Severity Indices of Personality Problems-Short Form in Old-Age Psychiatry

van Reijswoud, Barbera E; Debast, Inge; Videler, Arjan C; Rossi, Gina; Lobbestael, Jill; Segal, Daniel L; van Alphen, Sebastiaan P J

Published in:
Journal of Personality Assessment

DOI:
10.1080/00223891.2020.1743710

Publication date:
2021

Severity Indices of Personality Problems – Short Form (SIPP-SF) in Old Age Psychiatry: Reliability and Validity

Abstract

The Severity Indices of Personality Problems (SIPP-118 and SIPP-SF; Verheul et al., 2008) is a self-report questionnaire measuring severity of maladaptive personality functioning. Two studies demonstrated the utility of the short form (SIPP-SF) for elderly but research in clinical settings is lacking. Therefore we examined the psychometric properties of the SIPP-SF in N=124 Dutch elderly outpatients (age 60-85, *M*=69.8, *SD*=5.3). The SIPP-SF domains showed good to excellent internal reliability (Cronbach's alphas from α=.75 to .91) and discriminated between participants with and without a personality disorder as assessed with the Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II: First, Gibbon, Spitzer, Williams, & Benjamin, 1997). Convergent validity of the SIPP-SF was examined with instruments for measuring personality pathology in elderly (Informant Personality-questionnaire; HAP; Barendse & Thissen, 2006; Gerontological Personality disorders Scale; GPS; van Alphen, Engelen, Kuin, Hoijtink, & Derksen, 2006). The GPS generally correlated with the SIPP-SF domains with small to large effect sizes. For the HAP only one scale correlated with all SIPP-SF domains. No associations were found between the SIPP-SF and psychiatric symptomatology as measured by the BSI (Brief Symptom Inventory; BSI; Derogatis, 1975). The SIPP-SF appears a promising instrument for assessing maladaptive personality functioning in elderly outpatients.

Introduction

The assessment of personality disorders (PDs) in later life, age 65 years and older, is problematic because the diagnostic criteria of the Diagnostic and Statistical Manual of Mental Disorders (5th ed.; *DSM–5;* American Psychiatric Association, 2013) for PDs are not adequately attuned to older adults (e.g., Rossi, Van den Broeck, Dierckx, Segal, & van Alphen, 2014). As a result, almost a third of the PD symptoms defined in DSM-5 section II express differently later in life (Balsis, Gleason, Woods, & Oltmanns, 2007). This negatively affects the reliability, validity, and utility of the PD construct in older adults and could lead to misdiagnosing PDs in later life (Debast, Rossi & van Alphen, 2017). The prevalence of PDs among older adults in the general population is estimated around 8% (Schuster, Hoertel, Le Strat, Manetti, & Limosin, 2013). For older psychiatric outpatients percentages between 5 and 33% have been reported and the prevalence of (comorbid) PDs in older psychiatric inpatients has been estimated between 7% and 80% (van Alphen, Derksen, Sadavoy, & Rosowsky, 2012). These rates are problematic because PDs in old age are associated with a lower quality of life, more psychiatric comorbidity (Schuster et al., 2013) and more medical treatment (Friedman, Veazie, Chapman, Manning, & Duberstein, 2013). PDs are important to detect since there is accumulating evidence for the efficacy of psychotherapeutic treatment of PDs in adults (Cristea et al., 2017; Dixon-Gordon, Turner & Chapman, 2011; Stoffers et al., 2012). Recently, two studies have supported the efficacy of schema therapy for reducing PD symptoms in older adult. Such findings have increased optimism among those working with patients with PD's in later life (Videler, Rossi, Schoevaars, van der Feltz-Cornelis, & van Alphen, 2014; Videler et al., 2018).

Because the age neutrality of DSM PD criteria can be considered doubtful (Balsis et al., 2007 estimated 29% of criteria display measurement bias), two age-specific

personality measurement instruments have been developed for assessing PDs in geriatric psychiatry (Rossi et al., 2014). The first one is the Hetero-Anamnestische Persoonlijkheidsvragenlijst [Informant Personality Questionnaire] (HAP; Barendse et al., 2006), a Dutch informant questionnaire (HAP; Barendse & Thissen, 2006). The HAP items are based on detecting premorbid maladaptive and dysfunctional personality traits in retrospect. The psychometric qualities of this questionnaire are good (see Barendse, Thissen, Rossi, Oei, & Van Alphen, 2013). The second age-specific instrument is the Dutch Gerontological Personality disorders Scale (GPS; van Alphen, Engelen, Hoijtink, Kuin, & Derksen, 2006). The GPS is a screening instrument to detect PDs in older adults, and thus more generally captures the presence of PD pathology. The GPS consists of a patient and an informant version. Sensitivity and specificity of the GPS in samples of older adults both in psychiatric and general practice populations have shown to be fair (Penders, Duimel-Peeters, Rossi, Metsemakers, & van Alphen, 2015; van Alphen et al., 2006). Although these instruments can be used to screen for personality pathology in later life, they are nor developed nor validated for the use of detecting changes in components of personality functioning, for instance due to treatment, and hence cannot be used for assessing treatment efficacy in terms of personality functioning.

The alternative model for PDs in section III of the DSM-5 differentiates the severity of impaired personality functioning (Criterion A) from the presence of maladaptive traits (Criterion B). Criterion A is defined as impairment in the self and in the capacity for interpersonal functioning, and is dimensionally descripted. These core components of personality dysfunction have been found to discriminate between patients with and without a (traditionally diagnosed) PD (Berghuis, Kamphuis, Verheul, Larstone, & Livesley, 2013). Criterion A describes a view with levels in personality functioning. This view fits with the finding that dimensional models are more useful than categorical

ones for assessing dysfunctional traits and behavior patterns in older people with PDs (van Alphen, Rossi, Segal, & Rosowsky, 2013), since assessment of dimensional traits has been found to be less age biased (Oltmanns & Balsis, 2011). It has to be noted though, that some authors do not support the dimensional model (Shedler et al., 2010) but instead favor the categorical model.

The Severity Indices of Personality Problems (SIPP-118 and SIPP-SF; Verheul et al., 2008) is a promising instrument for assessing DSM-5 Criterion A (Bastiaansen, De Fruyt, Rossi, Schotte, & Hofmans, 2013). The SIPP was developed to differentiate between normal and clinical populations and to measure structural personality changes in treatment studies. It provides a set of five reliable and valid indices of core components of (mal) adaptive personality functioning which seems to be sensitive to change following treatment of patient populations (Johansen et al., 2016). The SIPP-118 has also shown to be a promising instrument for measuring personality pathology in adolescents (Feenstra, Hutsebaut, Verheul, & Busschbach, 2011). One main caveat however, is that the SIPP-118 has not been validated in older adults. Moreover, self-report questionnaires including a large number of items and semi-structured interviews are relatively time-consuming and intensive for older adults (van Alphen et al., 2006). For these reasons, shorter versions of self-report questionnaires are preferable in old age psychiatry. The short form of the SIPP-118 (SIPP-SF) has only half the number of items of the SIPP-118 (i.e., 60 items instead of the original 118), and it has been found to show good psychometric properties in a community sample with an overall mean age of 25 years (Ro & Clark, 2009). Two studies in community dwelling older adults demonstrated the utility of the SIPP short form (SIPP-SF) in later life. The first study evidenced the SIPP-SF`s construct validity in both older and younger adults by demonstrating a factorial structure of five higher order domains (Rossi, Debast & van Alphen, 2017). In older adults personality functioning

measured by the SIPP-SF was more strongly associated with pathological traits of the alternative model for PDs (namely Psychoticism, Disinhibition, Antagonism and Dissocial Behavior) than in younger adults. The second study showed that the SIPP-SF was an age neutral instrument for measuring three out of four domains (Self-Control, Identity Integration and Social Concordance) of personality functioning that closely correspond to criterion A of DSM-5. The SIPP-SF domains Self-Control and Identity Integration capture the self-dimension, and the domains Social Concordance and Relational Functioning capture the interpersonal dimension of the levels of personality functioning of the DSM-5 section III model (Debast, Rossi, & van Alphen, 2018). In both these studies, the SIPP-SF was compared with instruments that measure PDs as described in DSM-5 section III, yet no previous study used a categorical instrument for assessing DSM-section II PDs, like the Structured Clinical Interview for DSM-IV Axis II personality disorders (SCID-II; First, Gibbon, Spitzer, Williams, & Benjamin, 1997).

Since the SIPP is a useful instrument for measuring personality pathology in both adults and adolescents, it is important to investigate the usability of this instrument in older adults because this would offer an opportunity to measure core components of personality across the lifespan. The SIPP-SF seems a promising instrument to measure the core components of personality functioning in older adults and to measure changes in personality functioning due to treatment. Because research on the reliability and validity of the SIPP-SF in older adults is scarce, and completely lacking in clinical samples, the aim of the current study is to investigate the psychometric properties of the SIPP-SF in a clinical sample of older adults, namely psychiatric outpatients in the Netherlands. We will examine: 1) the internal reliability of the five domains of the SIPP-SF, 2) the criterion validity by comparing scores on the SIPP-SF between patients with and those without a SCID-II PD diagnosis and assessing the non-redundant contribution of the scales in

discriminating individuals with and without PDs, 3) associations between severity of personality pathology as measured by the SIPP-SF and DSM section II PDs as measured by the SCID-II (categorical and dimensional), 4) the convergent validity with the SIPP-SF and instruments developed specifically for assessing personality pathology in older adults by relating scores on the SIPP-SF with scores on the HAP and GPS and 5) associations between SIPP-SF scores with psychiatric symptomatology as measured by the Brief Symptom Inventory (BSI; Derogatis, 1975).

Method

Participants

Patients were recruited at two mental health institutes in the Netherlands. Sample one was collected at the old age psychiatry department of Mondriaan, including the Clinical Centre of Excellence for Personality Disorders in Older Adults. Sample two was collected at PersonaCura, an expertise center for PDs in later life of Mental Health Centre Breburg. The Medical Ethics Review Committee Zuyderland- Zuyd (METC-Z) gave approval for the research. Exclusion criteria were severe cognitive problems or dementia (MMSE $\leq$ 24/30), severe psychotic or bipolar problems, intellectual problems (an IQ measured or estimated as below 80) or actual alcohol or drug addiction or use during testing. In sample one, eight patients did not meet the criteria (too young, low intelligence or withdrawing informed consent later on) leaving 99 elderly patients with all sorts of psychiatric problems, whom all had one (or more) DSM-5 classification(s). In sample two there were nine patients that did not meet the criteria (too young) leaving 25 elderly patients who were referred for personality problems to Breburg Mental Health Centre, as can all be seen in Table 1. At the old age psychiatry department of Mondriaan the assessment battery included the SCID II, SIPP-SF, HAP, GPS, and BSI.

At Breburg the same instruments except the BSI were used. Of sample one, four participants did not fill in the BSI. In both samples twelve participants did not fill in the GPS. In 42 cases no informant was available for the informant questionnaire (HAP). In both samples gender was more or less equally divided, the majority was of average education and most participants were married or living together.

About here, insert Table 1

Instruments

The Severity Indices of Personality Problems Short Form (SIPP-SF)

(derived from the SIPP-118; Verheul et al., 2008; available online at https://www.deviersprong.nl/over-de-viersprong/over-de-viersprong-onderzoek/onderzoekslijn-diagnostiek/onderzoekslijn-assessment-en-indicatiestelling/sipp-main-menu/) is a short form of the Severity Indices of Personality Problems, developed in the Netherlands. The SIPP-SF is a dimensional self-report measure for the severity of personality pathology (i.e., severity indices of levels of personality functioning) and specifically developed for treatment outcome research. The 60 items measure the core components of (mal-) adaptive personality functioning with five domains (Self-Control, Identity Integration, Relational Capacities, Responsibility and Social Concordance). Respondents indicate to which extent they agree with statements over the last three months. The response categories range from 1 to 4 and are described as: 'fully disagree', 'partly disagree', 'partly agree' or 'fully agree'. An example of a question from a self-functioning scale is: "Sometimes I get so overwhelmed that I can't control my reactions". An example of a question from an interpersonal-functioning scale is: "I tend to think of myself as a loner". The response

categories range from 1 to 4 and are described as: 'fully disagree', 'partly disagree', 'partly agree' or 'fully agree'. The scores are clustered into five higher order domains, with higher domain scores indicating more adaptive functioning and lower scores indicating more maladaptive personality functioning. Cronbach's alpha values ranged from α=.83 (Social Concordance) to α=.89 (Self-Control and Identity Integration) (Ro & Clark, 2009).

The Dutch version of the SCID II (First et al., 1997; Weertman, Arntz & Kerkhofs, 2000) was used to diagnose DSM-5 section II PDs. This semi-structured interview covers all ten specific DSM-5 PDs as well as the PD Not Otherwise Specified (Other Specified PD in DSM-5). The interview contains 134 open-ended questions and begins with questions about behavior and relationships of the patient. Thereafter, per PD there are questions for all DSM-5 criteria. The SCID II consists of questions like: "When you are out in the public and see people talking, do you often feel that people are talking about you?" Each PD criterion is rated as: 1 'absent or false'; 2 'subthreshold'; or 3 'threshold or true'.

All clinicians conducting the SCID II interviews were extensively trained to ensure the quality of interviewing. The training was provided by the main researcher by giving oral education about the instrument to the individually clinicians in a 60-minute session. After that the clinicians observed two interviews done by an experienced interviewer and the clinicians performed two interviews under supervision before doing the interviews independently. The main researcher was available for consultation for the clinicians during the whole study. The SCID II has shown good inter-rater reliability for the presence or absence of a PD in previous research (Lobbestael, Leurgans & Arntz, 2011).

The Dutch Hetero-Anamnestische Persoonlijkheidsvragenlijst [Informant Personality Questionnaire] (HAP; Barendse et al., 2006) is an informant questionnaire. The HAP items are based on detecting premorbid maladaptive personality traits. The HAP is developed and validated for use in old age psychiatry and nursing homes. The HAP consists of 62 items retrospectively assessed personality traits which are subdivided into 10 scales: Socially Avoidant Behavior, Uncertain Behavior, Vulnerability in Interpersonal Relationships, Somatizing Behavior, Disorderly Behavior, Rigid Behavior, Perfectionistic Behavior, Antagonistic Behavior, Self-satisfied Behavior and Unpredictable and Impulsive Behavior. There are four scales to check the way of answering and rating of the informant. There are three response categories: 'yes', 'more or less' and 'no'. In the instructions of the HAP, a distinction is made between current psychological problems or psychiatric illnesses and the premorbid personality. The psychometric qualities of the HAP are good. The internal consistency of the 10 scales is good (α`s between= .63 and .85, AIC`s between= .23 and .53); the inter-rater and test-retest reliabilities are good to excellent (ICC between .60 and .98); the construct validity, as evidenced through factor analyses, showed the same factor structure in both nursing homes residents and elderly psychiatric patient populations (coefficients of congruence of respectively .99, .96 and .98 for corresponding factors 1, 2, and 3) (Barendse et al., 2013).

In the current study, Cronbach alpha values ranged from α=.44 (unacceptable) to .80 (good), yet the average inter-item correlations (AIC) were additionally calculated to correct for the small numbers of items in the subscales. We considered an AIC above .15 as acceptable (Clark & Watson, 1995). All AIC's were above .15, specifically, ranging from .15 (Rigid Behavior) to .47 (Somatizing Behavior).

The Gerontological Personality disorders Scale (GPS; van Alphen et al., 2006) is a screening instrument to detect PDs in older adults. The GPS consists of a patient and an informant version. Both versions consist of the scales: habitual behavior (GPS-HAB) and biographical information (GPS-BIO). The GPS-HAB scale assesses habitual behavior that reflects the expression of a number of PD features. In the GPS-BIO scale important and recurrent events or decisions in life are linked to the presence or absence of DSM-5 PDs. The internal consistency (Cronbach's alpha) of the two scales range from poor (GPS-HAB α=.57) to acceptable (GPS-BIO α=.77) (van Alphen, 2006). The test-retest reliability of the GPS-HAB and the GPS-BIO subscale items were moderate (Spearman's r=.72) and excellent (Spearman's r=.89), respectively. Sensitivity and specificity of the GPS patient version in an older psychiatric outpatient population was shown to be fair with sensitivity and specificity levels around 70% (van Alphen et al. 2006). In the current study only the GPS patient version was used.

For the total score Cronbach's alpha value was α=.71 (acceptable), for the GPS-HAB scale α=.52 (poor), and GPS-BIO α=.71 (acceptable). In addition, the average inter-item correlation (AIC) was calculated for the subscales to correct for the different numbers of items in the subscales. AIC's were .14 (GPS-HAB), and .20 (GPS-BIO). The AIC of the GPS-BIO scale was above the minimum level of .15 (Clark & Watson, 1995).

The Dutch version of the Brief Symptom Inventory (BSI; Derogatis, 1975; translated by de Beurs, 2006) was used to measure symptomatic distress. The BSI consists of 53 items covering nine symptom dimensions: Somatization, Obsession-Compulsion, Interpersonal Sensitivity, Depression, Anxiety, Hostility, Phobic Anxiety, Paranoid Ideation and Psychoticism. It contains three global indices of distress; Positive Symptom

Distress Index, Positive Symptom Total and Global Severity Index (GSI). Only the GSI was used in the current study. The GSI is a measurement for overall psychological distress reflecting the average score of all item responses. Scores range from 1 to 5, with a higher score indicating a higher level of psychological and emotional distress. This scale was used because it integrates all scales with different kind of symptoms and it is useful for measuring symptomatic distress for patients with divergent pathology as included in our study. Items are questions like: "feeling no interest in things". Respondents rate each item for the past seven days on a five point Likert scale described as: 'not at all' (0), 'a little bit' (1), 'moderately' (2), 'quite a bit' (3), to 'extremely' (4). Reliability of the Dutch version is good (Cronbach's alpha values ranging from α=.71 to .85) and the factorial structure is comparable to that of the original version (De Beurs, 2006).

In the current study, Cronbach's alpha value of the GSI scale was α=.97, which can be considered excellent.

Statistical Analyses

All statistics were performed using SPSS 22.0. Firstly, the internal reliability of the SIPP-SF was analyzed using Cronbach's alpha, AIC and inter-scale Pearson correlations (effect size *r*). Secondly, criterion validity of the SIPP-SF between patients with and those without a PD, as assessed by the SCID-II, was tested with independent sample *t*-tests. Bonferroni correction was used to correct for familywise error rates. The significance level for the analyses was set at p=.01 (.05/5). Effect sizes were computed by Cohen's *d*. In addition, to examine the value of the SIPP-SF scales in predicting the presence or absence of PDs, a binary logistic regression analysis was conducted. Thirdly, associations between severity of personality pathology as measured by the SIPP-SF and by the SCID II (categorical, as in number of diagnosable PDs and dimensional, as in the

number of PD criteria) were analyzed by Pearson correlations (effect size *r).* Fourthly, the convergent validity was evaluated by calculating Pearson correlations (effect size *r)* between the SIPP-SF and both the HAP scales and GPS scores. Finally, the Pearson correlations (effect size *r)* were calculated to evaluate associations between SIPP-SF domain scales and psychiatric symptomatology using the BSI GSI scale.

Results

The SCID II findings showed that 93 participants could be diagnosed with one or more PD(s), while 31 participants were not diagnosed with a PD. As can be seen in Table 2, this means that 75% of the participants were diagnosed with one (or more) PD(s), including Other Specified PD (OSPD). About 40 % had one PD, while 34.6 % had two or more PDs. Subsamples with specific PDs were too small to allow statistics for specific PDs. Therefore the distinction between having a PD or not was used in further analyses.

About here insert Table 2

Besides PDs, several comorbid psychiatric problems were present in the sample. Specifically, the number of non-PD DSM-5 diagnoses varied from none to five. The so-called “V-codes” were not included. In case of substance use or addiction, it either concerned a disorder in remission or a disorder in tobacco use. The most common diagnosis was depressive disorder.

Research Question 1: Internal Reliability of the SIPP-SF

The five domains of the SIPP-SF of participants showed Cronbach's alpha values ranging from α=.75 to .91 with a mean estimated alpha value of α=..82 (see Table 3). The AIC ranged from .20 (Relational capacities) to .46 (Identity integration), with a mean AIC of .31. These values indicate acceptable to excellent reliability of all the domain scales. Intercorrelations between the domains ranged from r =.39 (between Identity integration and Social Concordance) to .75 (between Self-control and Social Concordance), with a median correlation of r=.51 (see Table 4). These positive correlations of medium and large effect sizes confirm the homogeneity of the SIPP-SF.

Replace Table 3 at about here

Replace Table 4 at about here

Research Question 2: The Criterion Validity by Comparing Scores on the SIPP-SF Between Patients with and Those Without a SCID-II PD Diagnosis

The SIPP-SF scores on all domains showed statistical significant negative differences, at the .01 level, between patients with and without a PD. Effect sizes (d) ranged from .59 (Social Concordance) to .86 (Relational Capacities), indicating moderate to large differences, as can be seen in Table 3. Logistic regression was performed to assess the non-redundant impact of the SIPP-SF scales on the likelihood that the patient had a PD. The model for the SIPP-SF contained all five scales as predictor variables. A total of 124 cases were analyzed, and the full model significantly predicted the PD status (omnibus chi-square = 23.346; df = 5; p <.000; Hosmer & Lemeshow Chi-square = 16.805, df = 8; p< .05). The model accounted for between 10.3% and 18.0% of the variance in the PD status, and successfully predicted 93.5% of the patients with PDs.

Almost half (41.9%) of the predictions of patients without PDs were accurate. Overall, 80.6% of the predictions were correct, in comparison to 75.0% in the model only including the constant. Table 5 shows the coefficients, Wald statistic and probability values for each of the predictor variables. These values show that none of the SIPP-SF scales showed non-redundant contributions in predicting the PD status.

Replace Table 5 at about here

Research Question 3: Severity of Personality Pathology

As can be seen in Table 6, the number of DSM-5 PD criteria was negatively associated with SIPP-SF domains (medium effect sizes), thus with higher impairment of personality functioning. This was also seen for the number of present PDs.

Replace Table 6 at about here

Research Question 4: Convergent Validity with Personality Pathology Measures for Older Adults

All domain scores of the SIPP-SF correlated negatively with the GPS-total score and GPS-BIO scale score with small (Social Concordance), medium (Self-Control, Responsibility, Relational Capacities) and large (Identity Integration) effect sizes, as can be seen in Table 6. For the GPS-HAB scale score, the significant associations with Self-Control, Identity Integration and Responsibility were of medium effect size, and with Relational Capacities of small effect size. The association with Social Concordance was non-significant for the GPS-HAB scale score.

The majority of HAP scales did not correlate with the SIPP-SF domains, as can be seen in Table 6. Only one scale of the HAP, Unpredictable and Impulsive Behavior, correlated negatively with all SIPP-SF domains, yet only two correlations showed a medium effect (Self-Control and Social Concordance). The HAP scale Antagonistic Behavior correlated negatively partly with the SIPP-SF at three SIPP-SF domains, showing a medium effect (Self-Control, Relational Capacities end Social Concordance).

Research Question 5: Associations with Psychiatric Symptomatology

The domain scores of the SIPP–SF were not significantly correlated with the scale GSI of the BSI, as can be seen in Table 7. Descriptive statistics for the HAP, GPS and BSI GSI are provided in Table 8.

Replace Table 7 and 8 at about here

Discussion

The aim of this study was to investigate the psychometric properties of the SIPP-SF for assessing components of personality functioning in elderly outpatients. We found an acceptable to excellent internal consistency of all five SIPP-SF domains in this sample. For all SIPP-SF domains a statistical significant difference was found between patients with and those without a PD, as classified with the SCID II. Furthermore, the SIPP-SF showed good criterion validity in predicting a PD. This implies that the SIPP-SF can adequately differentiate between patients with and without a PD, and is related to personality pathology in elderly outpatients. Also, the SIPP-SF scales were associated with the severity of personality pathology, given the negative correlations of the scales

with the number of PDs and with the number of DSM-5 PD criteria. Moreover, all SIPP-SF domains correlated with the GPS-BIO subscale and total score of GPS patient version, a screening instrument for PDs in older adults. The GPS-HAB scale failed to show significant correlations with the subscale Social Concordance. One explanation for this might be that the majority of people with PDs are unaware of the effect their behavior has on others (Klonsky, Oltmanns, & Turkheimer, 2002).

The SIPP-SF did not correlate with most scales of the HAP. An explanation might be that the HAP is filled in by an informant and not by the participant self. Research showed that the informant sees a person and his/her pathology different than the participant sees him/herself. This can be due to a lack of insight that a person has in the effect of ones` behavior on others. Another explanation is unwillingness to disclose on a questionnaire or in an interview (Cruitt & Oltmanns, 2018). The scales from the HAP (ANT and UNP) that showed a negative correlation with medium effect with some, but not all, domains of the SIPP-SF, are scales that belong to the "impulsive and frustration tolerance" profile (Barendse & Thissen, 2006). For this profile it is described that most informants experience this behavior as egocentric and unpleasant and high scores might indicate an anti-social, borderline or passive-aggressive PD (Barendse & Thissen, 2006). This corresponds with the finding that self/informant accordance on PD traits is highest for cluster B pathology, excluding narcissism (Klonsky et al., 2002) and this might explain the correlations on these specific scales (ANT and UNP) with the SIPP-SF. In addition, the HAP uses a lot of behavior descriptions, whereas the SIPP-SF also includes questions about feelings and cognitions about oneself and other persons. Thus, the use of different kinds of questions in the instruments might measure different aspects of personality.

Unlike our expectations, no correlations were found between the BSI GSI and the SIPP-SF domain scores. This means that personality pathology as measured by the SIPP-SF was not associated with overall psychological distress as measured with the BSI GSI. This is in contrast with Feenstra and colleagues (2011) who found meaningful correlations between the SIPP-118 and the GSI scale of the SCL-90 (from which the BSI is derived) in adolescents. The variance of the GSI score showed no evidence for restricted range that would impede the correlation coefficient. One reason for this discrepancy can be that the study of Feenstra et al. (2011) used a more heterogeneous sample comprising, besides a sample of outpatients, high school students and inpatients. In a more heterogeneous sample, psychiatric symptoms are expected to show more variation due to respectively lower and higher symptomatology. Alternatively, the absence of correlations between the SIPP-SF and GSI might indicate that the SIPP-SF is not influenced by having other psychiatric symptomatology in this sample with high rates of PDs. It is possible that, in absence of a PD, the SIPP-SF is more sensitive to psychological distress. However further research with inpatients and a non-patients sample is needed to investigate this. Our findings indicate that the SIPP-SF is overall a good instrument to assess the severity of impaired personality functioning in elderly outpatients. This is an important finding, since the SIPP-SF is a relatively short instrument (60 questions versus 118 in the full SIPP version) and it is known that many elderly patients have difficulties with long instruments (Rossi et al., 2014).

Strengths

This study had some specific strengths. Firstly, this study has a relatively large sample for research with elderly outpatients. Studies in older adults are known for high dropout rates and recruitment difficulties (Provencher, Mortenson, Tanguay-Garneau, &

Dagenais, 2014). Moreover, there was a high rate of PDs in our sample. Our total sample consisted of 124 older patients, of which 75% was diagnosed with a PD. This high rate of PDs can be explained by the fact that in one sample (Breburg) only participants with a suspected PD were included and both participating psychiatric institutions are expertise centers for PDs in the elderly. Therefore, more people with PDs are probably referred to both institutions. This prevalence rate however is within the limits as described by van Alphen et al. (2012). Secondly, it was a strength that, in this study, almost all specific PDs were represented, except the histrionic PD. Nevertheless, the number of specific PDs was too small for further statistical analyses and only the distinction between having a PD or not was used in further analyses, therefore causing no problem for absence or overrepresentation of specific PDs. Thirdly, it was a strength that the SIPP-SF was investigated with instruments that are known to be applicable for the elderly, like the BSI, or specifically designed for age-specific personality assessment in old age, like the GPS and the HAP (Rossi et al., 2014). The final strength was the fact that we compared older patients with a PD and without a PD but with other DSM-5 diagnoses. Both clinical groups thus experienced general psychiatric distress. Due to differentiation and comparison of these groups, we decreased the odds that the SIPP-SF is measuring general psychiatric distress.

Limitations

There are several important limitations that need to be mentioned too. Firstly, our sample did not include participants from the general population where psychiatric symptoms are expected to be lower, or inpatients where psychiatric symptoms are expected to be higher. This might have given a different view on the associations with symptomatic distress. Secondly, some items of the SIPP-SF that refer to work can be

perceived as "not suitable" to elderly people. For instance, in the SIPP-SF one question is: "At work I get easily irritated about other people's ways of doing things". One can argue whether these items should be rewritten to general situations for use in the elderly to capture all components of personality pathology in later life. Moreover, the domain "relational capacities" was found to lack age-neutrality in recent research due to a different degree of expression of the same underlying construct in the older age-group (Debast, Rossi, & van Alphen, 2018). Thirdly, the GPS-HAB scale showed poor internal consistency in the current sample. This poor internal consistency might be explained by the fact that the GPS-HAB scale assesses habitual behaviors and consists of a short list of expressions that relate to behavior linked to various PDs and therefore the items are not necessarily correlated. Fourthly, one benefit of the SIPP is that is designed to capture personality change during treatment. Unfortunately a follow-up measurement was beyond the time-scope of our study. It is therefore a limitation of this study that test-retest reliability was not obtained and it deserves recommendation to do so in future research.

Conclusion

The SIPP-SF seems a promising instrument to be used in geriatric psychiatry for measuring the core components of (mal) adaptive personality pathology. Two main advantages of the SIPP-SF for application in later life are the relatively short form of the instrument, which makes it more suitable for use in older adults, and the close correspondence with the concept of the severity of impaired personality functioning as operationalized in Criterion A of DSM-5 section III. After all, treatment of PDs primarily aims at improving personality functioning. The best way of assessing improvement in personality functioning is by using an instrument that is designed to measure this, like the SIPP-SF. Several research questions remain to be answered. For clinical use of the SIPP-

SF a cut-off score for the domains might be useful to indicate pathological personality functioning . Furthermore, as the SIPP-118 is known to be able to measure the treatment efficacy for PDs in both adults and adolescents (Feenstra, et al., 2011, Verheul et al., 2008), further research on the capability of the SIPP-SF to assess treatment effects in older adults is desirable. Hopefully, our results will stimulate further research on older adults with PDs.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.) Washington DC: Author.

Balsis, S., Gleason, M. E. J., Woods, C. M., & Oltmanns, T. F. (2007). An item response theory analysis of *DSM-IV* personality disorder criteria across younger and older age groups. *Psychology and Aging*, *22*, 171-185.

Barendse, H.P. J., & Thissen, A.J.C. (2006). *Hetero-Anamnestische Persoonlijkheidsvragenlijst (de HAP): handleiding* [Informant Personality Questionnaire (the HAP): manual]. Den Bosch: Barendse & Thissen.

Barendse, H.P.J., Thissen, A.J.C., Rossi, G., Oei, T.I., & Van Alphen, S.P.J. (2013). Psychometric properties of an informant personality questionnaire (the HAP) in a sample of older adults in the Netherlands and Belgium. *Aging & Mental Health, 17* (5), 623-629.

Bastiaansen, L., De Fruyt, F., Rossi, G., Schotte, C., & Hofmans, J. (2013). Personality disorder dysfunction versus traits; structural and conceptual issues. *Personality disorders, 4*, 293-303.

Berghuis, H., Kamphuis, J.H., Verheul, R., Larstone, R., & Livesley, J. (2013). The General Assessment of Personality Disorder (GAPD) as an instrument for assessing the core features of personality disorders. *Clinical Psychology and Psychotherapy, 20*, 544-557.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment,* 7, 309-319.

Cristea, I.A., Gentili, C., Cotet, C.D., Palomba, D., Barbui, C., & Cuijpers, P. (2017). Efficacy of psychotherapies for borderline personality disorder: A systematic review and meta-analysis. *JAMA Psychiatry, 74*, 319-328.

Cruitt, P. J., & Oltmanns, T. F. (2018). Incremental validity of self-and informant report of personality disorders in later life. *Assessment*, *5,* 324-335.

Debast, I., Rossi, G., & Van Alphen, S. P. J. (2017). Construct validity of the DSM-5 Section III maladaptive trait domains in older adults. *Journal of personality disorders*, *31*(5), 671-688.

Debast, I., Rossi, G., Van Alphen, S.P.J. (2018). Age-neutrality of a brief assessment of the Section III Alternative Model for Personality Disorders in older adults. *Assessment, 25*(3), 310-323.

De Beurs, E. (2006) *Brief Symptom Inventory (BSI): Handleiding.* Leiden, The Netherlands: PITS B.V.

Derogatis, L.R. (1975). *Brief Symptom Inventory*. Baltimore, MD: Clinical Psychometric Research.

De Viersprong. *SIPP-SF.* [ca. 2006]. Retrieved on 06-04-2015, from https://www.deviersprong.nl/wp-content/uploads/2016/01/SIPP-SF_English_NEW_2006.pdf

Dixon-Gordon, K.L., Turner, B.J., & Chapman, A.L. (2011). Psychotherapy for personality disorders. *International Review of Psychiatry*, *23,* 282-302.

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

First, M.B., Gibbon, M., Spitzer, R.L., Williams, J.B.W., & Benjamin, L.S. (1997). *Structured Clinical Interview for DSM-IV Axis II personality disorders (SCID II).* Washington, DC: American Psychiatric Press.

Friedman, B., Veazie, P.J., Chapman, B.P., Manning, W.G., & Duberstein, P.R. (2013). Is personality associated with health care use by older adults? *The Milbank*

Quarterly, 91(3), 491-527.

Johansen, M.S, Normann-Eide, E., Normann-Eide, T,. Klungs Yr, O., Kvarstein, E. & Wilberg, T. (2016). Relationships between affect consciousness and personality disorders: A prospective study. *Journal of Personality Disorders, 30(5),* 633-652.

Klonsky. E.D. Oltmanns, T.F., & Turkheimer, E. (2002). Informant-reports of personality disorder: Relation to self-reports and future research directions. *Clinical Psychology: Science and practice, 9*(3), 300-311.

Lobbestael, J., Leurgans, M., & Arntz, A. (2011). Inter-rater reliability of the structured clinical interview for DSM-IV axis I disorders (SCID I) and axis II disorders (SCID II). *Clinical psychology and psychotherapy, 18,* 75-79.

Oltmanns, T.F., & Balsis, S. (2011). Personality disorders in later life: Questions about the measurement, course, and impact of disorders. *Annual Review of Clinical Psychology, 27*(7), 321–349.

Penders, K.A.P., Duimel-Peeters, I.G.P., Rossi, G., Metsemakers, J.F.M., & Van Alphen, S.P.J. (2015). The diagnostic accuracy of the Gerontological Personality disorder Scale (GPS) in Dutch general practice. *Aging & Mental Health, 20*(3), 318-328.

Provencher, V., Mortenson, W.B., Tanguay-Garneau, L, & Dagenais, M. (2014). Challenges and strategies pertaining to recruitment and retention of frail elderly in research studies: A systematic review. *Archives of Gerontology and Geriatrics. 59*(1), 18-24

Ro, E., & Clark, L.E. (2009). Psychosocial functioning in the context of diagnosis: Assessment and theoretical issues. *Psychological Assessment, 21*(3), 313-324.

Rossi, G., Debast, I. & Van Alphen, S.P.J. (2017). Measuring personality functioning in older adults; construct validity of the severity Indices of personality Functioning – Short Form (SIPP-SF). *Aging and mental health. 21*(7), 703-711

Rossi, G., Van den Broeck, J., Dierckx, E., Segal, D.L., & Van Alphen, S.P.J. (2014). Personality assessment among older adults: The value of personality questionnaires unraveled. *Aging & Mental Health, 18*(8), 936-940.

Schuster, J.P., Hoertel, N., Le Strat, Y., Manetti, A. & Limosin, F. (2013). Personality disorders in older adults: Findings from the National epidemiologic Survey on Alcohol and Related Conditions. *American Journal of Geriatric Psychiatry, 21*(8), 757-68.

Shedler, J., Beck, A., Fonagy, P., Gabbard, G.O., Gunderson, J., Kernberg, O., …Westen, D. (2010). Personality disorders in DSM-5. *American Journal of Psychiatry*, *167*(9), 1026-1028

Stoffers, J.M., Völlm, B.A., Rücker, G., Timmer, A., Huband, N., & Lieb, K. (2012). Psychological therapies for people with borderline personality disorder. *Cochrane Database of Systematic Reviews, 15*(8). CD005652.

Van Alphen, S.P.J. (2006). Diagnostiek van persoonlijkheidsstoornissen bij ouderen. Een bijdrage aan de ontwikkeling van een screeningsinstrument [Diagnostics of personality disorders in the elderly. A contribution to the development of a screening instrument]. (Doctoral dissertation). Retrieved on 28-08-2018, from https://repository.ubn.ru.nl/bitstream/handle/2066/27409/27409.pdf

Van Alphen, S.P.J., Derksen, J.J.L. Sadavoy J., & Rosowsky E., (2012). Features and challenges of personality disorders in late life. *Aging & Mental Health,16*(7), 805-10.

Van Alphen, S.P.J., Engelen, G.J.J.A., Kuin, Y., Hoijtink H.J.A., & Derksen, J.J.L. (2006). A preliminary study of the diagnosis accuracy of the Gerontological Personality disorders Scale (GPS). *International Journal of Geriatric Psychiatry, 21*(9), 862-8.

Van Alphen, S.P.J., Rossi, G., Segal, D.L., & Rosowsky, E. (2013). Issues regarding the proposed DSM-5 personality disorders in geriatric psychology and psychiatry. *International Psychogeriatrics, 25*, 1-5.

Verheul, R., Andrea, H., Berghout, C.C., Dolan, C., Busschbach, J.J.V., Van de Kroft, … Fonagy, P. (2008). Severity indices of personality problems (SIPP-118): Development, factor structure, reliability, and validity. *Psychological Assessment, 20*(1), 23-34.

Videler, A.C., Rossi, G., Schoevaars, M., Van der Feltz-Cornelis, C.M., & Van Alphen, S.P.J. (2014). Effects of schema group therapy in older outpatients: A proof of concept study. *International Psychogeriatrics, 26*(10), 1709-17.

Videler, A.C., Van Alphen, S.P.J., van Royen, R.J.J., Van der Feltz-Cornelis, C.M., Rossi, G., & Arntz, A. (2018). Schema therapy for personality disorders in older adults: A multiple-baseline study. *Aging & Mental Health, 22*(6), 738-747.

Weertman, A., Arntz, A., & Kerkhofs, M.L.M. (2000). *Gestructureerd diagnostisch interview voor DSM-IV persoonlijkheidsstoornissen (SCID II)* [Structural and Clinical Interview for DSM-IV personality disorders (SCID II)]. Lisse, The Netherlands: Swets Test Publisher.

Table 1.

Overview of the Samples, With a Total of n=124 Participants.

	Sample	
	1	2
n	99	25
% Female	51.5	60.0
Mean age (*SD*)	70.6 (5,3)	66.9 (4,6)
% Educational level		
Low	20.2	20.0
Average	53.5	56.0
High	26.3	24.0
% Marital status		
Married/living together	68.7	58.3
Single/divorced/widow(er)	31.3	41.7
Assessment battery	SCID II, SIPP-SF, HAP, GPS, BSI	SCID II, SIPP-SF, HAP, GPS

Table 2.

PD Frequencies of all participants.

Personality disorder	*Percentage of cases*
Avoidant	13.7%
Dependent	5.6%
Obsessive-compulsive	16.9%
Paranoid	3.2%
Schizoid	0.8%
Schizotypal	0.8%
Histrionic	0%
Narcissistic	4.0%
Borderline	14.5%
Antisocial	7.3%
OSPD (Other Specified PD)	60.5%
Any PD	75.0%

Note. PD = personality disorder.

Table 3.

Mean Domain Scores and Standard Deviations of Patients Without PD and Patients With PD as measured by the SCID II, independent t-tests and effect size (d)

		Sample					
		Patients without PD (n=31)		*Patients with PD (n=93)*			
Domain	*Cronbach's Alpha*	*M*	*SD*	*M*	*SD*	*t (134)*	*effect size (d)*
Self-control	.86	3.37	0.61	2.84	0.64	4.04*	.84
Identity Integration	.91	3.12	0.75	2.52	0.75	3.92*	.81
Responsibility	.79	3.50	0.46	3.14	0.52	3.40*	.72
Relational capacities	.75	3.06	0.52	2.61	0.53	4.02*	.86
Social concordance	.79	3.35	0.47	3.04	0.55	2.80*	.59

Note. Equal variances were assumed for all domains. * $p<.01$ (using Bonferroni-correction)

Table 4.

Pearson correlations of the SIPP-SF scales.

	Self-control	Identity integration	Responsibility	Relational Capacities	Social Concordance
Self-control	-				
Identity integration	.60**	-			
Responsibility	.50**	.51**	-		
Relational capacities	.55**	.61**	.41**	-	
Social concordance	.75**	.39**	.44**	.53**	-

Note. **Correlations are significant at the .01 level. n=124.

Table 5.

Logistic regression analyses predicting the likelihood of having a PD based on the SIPP-SF scales.

SIPP-SF scales	*B*	SE	Wald	*p*	Exp (*B*)	95% CI
Self-control	-.812	.641	1.605	.205	.444	[.126, 1.560]
Identity Integration	-.246	.441	.312	.576	.782	[.330, 1.854]
Responsibility	-.663	.591	1.259	.262	.516	[.162, 1.640]
Relational capacities	-.832	.573	2.109	.146	.435	[.142, 1.338]
Social concordance	.250	.731	.117	.732	1.285	[.307, 5.381]
Constant	8.093	2.134	14.383	.000	3270.547	

Note. df= 1; Reference category of the dependent: no pd ; Nagelkerke's R-Square=.254.

Table 6.

Pearson Correlations Between The SIPP-SF Domain Scores and the Number of PD Criteria and Number of Diagnosable PDs, GPS scales and HAP dimensions.

	SIPP-SF domains				
	Self-Control	Identity Integration	Responsibility	Relational Capacities	Social Concordance
SCID-II (n=124)					
Number of PD criteria	-.44***	-.30**	-.30**	-.37***	-.44***
Number of diagnosable PDs	-.50***	-.38***	-.38***	-.42***	-.49***
GPS (n=112)					
BIO score	-.35**	-.53**	-.36**	-.40**	-.21*
HAB score	-.38**	-.42**	-.43**	-.29**	-.18
Total score	-.44**	-.60**	-.47**	-.44**	-.24**
HAP (n=82)					
SOC	-.21	-.16	-.15	-.28**	-.29**
UNC	-.03	-.03	-.04	.02	.07
VUL	-.28*	-.11	-.08	-.10	-.27**
SOM	-.13	.08	.17	.15	-.01
DIS	.03	.01	-.19	-.10	-.01
RIG	-.19	-.15	.04	-.10	-.16

PER	-.01	-.01	.18	-.02	-.06
ANT	-.38**	-.15	-.13	-.30**	-.49**
SEL	-.20	-.17	-.10	-.19	-.24*
UNP	-.46**	-.25*	-.23*	-.27**	-.45**

Note. ** Correlation is significant at the $p<.01$ level (2-tailed). *** Correlation is significant at the $p<.001$ level (2-tailed). SOC=socially avoidant behavior, UNC=uncertain behavior, VUL=vulnerability in interpersonal relationships, SOM=somatizing behavior, DIS=disorderly behavior, RIG= rigid behavior, PER=perfectionistic behavior, ANT=antagonistic behavior, SEL=self-satisfied behavior, UNP=unpredictable and impulsive behavior.

Table 7.

Pearson Correlations between SIPP-SF Domain Scores and BSI subscale GSI.

SIPP-SF Domain	*BSI*
Self-Control	-.15
Identity Integration	-.14
Responsibility	.01
Relational Capacities	-.09
Social Concordance	-.07

Note. * Correlation is significant at the $p<.05$ level (2-tailed).** Correlation is significant at the $p<.01$ level (2-tailed). $n=95$.

Table 8.

Descriptive statistics for the HAP, GPS and BSI.

	Scale	Min	Max	Mean	SD	Variance
HAP	SOC	.00	10.00	2.96	2.73	7.42
	UNC	.00	10.00	3.64	2.89	8.33
	VUL	.00	12.00	5.83	3.31	10.95
	SOM	.00	8.00	2.46	2.49	6.18
	DIS	.00	8.00	2.19	2.41	5.82
	RIG	.00	8.00	3.77	2.08	4.32
	PER	.00	8.00	4.24	2.40	5.77
	ANT	.00	18.00	6.09	4.04	16.34
	SEL	.00	10.00	2.27	2.07	4.29
	UNP	.00	12.00	4.17	3.17	10.08
GPS	BIO	.00	8.00	3.83	2.15	4.77
	HAB	.00	7.00	3.00	1.68	2.81
	Total score	.00	14.00	6.83	3.15	9.91
BSI	GSI	.02	3.43	1.30	.83	.68

Note. SOC=socially avoidant behavior, UNC=uncertain behavior, VUL=vulnerability in interpersonal relationships, SOM=somatizing behavior, DIS=disorderly behavior, RIG= rigid behavior, PER=perfectionistic behavior, ANT=antagonistic behavior, SEL=self-satisfied behavior, UNP=unpredictable and impulsive behavior.