Avoidant personality disorder severity index: dimensional structure and psychometric properties

Baljé A.E., Karch J.D., Greeven A., van Giezen A.E., Muste E.H., Arntz A., Spinhoven P.

Personality and Individual Differences 213 (2023) 112268

Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

Avoidant Personality Disorder Severity Index: Dimensional structure and psychometric properties

Astrid E. Baljé [a,b,*], Julian D. Karch [b], Anja Greeven [a,b], Anne E. van Giezen [b], Eelco H. Muste [c], Arnoud Arntz [d], Philip Spinhoven [b]

[a] Department of Anxiety Disorders/Psyq, The Hague, the Netherlands
[b] Institute of Psychology/Leiden University, Leiden, the Netherlands
[c] Cirva GGZ/Dordrecht, the Netherlands
[d] Programme group Clinical Psychology/University of Amsterdam, Amsterdam, the Netherlands

ARTICLE INFO

Keywords:
Psychometric properties
Symptom measure
Standardized interview
Avoidant personality disorder

ABSTRACT

Introduction: The Avoidant Personality Disorder Severity Index (AVPDSI) is developed to provide a standardized clinical interview to measure the severity of and changes in manifestations of avoidant personality disorder (AVPD) in terms of behavior and anxiety, as defined in the Diagnostic and Statistical Manual of Mental Disorders-IV/5 (DSM-IV/5). This first assessment of the AVPDSI examined its dimensional structure and psychometric properties.

Methods: We investigated factor structure, reliability, interrater reliability, validity, clinical cutoff scores, and sensitivity to change in a mixed sample of patients with a primary diagnosis of AVPD, borderline personality disorder (BPD), social anxiety disorder (SAD), panic disorder (PAD) and in non-patients (NP) (n = 305).

Results: The underlying structure indicated a behavior and anxiety dimension. The instrument showed good internal reliability, inter-rater agreement, and a clinical cutoff point with good sensitivity and specificity. Convergent and discriminant validity were good and preliminary results in a treated subsample (n = 20) indicated sensitivity to change.

Conclusion: In our opinion, a new and valid instrument has been introduced, enabling researchers and clinicians to reliably measure the severity of and changes in manifestations of AVPD.

1. Introduction

Avoidant personality disorder (AVPD) is one of the most prevalent personality disorders (PDs) (Weinbrecht et al., 2016). The massive use of avoidance distinguishes AVPD from other PDs, significantly impacts daily functioning, and often has severe consequences both for patients and society (Simonsen et al., 2019). Patients feel inhibited, inadequate, perceive themselves as inferior to others, and avoid social interactions to escape distress (American Psychiatric Association, 2000; American Psychiatric Assosiation, 2013). AVPD is associated with lower quality of life, lower levels of social support, a lower educational level (Olsson & Dahl, 2012; Wilberg et al., 2009) and with comorbid disorders such as dysthymia, major depression, and social anxiety disorder (Lampe et al., 2003; Weinbrecht et al., 2016). Societal costs result, for instance, from unemployment, and the use of social welfare (Olsson & Dahl, 2012). Though promising treatment results have been reported (Bamelis et al., 2014), there is still a lot to be learned with respect to the optimal treatment approach (Simonsen et al., 2019).

In line with this, a measurement tool that can assess fine-grained manifestations of and detailed changes in AVPD symptoms over the shorter term could benefit research and customized care. The dichotomous and coarse-grained design of existing instruments (e.g., Structured Clinical Interview for DSM-IV for Axis II Personality Disorders (SCID- II)) allows for diagnosis but not for a detailed examination of individual manifestations of and gradual changes in AVPD pathology. They mainly evaluate long-term functioning and the presence of patterns that are relatively stable across time (Arntz et al., 2003).

To our knowledge, an instrument to assess current severity and changes over time is not yet available. By translating each of the AVPD criteria of the DSM in concrete behavioral manifestations and

* Corresponding author at: Department of Anxiety Disorders/Psyq, Lijnbaan 4, 2512 VA The Hague, the Netherlands.
E-mail addresses: astrid.balje@psyq.nl (A.E. Baljé), j.d.karch@fsw.leidenuniv.nl (J.D. Karch), a.greeven@psyq.nl (A. Greeven), giezen@fsw.leidenuniv.nl (A.E. van Giezen), eelco@muste.nu (E.H. Muste), A.R.Arntz@uva.nl (A. Arntz), spinhoven@fsw.leidenuniv.nl (P. Spinhoven).

https://doi.org/10.1016/j.paid.2023.112268
Received 21 June 2022; Received in revised form 16 May 2023; Accepted 23 May 2023
Available online 21 June 2023

accompanying levels of anxiety, we aimed to develop a complement to the existing assessment arsenal: i.e., not an instrument intended for diagnostic purposes, but for assessing the severity of and changes in AVPD manifestations in daily life.

This resulted in the development of a semi-structured interview named the Avoidant Personality Disorder Severity Index (AVPDSI). Its items measure the frequency of specific manifestations of AVPD. It objectively determines the current severity level, and gives insight into its tangible effects on a personal level. Since the primary treatment target of clinicians is to reduce the severity of disorders, the AVPDSI provides them with valuable information on how severely a specific patient is impaired in daily life and how treatment affects this severity over time. Since PD patients might vary in the extent and severity of their pathology and their response to treatment, more specific insight into these manifestations may aid in the development of more individualized treatments for those struggling with AVPD. Furthermore, since the AVPDSI provides a quantitative index, it can also advance treatment research by allowing the effect of treatments to be compared across individuals. The objective of the current study was to examine the dimensional structure and psychometric properties of the AVPDSI.

2. Methods

2.1. Participants and procedures

The sample consisted of 305 Dutch participants, between 18 and 65 years, recruited from three large Dutch mental health care institutions (PsyQ, Indigo, and De Viersprong) and from the general population via snowball sampling. The AVPD group consisted of 186 patients, of which the majority (n = 138) were recruited as part of a randomized controlled trial (RCT) (Baljé et al., 2016). The comparison groups, whose size and characteristics can be found in Table 1, consisted of patients with social anxiety disorder (SAD), panic disorder (PAD), borderline personality disorder (BPD), and non-patients (NP). For AVPD and BPD patients, diagnostics at the start of treatment consisted of the SCID-II (First et al., 1997) and the Mini-International Neuropsychiatric Interview (MINI) (Sheehan et al., 1998) or Structured Clinical Interview for DSM- IV for axis I Disorders (SCID-I) (First et al., 1996). SAD and PAD patients were diagnosed through an intake interview, mostly combined with the MINI.

Inclusion criteria were a primary diagnosis of respectively AVPD, BPD, SAD, or PAD, age above 18 years, and sufficient mastery of the Dutch language. Additional criteria for AVPD patients of the RCT were applied (see Baljé et al., 2016). To reduce the contamination of the dataset due to comorbidities, AVPD and BPD patients with respectively more than two BPD or two AVPD traits were excluded, as well as patients with a diagnosis of both SAD and PAD. Non-patients were permitted a maximum score on the Liebowitz Social Anxiety Scale - Self Report (LSAS-SR) of 47 as this cut-off score maximizes specificity in correctly identifying non-socially anxious participants (Rytwinski et al., 2009).

After recruitment, eligible participants gave written consent and completed an online survey distributed via Qualtrics software comprising socio-demographic questions and several self-report questionnaires (see below). Following this, participants were invited for the AVPDSI interview. The interview took 60–90 minutes, was conducted either face-to-face or by phone, and was audiotaped for quality control and inter-rater reliability purposes. To assess sensitivity to change, a subsample of 20 AVPD patients of De Viersprong, participated in the survey and AVPDSI interview before and one month after outpatient or part-time treatment.

2.2. Measures

2.2.1. Avoidant Personality Disorder Severity Index (AVPDSI)

The AVPDSI was developed analogous to the Borderline Personality Disorder Severity Index (BPDSI) (Arntz et al., 2003; Giesen-Bloo et al., 2010). The interview provides a quantitative index of the severity of AVPD pathology over the preceding month (see Appendix A). Item development and selection were based on literature study, exploration of existing instruments, and expert interviews (see Appendix B). The AVPDSI consists of 106 questions related to 53 items. Each item exists of a pair of two questions. The questions are measured on a 6-point Likert scale. All seven AVPD criteria in the DSM IV and 5 are represented by a subscale comprising behavioral manifestations and accompanying anxiety levels. For each pair, the first question relates to the frequency of some kind of approach behavior, and the second question to the frequency of the associated anxiety. For example, one of the items representing the DSM AVPD trait "Is unwilling to get involved with people unless certain of being liked" is "Taking part in an activity with another person". Its manifestations are assessed by asking people about the past month, how many times they spend their free time with one other person, and how many times, in relation to these situations, they experienced fear that people would not like them.

Because people are not well aware of what they avoid, the choice has been made to ask for rapprochement. Reversing these scores results in an index of avoidance. When constructing the AVPDSI, developing the behavioral questions for subscales 4 and 6 was challenging. Scale 4 contains items about being preoccupied with criticism and rejection in social situations (DSM-IV: APD criterion 4) and scale 6 contains items concerning seeing oneself as socially inept, unappealing, or inferior (DSM-IV: APD criterion 6). In these two scales, behavior refers to covert behavior or internal processes, namely thinking about how others will react or how you consider yourself, as opposed to the other scales where behavior is more overt and can be observed.

2.2.2. Liebowitz Social Anxiety Scale – Self Report (LSAS-SR)

The LSAS-SR assesses the dimensional severity of social anxiety

Table 1
Group characteristics on all diagnostic groups and non-patients.

Characteristics	AVPD (N = 186)	BPD (N = 28)	SAD (N = 27)	PAD (N = 30)	NP (N = 34)	Analyses	
	Mean(SD)	Mean(SD)	Mean(SD)	Mean(SD)	Mean(SD)	F(df)	p
Age	31.2 (9.0)	32.6 (9.3)	32.8 (8.1)	36.6 (14.0)	29.3 (9.9)	1.66 (4, 68.1)	0.16
Education[1]	5.8 (1.6)	5.2 (1.8)	5.9 (1.4)	4.7 (1.7)	7.0 (1.2)	12.89 (4, 70.3)	<0.001
	% (N)	% (N)	% (N)	% (N)	% (N)	Chi^2 (df)	p
Sex: female	52.2 (97)	85.7 (24)	33.3 (9)	56.7 (17)	70.6 (24)	19.73 (4)	0.001
Work/study	56.6 (105)	39.3 (11)	74.1 (20)	66.7 (20)	88.2 (30)	15.94 (4)	0.003
Living alone	40.3 (75)	32.1 (9)	51.9 (14)	6.7 (2)	32.4 (11)	20.11 (4)	<0.001

Note: [1] Education: 8 levels, varying from 1 = primary school (partially finished) to 8 = academic education (Bachelors or Master's degree). AVPD = avoidant personality disorder, BPD = borderline personality disorder, SAD = social anxiety disorder, PAD = panic disorder, NP = non-patients, SD = standard deviation, df = degrees of freedom, p = p-value.

disorder symptoms in the past week (Liebowitz, 1987). The 24 items, 11 about social and 13 about performance situations, are all scored with respect to anxiety and avoidance. The LSAS-SR has good psychometric properties (Baker et al., 2002) and can be used to reliably classify individuals with and without social anxiety disorder (Rytwinski et al., 2009).

2.2.3. Acceptance and Action Questionnaire (AAQ-II)

The AAQ-II is a 10-item self-report Likert scale measuring experiential avoidance or psychological inflexibility (Bernaerts et al., 2012). It has a good level of reliability and validity (Bond et al., 2011). In the current study, scoring was reversed so that higher scores represented more experiential avoidance.

2.2.4. Inventory of Depressive Symptoms – Self Report (IDS-SR)

The IDS-SR consists of 30 items and measures depressive symptoms severity based on DSM-IV. It has highly acceptable psychometric properties with reasonable internal consistency, inter-rater reliability, and concurrent and discriminant validity (Trivedi et al., 2004).

2.2.5. Schema Mode Inventory-2 (SMI-2)

The SMI assesses different schema modes (Young et al., 2007). The current study used three modes of the SMI-2: the avoidant protector (AVP, 10 items, characterized by situational/behavioral avoidance), enraged child (EC, seven items, characterized by angry feelings as a reaction to unfulfilled needs or unfair treatment) and attention- and approval seeker (AAS, six items, characterized by attention-attracting behavior and a dramatic interaction style). The internal consistency of these three modes was in the range 0.86–0.92.

2.2.6. Structured Clinical Interview for DSM Disorder-II Personality Questionnaire (SCID-screen)

The SCID-II interview used to diagnose DSM-PDs, has a self-report version, which can be used as a screener: the SCID-screen (Ekselius et al., 1994). The Avoidant, Histrionic and Narcissistic PD (7, 7, and 17 questions, respectively) subsets were used. Overall, internal consistency and inter-rater reliability of the DSM-IV personality disorder criteria, assessed using the pre-screener, meet or exceed standard cutoffs (Ryder et al., 2007).

3. Statistical analysis

Statistical analyses were performed in Statistical Package for the Social Sciences (SPSS) version 25 and in R packages lavaan (Rosseel, 2012) and psych (Revelle, 2015).

3.1. Dimensional structure

Confirmatory factor analyses (CFA) were performed to assess the dimensional structure of the AVPDSI. Model fit was evaluated using the chi-square statistic, Tucker-Lewis Index (TLI), the comparative fit index (CFI), and the standardized root mean square residual (SRMR). Both a one-factor structure and a two-factor structure were tested. In the two-factor model, each of the 53 avoidance and 53 anxiety items were supposed to load on a separate factor for avoidance, respectively anxiety. For both models, residual variances and factor loadings were not constrained. Power analysis for both models was performed using the method described in Moshagen and Erdfelder (2016). The lowest root mean squared error value within the 90 % confidence interval was used as effect size to obtain a conservative estimate. Subsequently, an exploratory factor analysis (EFA) using minimum residuals (Harman & Jones, 1966) and oblimin rotation was executed to explore the underlying dimensional structure of the AVPDSI further.

3.2. Reliability

Internal reliability was assessed by Cronbach's alpha. To evaluate interrater reliability a total of 48 audiotapes were selected from all recorded interviews using block randomization to represent all diagnoses. Second raters (three psychologists and three master students in clinical psychology) were trained to use and rate the AVPDSI. Inter-rater reliability was estimated by computing intraclass correlation coefficients (ICC using a one-way random-effects model) for each avoidance ($n = 53$) and anxiety question ($n = 53$), final avoidance and anxiety subscales, and the total AVPDSI scale.

3.3. Validity

Construct validity was tested by examining differences between the diagnostic groups using a one-way between-subjects Welch's Analysis of Variance (ANOVA) (Boateng et al., 2018), followed by contrast tests of differences between the AVPD group and other groups and by computing the effect sizes of these differences with Cohen's d_s (Lakens, 2013). Differences were also investigated in an ANCOVA controlling for possible effects of gender, age and educational level as covariates using bootstrap as inference approach, as this is robust to potential violations of homoscedasticity.

Convergent and divergent validity was assessed by calculating Pearson's correlations (Boateng et al., 2018; Churchill, 1979) between the AVPDSI and its subscales and related measures adjusted for attenuation. We used Cronbach's alpha of these measures in our sample as reliability estimates. We expected correlations with similar constructs to be above 0.50, with related constructs to be between 0.30 and 0.50, and with unrelated constructs to be lower than 0.30 (Prinsen et al., 2018).

3.4. Clinical cutoff scores

Receiver operating characteristics (ROC) analyses were performed to determine optimal cutoff scores between AVPD and non-patients, based on Youden- and Liu's indices and Euclidean distances (Liu, 2012; Youden, 1950).

3.5. Sensitivity to change

Sensitivity to change of the AVPDSI was explored by executing paired sample *t*-tests and calculating Cohen's d_{rm} (Lakens, 2013) of the subscales and total scale before and one month after treatment.

3.6. Missing data

AVPDSI questions of subscales related to DSM-IV/5 criterion one concern avoidance and anxiety in work- or school-related situations. Some participants were not or so rarely involved in these situations that these questions did not apply. There were a few other missing data for the AVPDSI (0.07 %). Missing data were handled in CFA using full information maximum likelihood (FIML) and in EFA and reliability analyses using the correlation matrix based on pairwise complete observations. In all other analyses, missing values were treated by mean substitution using the available scores of that person on the (sub)scales.

4. Results

4.1. Dimensional structure of the AVPDSI

4.1.1. Confirmatory factor analyses

Both the hypothesized one-factor ($\chi 2 < .001; CFI = .49; TLI = .48; SRMR = .10$) and two-factor model ($\chi 2 < .001; CFI = .50; TLI = .49; SRMR = .10$) failed to provide an adequate fit (see Appendix C Table C.1). The estimated power was >99 % for both models (see Appendix D for power analysis).

4.1.2. Exploratory factor analyses

Inspection of the data did not reveal substantial violations of normality. Item pair 18a and 18b (referring to constantly focusing on someone's reaction and facial expression) was removed because, in retrospect, the behavioral question was not phrased clearly, was not related to overt behavior, and some participants reported difficulties with understanding the difference between the behavioral and anxiety question as also indicated by the high correlation between both questions. No other items were excluded. For the total scale and both subscales, reliability-if-item-deleted analysis showed that reliability did not improve by removing any item.

Since the CFA rejected the hypothesized structure, we subsequently performed an EFA to examine whether an alternative structure was supported by the data. Specifically, we used EFA using minimum residuals and oblimin rotation. A first EFA on all items showed that some behavioral questions loaded on the same factor as the anxiety questions (see Appendix E Fig. E.1). A closer look at the inter-item correlations indicated that the correlations of the 15 item pairs of scale 4 and 6 were high. Values varied from 0.62 to 0.91, with 12 out of 15 well above 0.7. Because of the similarity of the questions of each pair and the high inter-item correlations, it was decided to compute the mean scores for each question pair of scale 4 and 6, which were used in the remaining analyses.

Subsequently, an EFA was conducted on the resulting 89 items. The Kaiser-Meyer-Olkin measure (KMO) was 0.74, indicating sampling adequacy for the analysis. Bartlett's test of sphericity, $\chi 2\,(3916) = 16830.92, p < .001$ provided evidence against the null hypothesis of the items being uncorrelated. Based on the scree plot (elbow method, see Fig. 1), the correlational data, and interpretability, a two-factor solution was chosen. The item loadings strongly suggest a "behavior" factor and an "anxiety" factor, since using a threshold of 0.2, all behavior items load on the first factor only and all anxiety items on the second factor only (see also Fig. 2). From here on, these scales will be labeled: AVPDSI-Behavior (AVPDSI-B) and AVPDSI-Anxiety (AVPDSI-A).

The AVPDSI-B factor encompasses the 37 items of scales 1A, 2A, 3A, 5A and 7A, which all relate to avoidant behavior, whereas the AVPDSI-A factor is made up of the 37 items of scale 1B, 2B, 3B, 5B, 7B and the 15 combined behavior and anxiety items of scale 4 and 6. The AVPDSI-B factor accounts for 9 %, and the AVPDSI-A factor for 22 % of the variance. The correlation between both factors was 0.12. Factor loadings of the AVPDSI-A factor were all above 0.4. With respect to the AVPDSI-B factor, 12 of the 37 loadings were lower than 0.4 (Appendix C Table C.2 shows the factor loadings and the communality scores after rotation).

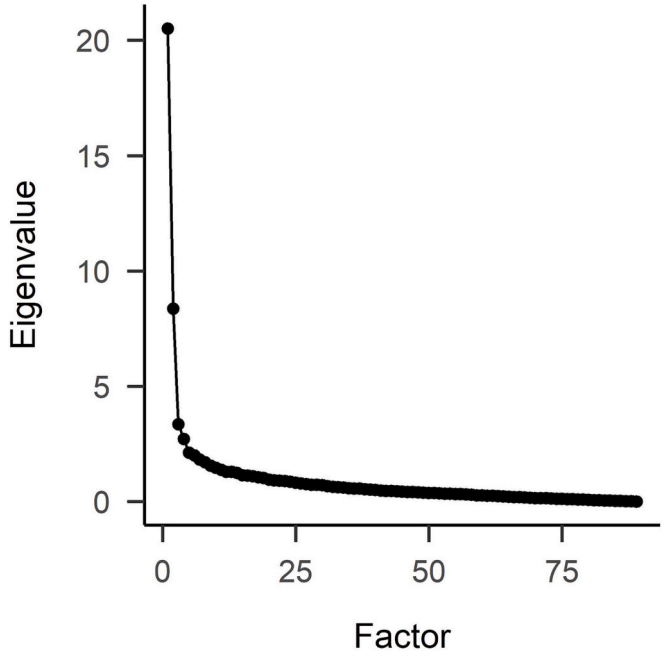

Fig. 1. Scree plot.

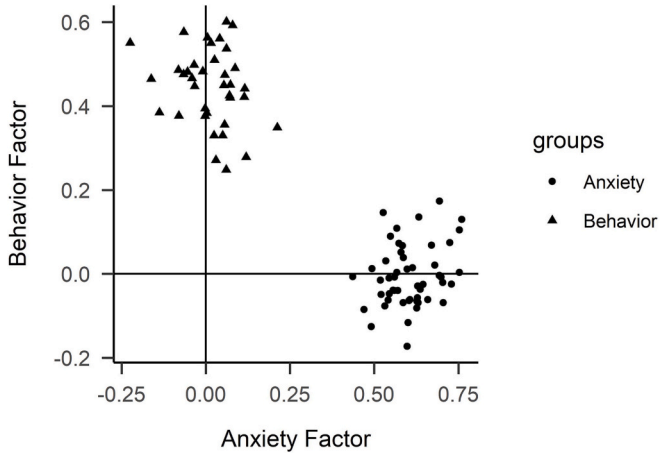

Fig. 2. Factor plot.

In line with the results of the exploratory factor analysis, the AVPDSI-B scale (37 items) consists of the items of scale 1A, 2A, 3A, 5A and 7A. The AVPDSI-A scale consists of the items of scale 1B, 2B, 3B, 5B, 7B (37 items) and the combined items of the corresponding behavior and anxiety questions of scale 4 (7 items) and 6 (8 items). Both subscale scores are obtained by averaging the respective items and multiplying this average by 10. The AVPDSI total score is calculated by adding the AVPDSI-B and AVPDSI-A score (see Appendix F for the scoring key).

4.2. Reliability

Reliability of the AVPDSI score was high (see Table 2). The correlation between the subscales was low (0.12 in the total and 0.19 in the AVPD sample). Consequently, we performed all remaining psychometric analyses separately for the AVPDSI-B, AVPDSI-A, and total AVPDSI scores.

Mean inter-item correlations in the total sample for the total scale and both subscales were satisfactory to good (see Table 2; (Clark & Watson, 1995)). The relatively lower mean inter-item correlations in the total scale compared to those in both subscales (see Table 2) are understandable in light of the low correlation between the two scales.

4.3. Interrater reliability

The median ICCs for individual avoidance and anxiety questions were generally excellent (Koo & Li, 2016), respectively 0.97 and 0.99 (range .56 – .99 and .76 – .99; see Appendix C Table C.3). There was a high level of absolute agreement in ratings of individual questions (avoidance 94.7 %, anxiety 96 %). Both the AVPDSI-B scale (ICC = 0.993) and the AVPDSI-A scale (ICC = 0.997), and the total scale (ICC = 0.998) had excellent interrater reliability.

4.4. Validity

4.4.1. Comparison with diagnostic groups

One-way between-subjects ANOVA indicated significant differences between the diagnostic groups for the AVPDSI-B, AVPDSI-A, and total AVPDSI scale, respectively $F(4, 69.49) = 32.94$, $p < 0.001$, $F(4, 73.99) = 65.18$, <0.001 and $F(4, 71.82) = 92.49$, $p < 0.001$. Contrasts revealed that mean scores of the AVPD group differed from BPD, SAD, PAD, and NP, except for AVPD versus SAD (AVPDSI-B), and AVPD versus BPD (AVPDSI-A) (see Table 3). Correction for age, educational level, and gender led to similar results. Cohen's d_s of significant contrasts ranged from 0.41 to 2.83.

Table 2
Reliability of the AVPDSI scales for the total and AVPD sample.

Scale	Number of items	Cronbach alpha (95 % CI)	Inter-item correlation		Item-rest correlation	
			Mean	Median	Mean	Range
Total (n = 313)						
AVPDSI-B	37	0.90 [0.89, 0.92]	0.20	0.19	0.43	.25 – .58
AVPDSI-A	52	0.97 [0.96, 0.97]	0.37	0.36	0.59	.43 – .76
AVPDSI	89	0.95 [0.95, 0.96]	0.18	0.14	0.46	.05 – .75
AVPD (n = 194)						
AVPDSI-B	37	0.85 [0.81, 0.88]	0.13	0.13	0.37	– .01 – .52
AVPDSI-A	52	0.95 [0.94, 0.96]	0.28	0.27	0.51	.32 – .67
AVPDSI	89	0.91 [0.89, 0.93]	0.09	0.07	0.39	– .15 – .63

AVPDSI = avoidant personality disorder severity index, AVPD = avoidant personality disorder, AVPDSI-B = avoidant personality disorder severity index - behavior, AVPDSI-A = avoidant personality disorder severity index - anxiety.

4.4.2. Convergent and discriminant validity

Both AVPDSI subscales and the total AVPDSI scale showed adjusted correlations larger than 0.5 with three LSAS-SR subscales (Avoidance, Anxiety, and total LSAS-SR), the AVP mode of the SMI, and the AVPD as measured with the SCID screener (see Table 4), indicating good convergent validity. As measured with the IDS-SR and the AAQ-II, correlations with less related constructs were somewhat lower. The negative correlation of −0.56 of Avoidance with attention and approval-seeking (SMI AAS) was also expected. Relatively lower correlations were also found between the AVPDSI total and AVPDSI-A scale with histrionic PD. Low correlations with the EC mode of the SMI and the Narcissistic PD of the SCID screener indicate discriminant validity of the AVPDSI.

4.5. Clinical cutoff scores

All three indices (Youden, Liu, Euclidean distance) gave the same results. The cutoff for the total score on the AVPDSI between AVPD and NP turned out to be 37.04 with a corresponding sensitivity of 0.98 and specificity of 0.94. The cutoff of the AVPDSI-B and AVPDSI-A scale were respectively 26.08 and 9.86. Sensitivity and specificity were for AVPDSI-B, both 0.88, and for AVPDSI-A 0.90 and 0.88 (see Appendix E Fig. E.2 for the ROC curves and cutoffs).

4.6. Sensitivity to change

Paired *t*-tests showed a significant decrease in the mean scores of the AVPDSI-B ($t(19) = 4.20$, $p < .001$, $M_d = -4.46$) and AVPDSI ($t(19) = 2.32$, $p < .003$, $M_d = -4.22$) after treatment. The effect size of the decrease was large for the AVPDSI-B scale (Cohen's d_{rm} 0.83) and medium for the AVPDSI total scale (Cohen's d_{rm} 0.47). On the AVPDSI-A scale, no significant change in mean scores was found ($t(19) = -0.03$, $p < .98$, $M_d = -0.05$) with an effect size of zero (Cohen's $d_{rm} - 0.01$).

Next, sensitivity to change of the AVPDSI was compared with sensitivity to change of related measures. Comparison with the LSAS-SR showed a similar effect size for the AVPDSI-B scale and the avoidance scale of the LSAS-SR of respectively a Cohen's d_{rm} 0.83 and 0.82. Contrary to the AVPDSI-A where no effect was found, the LSAS-SR anxiety scale had a Cohen's d_{rm} of 0.57. The effect size of the APVDSI total is somewhat smaller than the LSAS-SR total, 0.47 versus 0.73. This corresponds to the absence of a significant change in the AVPDSI-A scale.

5. Discussion

The primary purpose of this study was to evaluate the psychometric properties of the AVPDSI: a new semi-structured interview to measure short-term changes in the severity of AVPD manifestations. In general, the results of this study were positive: the AVPDSI appeared to be a valid and reliable instrument, sensitive to change in our mixed sample consisting of patients with AVPD, BPD, PD, SAD, and a group of healthy controls, providing researchers and clinicians with a suitable outcome measure for both clinical practice and treatment studies.

Though somewhat different from our initially hypothesized structure, a meaningful distinction emerged, representing a behavior and anxiety dimension. The resulting behavior (AVPDSI-B) and anxiety (AVPDSI-A) sub-scales were reliable and valid. Correlations of the AVPDSI-B with the avoidance scales of the LSAS-SR and SMI confirm that measuring the reversed score of the frequency of approach behavior provides an index of the severity of avoidance. The higher mean scores on this AVPDSI-B scale for the AVPD group compared to other diagnostic groups also point in this direction. The AVPDSI and its sub-scales have good reliability and excellent interrater reliability. Group comparisons and examined relations with other constructs indicate good convergent and discriminant validity and the ability of the AVPDSI to differentiate between diagnostic groups. We derived a clinical cutoff point, which had good sensitivity and specificity, and thus can be used to distinguish AVPD patients from non-patients.

The absence of a significant difference on the AVPDSI-B sub-scale between AVPD and SAD can be explained by the similarity between both diagnostic groups (Reich, 2014). Secondly, the sample size of our SAD group was relatively small, making it difficult to detect an actual difference between both groups. Thirdly, the absence of AVPD in the SAD group was not formally checked by a diagnostic interview. Finally, SAD was present in all AVPD patients who participated in the RCT (Baljé et al., 2016).

Corrected correlations of AVPDSI subscales with external measures showed some similarity. Unexpected were the similar correlations with the LSAS subscales. This finding can be explained by the high

Table 3
Comparison of AVPDSI scores, AVPD group versus other diagnostic groups.

Group	N	AVPDSI - B		AVPDSI - A		AVPDSI	
		M (SD)	ES d_s	M (SD)	ES d_s	M (SD)	ES d_s
AVPD	186	32.6 (5.9)		20.6 (9.0)		53.2 (9.7)	
BPD	28	30.2 (5.6)*	0.41	16.7 (9.6)	0.43	47.0 (11.1)**	0.63
SAD	27	31.7 (6.6)	0.15	14.6 (8.1)**	0.67	46.4 (7.9)**	0.72
PAD	30	30.0 (6.0)*	0.45	7.1 (7.0)**	1.54	37.1 (8.4)**	1.69
NP	34	21.2 (5.2)**	1.98	5.4 (4.2)**	1.79	26.6 (7.2)**	2.83

Note: Effect sizes: small 0.2, medium 0.5, large 0.8. * $p < 0.05$, ** $p < 0.01$, 2 tailed. AVPDSI = avoidant personality disorder severity index, AVPDSI-B = avoidant personality disorder severity index - behavior, AVPDSI-A = avoidant personality disorder severity index - anxiety, M = mean, SD = standard deviation, ES d_s = effect size d_s (Cohen's d_s), AVPD = avoidant personality disorder, BPD = borderline personality disorder, SAD = social anxiety disorder, PAD = panic disorder, NP = non-patients.

Table 4
Corrected and uncorrected Pearson's correlations of the AVPDSI-B and AVPDSI-A scales and AVPDSI with other questionnaires and/or scales.

Scale/questionnaire	AVPDSI-B		AVPDSI-A		AVPDSI	
	Correlation [95 % CI][b]	Corrected correlation[a]	Correlation [95 % CI][b]	Corrected correlation[a]	Correlation [95 % CI][b]	Corrected correlation[a]
LSAStot	0.52** [0.44,0.60]	**0.58**	0.59** [0.51,0.66]	**0.62**	0.74** [0.69,0.79]	**0.80**
LSASav	0.53** [0.44,0.60]	**0.59**	0.56** [0.48,0.64]	**0.60**	0.72** [0.67,0.77]	**0.79**
LSASanx	0.49** [0.40,0.58]	**0.55**	0.59** [0.51,0.66]	**0.62**	0.73** [0.67,0.78]	**0.79**
IDS	0.39** [0.29,0.48]	0.44	0.53** [0.44,0.60]	**0.57**	0.62** [0.55,0.68]	**0.69**
AAQ-II	0.33** [0.23,0.43]	0.38	0.53** [0.45,0.61]	**0.58**	0.60** [0.52,0.66]	**0.67**
SMI AVP	0.56** [0.48,0.63]	**0.64**	0.52** [0.44,0.60]	**0.56**	0.71** [0.65,0.76]	**0.78**
SMI EC	0.002 [−0.11,0.11]	*0.002*	0.16** [0.05,0.27]	*0.18*	0.13** [0.02,0.24]	*0.15*
SMI AAS	−0.48** [−0.56,-.0.38]	**−0.58**	−0.12 [−0.23,-0.01]	*−0.14*	−0.35** [−0.44,-0.25]	−0.41
SCS AVPD[c]	0.51**[0.42,0.60]	**0.62**	0.62** [0.54,0.69]	**0.71**	0.75** [0.69,0.80]	**0.88**
SCS NAR[d]	0.01 [−0.11,0.12]	*0.01*	0.13* [0.01,0.24]	*0.15*	0.10* [−0.02,0.22]	*0.13*
SCS HIS[e]	−0.29** [−0.40,-0.18]	−0.47	−0.09 [−0.21,0.03]	*−0.14*	−0.23* [−0.34,-0.11]	−0.35

Note: Correlation (2 tailed) is significant at: *0.05 level and **0.01 level. Note [a] corrected correlation is based on the internal consistency of these measures in our sample, i.e., by estimating correlations if reliabilities would be perfect, to give a better estimate of true correlations. Cronbach alpha's: LSAS total 0.97, LSAS avoidance 0.94, LSAS anxiety 0.95, IDS 0.90, AAQ-II 0.88, SMI AVP 0.91, SMI EC 0.86, SMI AAS 0.79, SCS AVPD 0.80, SCS NAR 0.70, SCS HIS 0.46. Corrected correlation: **Bold:** correlations of |0.5| and higher; Normal font: correlations between |0.3 and 0.5| *Italic*: Correlations of |0.3| and smaller. Note [b] Values in square brackets indicate the 95 % confidence interval of the correlations. Notes [c] n = 270; [d] n = 274; [e] n = 275 (sample size differs due to missing values in the APD sample). Abbreviations: AVPDSI = Avoidant personality disorder severity index, r = Pearson correlation, r corr. = corrected Pearson correlation, CI = Confidence interval, LSAS = Liebowitz social anxiety scale, tot = total, av = avoidance, anx = anxiety, IDS = Inventory of depressive symptoms, AAQ-II = Acceptance and action questionnaire, SMI = Schema mode inventory, AVP = Avoidant protector, EC = Enraged child, AAS = Attention- and approval seeker, SCS = SCID screener, AVPD = Avoidant personality disorder, NAR = Narcissistic personality disorder, HIS = Histrionic personality disorder.

correlations of the LSAS anxiety and behavior scales with the LSAS total score (0.98), indicating that they are indistinguishable from each other (Fresco et al., 2001; Heimberg et al., 1999).The behavior scale of the AVPDSI showed, compared to the anxiety scale, as expected, a somewhat higher corrected correlation with the avoidant protector mode that relates to situational avoidance behavior (SMI AVP; 0.64 vs. 0.56). The anxiety scale correlated stronger with the AAQ which measures experiential avoidance, i.e., avoidance of distressing internal experiences. Furthermore, important differences were found between the AVPDSI-B and AVPDSI-A in correlations with the avoidant protector mode and the histrionic PD scale. The higher negative correlations of the behavioral scale compared to the anxiety scale with the SMI AAS; (−0.58 vs. −0.14) and the SCS HIS (−0.47 vs. −0.14) were as hypothesized. Both measures ask about histrionic behaviors, such as attention seeking and being the center of attention, which are exactly opposite to the behavior of AVPD patients. Furthermore, the correlation of the total AVPDSI scale with external criteria was higher than that of its subscales, indicating that each subscale contributes individually to the total correlation. In sum, each scale seems to focus on related but distinct properties indicating that both the behavior and the anxiety indices contribute to the dysfunction of patients with AVPD.

The results regarding sensitivity to change might indicate that reducing avoidance behavior precedes a decrease in anxiety symptoms (which remained stable). For example, during cognitive behavioral therapy for clients with symptoms of social anxiety disorder, avoidance behavior proved to be predictive of general anxiety at a subsequent time point, but not vice versa (Lervik et al., 2021).

The study also has some limitations. Firstly, our initially hypothesized factor structure was not confirmed by CFA. This finding was explained by EFA, which presented a meaningful two-factor structure with a behavior and anxiety dimension as hypothesized, though somewhat different from the tested two-factor model. It showed that some questions originally intended to measure behavioral manifestations loaded on the anxiety factor. These were the questions on AVPD criteria: being preoccupied with criticism and rejection in social situations (DSM-IV: AVPD criterion 4) and concerning seeing oneself as socially inept, unappealing, or inferior (DSM-IV: APD criterion 6). A closer look revealed that both criteria refer to covert behavior or internal processes, namely thinking about how others will react or how you consider yourself. This contrasts with the other five AVPD traits, which concern more overt and observable behavior. These findings, together with the high correlation of question pairs of trait 4 and 6, led to the decision to combine the scores of these question pairs in the total anxiety score. Future work should confirm the structure identified by EFA with CFA in a new data set, since performing CFA after EFA on the same data leads to invalid results (Fokkema & Greiff, 2017). Secondly, we were not able to rule out psychiatric problems in the NP group or the presence of PDs in the PAD and SAD group by clinical interview due to practical and financial constraints. This might have influenced differences between groups, due to the possible presence of SAD in NP or AVPD in SAD and PAD patients. Thirdly, the results of the sensitivity to change analysis might have been influenced by the small sample size and the relatively short time frame between completing treatment and the interview. Future research should focus on replication, addressing abovementioned issues, preferably in subgroups clearly distinguished using diagnostic interviews, and larger samples of comparison groups.

Despite these limitations, a new reliable, and valid instrument for measuring the severity of AVPD according to DSM-IV/5 criteria has been introduced. The AVPDSI enables researchers and/or clinicians to take individual and gradual differences into account, enabling research and personalized care for a neglected diagnostic group in clinical research on PDs.

Statement of ethics

The Medical Ethical Committee (MEC) of the Leiden University Medical Centre (LUMC) in Leiden approved the study (protocol number p12.165). All participants gave written informed consent.

Funding

The Parnassia Group supported the study. Authors received external financial funding from Stichting MIND (Fonds Psychische Gezondheid).

Credit authorship contribution statement

AB designed, conducted the study and wrote the manuscript. AB and JK performed the statistical analyses. AG, AvG, EM, AA, and PhS participated in the design and coordination and helped to draft the manuscript. All authors read and approved the final manuscript.

Declaration of competing interest

The authors have no conflicts of interest to declare.

Data availability

Participants of this study did not agree for their data to be shared publicly, so supporting data is unavailable. The AVPDSI (Dutch and English version) is available upon request.

Acknowledgments

The authors gratefully thank all individuals who participated in the study: clinical experts, clinicians, research assistants, and students who contributed to the development of the AVPDSI and data collection; and the organizations who enabled the data collection: Psyq, The Viersprong, and Indigo. Also, we thank Dr. Maria Sherwood-Smith and Maurice van Sabben for their contribution to translating the Dutch version of the AVPDSI into English.

Appendices. Supplementary data

Supplementary data/appendices to this article can be found online at https://doi.org/10.1016/j.paid.2023.112268.

References

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th-TR ed.). American Psychiatric Association. https://doi.org/10.1016/j.psychres.2011.06.006

American Psychiatric Assosiation. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). American Psychiatric Association.

Arntz, A., Van den Hoorn, M., Cornelis, J., Verheul, R., Van den Bosch, W. M. C., & De Bie, A. J. H. T. (2003). Reliability and validity of the borderline personality disorder severity index. *Journal of Personality Disorders, 17*(1), 45–59. https://doi.org/10.1521/pedi.17.1.45.24053

Baker, S. L., Heinrichs, N., Kim, H. J., & Hofmann, S. G. (2002). The Liebowitz social anxiety scale as a self-report instrument: A preliminary psychometric analysis. *Behaviour Research and Therapy, 40*(6), 701–715. https://doi.org/10.1016/S0005-7967(01)00060-2

Baljé, A., Greeven, A., van Giezen, A., Korrelboom, K., Arntz, A., & Spinhoven, P. (2016). Group schema therapy versus group cognitive behavioral therapy for social anxiety disorder with comorbid avoidant personality disorder: Study protocol for a randomized controlled trial. *Trials, 17*(1), 1–13. https://doi.org/10.1186/s13063-016-1605-9

Bamelis, L. L. M., Evers, S. M. A. A., Spinhoven, P., & Arntz, A. (2014). Results of a multicenter randomized controlled trial of the clinical effectiveness of schema therapy for personality disorders. *American Journal of Psychiatry, 171*(3), 305–322. https://doi.org/10.1176/appi.ajp.2013.12040518

Bernaerts, I., De Groot, F., & Kleen, M. (2012). De AAQ-II (Acceptance and Action Questionnaire-II), een maat voor experiëntiële vermijding: Normering bij jongeren = The Acceptance and Action Questionnaire-II (AAQ-II). *Gedragstherapie, 45*(4), 389–399.

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health, 6*(June), 1–18. https://doi.org/10.3389/fpubh.2018.00149

Bond, F. W., Hayes, S. C., Baer, R. A., Carpenter, K. M., Guenole, N., Orcutt, H. K., … Zettle, R. D. (2011). Preliminary psychometric properties of the acceptance and action questionnaire-II: A revised measure of psychological inflexibility and experiential avoidance. *Behavior Therapy, 42*(4), 676–688. https://doi.org/10.1016/j.beth.2011.03.007

Churchill, G. A. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research, 16*(1), 64. https://doi.org/10.2307/3150876

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*(3), 309–319. https://doi.org/10.1037/1040-3590.7.3.309

Ekselius, L., Lindström, E., von Knorring, L., Bodlund, O., & Kullgren, G. (1994). SCID II interviews and the SCID Screen questionnaire as diagnostic tools for personality disorders in DSM-III-R. *Acta Psychiatrica Scandinavica, 90*(2), 120–123. https://doi.org/10.1111/j.1600-0447.1994.tb01566.x

First, M. B., Gibbon, M., Spitzer, R. L., Williams, J. B. W., & Benjamin, L. S. (1997). *User's guide for the structured clinical interview for DSM-IV axis II personality disorders (SCID-II)*. American Psychiatric Association.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1996). *Structured clinical interview for DSM-IV axis I disorders – Patient edition (SCID-I/P), version 2.0*. New York State Psychiatric Institute.

Fokkema, M., & Greiff, S. (2017). How performing PCA and CFA on the same data equals trouble. In *, Vol. 33, Issue 6. European Journal of Psychological Assessment* (pp. 399–402). Hogrefe Publishing GmbH. https://doi.org/10.1027/1015-5759/a000460.

Fresco, D. M., Coles, M. E., Heimberg, R. G., Liebowitz, M. R., Hami, S., Stein, M. B., & Goetz, D. (2001). The Liebowitz Social Anxiety Scale: A comparison of the psychometric properties of self-report and clinician-administered formats. *Psychological Medicine, 31*(6), 1025–1035. https://doi.org/10.1017/S0033291701004056

Giesen-Bloo, J. H., Wachters, L. M., Schouten, E., & Arntz, A. (2010). The Borderline Personality Disorder Severity Index-IV: Psychometric evaluation and dimensional structure. *Personality and Individual Differences, 49*(2), 136–141. https://doi.org/10.1016/j.paid.2010.03.023

Harman, H. H., & Jones, W. H. (1966). Factor analysis by minimizing residuals (minres). *Psychometrika, 31*(3), 351–368. https://doi.org/10.1007/BF02289468

Heimberg, R. G., Horner, K. J., Juster, H. R., Safren, S. A., Brown, E. J., Schneier, F. R., & Liebowitz, M. R. (1999). Psychometric properties of the Liebowitz Social Anxiety Scale. *Psychological Medicine, 29*(1), 199–212. https://doi.org/10.1017/S0033291798007879

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*(NOV), 1–12. https://doi.org/10.3389/fpsyg.2013.00863

Lampe, L., Slade, T., Issakidis, C., & Andrews, G. (2003). Social phobia in the Australian National Survey of Mental Health and Well-Being (NSMHWB). *Psychological Medicine, 33*(4), 637–646. https://doi.org/10.1017/S0033291703007621

Lervik, L. V., Hoffart, A., Knapstad, M., & Smith, O. R. F. (2021). Exploring the temporal associations between avoidance behavior and cognitions during the course of cognitive behavioral therapy for clients with symptoms of social anxiety disorder. *Psychotherapy Research*. https://doi.org/10.1080/10503307.2021.1930243

Liebowitz, M. R. (1987). Social phobia. *Modern Problems of Pharmacopsychiatry, 22*, 141–173.

Liu, X. (2012). Classification accuracy and cut pointselection. *Statistics in Medicine, 31*(23), 2676–2686. https://doi.org/10.1002/sim.4509

Moshagen, M., & Erdfelder, E. (2016). A new strategy for testing structural equation models. *Structural Equation Modeling, 23*(1), 54–60. https://doi.org/10.1080/10705511.2014.950896

Olsson, I., & Dahl, A. A. (2012). Avoidant personality problems - Their association with somatic and mental health, lifestyle, and social network. A community-based study. *Comprehensive Psychiatry, 53*(6), 813–821. https://doi.org/10.1016/j.comppsych.2011.10.007

Prinsen, C. A. C., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., de Vet, H. C. W., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research, 27*(5), 1147–1157. https://doi.org/10.1007/s11136-018-1798-3

Reich, J. (2014). Avoidant personality disorder and its relationship to social anxiety disorder. In *Social anxiety: Clinical, developmental, and social perspectives* (3rd ed., pp. 27–44). Elsevier Inc.. https://doi.org/10.1016/B978-0-12-394427-6.00002-9

Revelle, W. (2015). Procedures for personality and psychological research. *R Package Version, 1*(8).

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2). https://doi.org/10.18637/jss.v048.i02

Ryder, A. G., Costa, P. T., & Bagby, R. M. (2007). Evaluation of the SCID-II personality disorder traits for DSM-IV: Coherence, discrimination, relations with general personality traits, and functional impairment. *Journal of Personality Disorders, 21*(6), 626–637. https://doi.org/10.1521/pedi.2007.21.6.626

Rytwinski, N. K., Fresco, D. M., Heimberg, R. G., Coles, M. E., Liebowitz, M. R., Cissell, S., … Hofmann, S. G. (2009). Screening for social anxiety disorder with the self-report version of the liebowitz social anxiety scale. *Depression and Anxiety, 26*(1), 34–38. https://doi.org/10.1002/da.20503

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., … Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *The Journal of Clinical Psychiatry, 59*(20), 22–33.

Simonsen, S., Eikenæs, I. U. M., Nørgaard, N. L., Normann-Eide, E., Juul, S., & Wilberg, T. (2019). Specialized treatment for patients with severe avoidant personality disorder: Experiences from Scandinavia. *Journal of Contemporary Psychotherapy, 49*(1), 27–38. https://doi.org/10.1007/s10879-018-9395-x

Trivedi, M. H., Rush, A. J., Ibrahim, H. M., Carmody, T. J., Biggs, M. M., Suppes, T., … Kashner, T. M. (2004). The Inventory of Depressive Symptomatology, clinician rating (IDS-C) and self-report (IDS-SR), and the Quick Inventory Depressive Symptomatology, clinician rating (QIDS-C) and self-report (QIDS-SR) in public sector patients with mood disorders: A psychome. *Psychological Medicine, 34*(1), 73–82. https://doi.org/10.1017/S0033291703001107

Weinbrecht, A., Schulze, L., Boettcher, J., & Renneberg, B. (2016). Avoidant personality disorder: A current review. *Current Psychiatry Reports, 18*(3), 1–8. https://doi.org/10.1007/s11920-016-0665-6

Wilberg, T., Karterud, S., Pedersen, G., & Urnes, Ø. (2009). The impact of avoidant personality disorder on psychosocial impairment is substantial. *Nordic Journal of Psychiatry, 63*(5), 390–396. https://doi.org/10.1080/08039480902831322

Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer, 3*(1), 32–35. https://doi.org/10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3

Young, Y. E., Arntz, A., Atkinson, T., Lobbestael, J., Weishaar, M. E., van Vreeswijk, M. F., & Klokman, J. (2007). *The schema mode inventory*. Schema Therapy Institute.