psychiatry Research. 1991;1:13-26.

30. Trautmann-Sponsel R-D, Zaudig M. Diagnostik, Differentialdiagnostik und Komorbidität der Persönlichkeitsstörungen nach ICD-10 und DSM-IV. Persönlichkeitsstörungen, Theorie und Therapie. 1997;1(2):51-62.

31. Westen D, Arkowitz-Westen L. Limitations of Axis II in Diagnosing Personality Pathology in Clinical Practice. Am J Psychiatry. 1998;155(12):1767-71.

32. Westen D, Shedler J, Durrett C, Glass S, Martens A. Personality Diagnoses in Adolescence: DSM-IV Axis II Diagnoses and an Empirically Derived Alternative. Am J Psychiatry. 2003;160(5):952-66.

33. Westen D, Shedler J. A prototyp matching approach to personality disorders: towards DSM -V. J Personal Disord. 2000;14(2):109-26.

34. Westen D, Shedler J. Revising and Assessing Axis II, Part I: Developing a Clinical and Empirically Valid Assessment Method. Am J Psychiatry. 1999;156(2):258-72.

35. Westen D, Shedler J. Revising and Assessing Axis II, Part II: Toward an Empirically Based and Clinically Useful Classification of Personality Disorders. Am J Psychiatry. 1999;156(2):273-85.

36. Westen D. Divergences between clinical and research methods for assessing personality disorders: implications for research and the evolution of axis II. Am J Psychiatry. 1997;154:895-903.

37. Widiger T. Personality disorder dimensional models proposed for DSM-IV. J Personal Disord. 1991;5:386-98.

38. Widiger TA, Cadoret R, Hare R. DSM-IV antisozial personality disorder trail. J Abnorm Psychol. 1996;105:3-16.

39. Widiger TA, Costa PT. Personality and Personality Disorders. J Abnorm Psychol. 1994;103(1):78-91.

40. Wittchen H-U, Zaudig M, Fydrich T. Strukturiertes klinisches Interview für DSM-IV (Skid) Achse I und II. Göttingen: Hogrefe; 1997.

41. Wittchen H-U. Klassifikation und Diagnostik von Persönlichkeitsstörungen. In: Schmitz B, Fydrich T, Limbach K, editors. Persönlichkeitsstörungen: Diagnostik und Psychotherapie. Weinheim: Psychologie Verlags Union; 1996. p. 27-41.

42. Zimmermann M. Diagnosing Personality Disorders. Arch Gen Psychiatry. 1994;51:225-45.

Corresponding author:

Dr. Anke Höflich

Clinic and Policlinic for Psychosomatic Medicine and Psychotherapy, Johannes-Gutenberg-University-Clinic, Untere Zahlbacher Str. 8, 55131 Mainz, Tel.: 49-06131-175485, Fax: 49-06131-176688 hoeflich@uni-mainz.de