

Personality disorders in older adults

Penders, Krystle A P; Rossi, Gina; Debast, Inge; Segal, Daniel L; Peeters, Inge G P; Metsemakers, Job F M; van Alphen, Sebastiaan P J

Published in:
Personality and Mental Health

DOI:
[10.1002/pmh.1592](https://doi.org/10.1002/pmh.1592)

Publication date:
2024

License:
Unspecified

Document Version:
Accepted author manuscript

[Link to publication](#)

Citation for published version (APA):

Penders, K. A. P., Rossi, G., Debast, I., Segal, D. L., Peeters, I. G. P., Metsemakers, J. F. M., & van Alphen, S. P. J. (2024). Personality disorders in older adults: Differences in self-informant ratings. *Personality and Mental Health*, 18(1), 32-42. <https://doi.org/10.1002/pmh.1592>

Copyright

No part of this publication may be reproduced or transmitted in any form, without the prior written permission of the author(s) or other rights holders to whom publication rights have been transferred, unless permitted by a license attached to the publication (a Creative Commons license or other), or unless exceptions to copyright law apply.

Take down policy

If you believe that this document infringes your copyright or other rights, please contact openaccess@vub.be, with details of the nature of the infringement. We will investigate the claim and if justified, we will take the appropriate steps.

Personality Disorders in Older Adults: Differences in Self-Informant Ratings

Krystle A.P. Penders, PhD^{ab}, Gina Rossi, PhD^a, Inge Debast, PhD^a, Daniel L. Segal, PhD^c, Inge G.P. Peeters, PhD^{de}, Job F.M. Metsemakers, PhD^d & Sebastiaan P.J. van Alphen, PhD^{afg}

^a Faculty of Psychology & Educational Sciences, Department of Clinical & Lifespan Psychology, Vrije Universiteit Brussel (VUB), Brussels, Belgium

^b Department of Treatment and Guidance, Envida, Maastricht, The Netherlands

^c Department of Psychology, University of Colorado, Colorado Springs, Colorado, USA

^d Department of Family Medicine, School CAPHRI, Care and Public Health Research Institute, Maastricht University (UM), Maastricht, The Netherlands

^e Academy for Postgraduate Medical Training, Maastricht University Medical Centre + (MUMC+), Maastricht, The Netherlands

^f Department of Old Age Psychiatry, Mondriaan Hospital, Heerlen-Maastricht, The Netherlands

^g School of Social and Behavioral Sciences, Department of Medical and Clinical Psychology, Tilburg University, Tilburg, The Netherlands

Funding acknowledgement

This work was supported by the National Care for the Elderly Program from The Netherlands Organization for Health Research and Development (ZonMw) under Grant 311070201.

Correspondence

Correspondence concerning this article should be addressed to Krystle Penders, E-mail: krystle.penders@vub.be, Telephone number: +31610896125.

Declaration of Conflicting Interests

The authors declare that there is no conflict of interest.

SELF- AND INFORMANT-REPORTED PERSONALITY PATHOLOGY

Abstract

Previous research on self-informant reports in assessing personality disorders (PDs) has been mainly focused on adults, leaving older adults understudied. We examined self-informant agreement in PD screening among older adults (≥ 60 years) using the Gerontological Personality disorders Scale (GPS). Potential differences such as who reports more personality pathology on a PD screener (i.e. GPS), item accessibility, and the effect of relational aspects were studied as well. Data of 326 older adult-informant dyads, of which the older adults were sampled from five general practices in the Netherlands, were used. Results indicate that self-informant agreement ranged from $r = .26-.73$, with lower concordance on the GPS-subscale measuring intrapersonal aspects of personality pathology. Informants were more sensitive to habitual pathological personality features than older adults. Two GPS-items showed differential item functioning across self- and informant-report. Of relational aspects, only congeniality affected the GPS-iv scores; lower ratings on congeniality were associated with higher GPS-iv scores (i.e. higher reporting of personality problems).

What this paper adds:

- It underscores the importance of including both self- and informant-reports in assessing PDs in older adults as they provide complementary perspectives.

Applications of study findings:

- Self-report is especially useful when addressing intrapsychic processes such as thoughts and feelings.
- Informant-report is most appropriate for assessing highly observable and interpersonal behavior.

SELF- AND INFORMANT-REPORTED PERSONALITY PATHOLOGY

Keywords: personality disorders; older adults; self-report; informant-report; Gerontological

Personality disorders Scale (GPS)

Indication: Research article



Personality Disorders in Older Adults: Differences in Self-Informant Ratings

Introduction

Although the value of combining self- and informant-report in the assessment of personality disorders (PDs) is widely accepted (e.g. Oltmanns et al., 2014), this approach still invokes several questions. These mainly include issues concerning self-informant agreement such as which source indicates more personality pathology, what information is more accessible to which source, and whether the nature of the relationship and the personality pathology in question affect the agreement (Klonsky et al., 2002, Cooper et al., 2012, Cooper et al., 2014, Vazire and Carlson, 2011, Watson et al., 2000, Leising et al., 2010, Griens et al., 2002, Santor et al., 1997, Mosterman and Hendriks, 2011, Lieberman et al., 2016). To date, only a few self-informant studies have included midlife or older adults (e.g. Oltmanns et al., 2014, Balsis et al., 2018) whereas most PD self-informant agreement research has focused on adult samples (19-65 years). The current study therefore explores all the above-mentioned challenges to assessment in a large sample of Dutch older adults (≥ 60 years).

Until recently, self-informant agreement for measures of PDs, no matter the relationship of the informant, appeared to be low to moderate (Klonsky et al., 2002, Cooper et al., 2012, Samuel, 2015, Balsis et al., 2018). Recent meta-analyses hold a somewhat more optimistic view, with an average agreement level of $r = .37$, in which this improvement of agreement seems to be attributable to assessing PD traits dimensionally (with several items per facet scale) as opposed to dichotomous categories (often relying on one item per DSM criterion; Oltmanns and Oltmanns, 2021).

Self- and informant-report appear to have incremental validity above each other. Overall, informant-report seems to reveal more PD symptoms compared to self-report (Oltmanns et al., 2014, Cooper et al., 2012, Cooper et al., 2014) even at lower levels of PD

SELF- AND INFORMANT-REPORTED PERSONALITY PATHOLOGY

severity (Balsis et al., 2018). The self-informant discrepancy might be explained by the lack of both inter- and intrapersonal insight as well as by egosyntonicity, which is thought to be inherent to personality pathology (Klonsky et al., 2002), including among older adults (Segal, Coolidge, & Rosowsky, 2006). Low agreement also seems to be partly attributable to the nature of the precise information that is being assessed. Specifically, low observable aspects (i.e. internal processes, such as feelings and thoughts) and low evaluative traits (i.e. low valence of either desirability or undesirability, such as anxiety or self-esteem) are believed to be best retrieved with self-report, whereas informant-reports are preferred when traits are high observable and high evaluative (e.g. creativity or being intelligent; Vazire and Carlson, 2011). Besides trait observability and evaluativeness, multiple other factors seem to play a role in self-informant agreement in PDs as well, such as acquaintanceship and liking (Watson et al., 2000, Leising et al., 2010), although recent moderation analyses lacked significant findings (Oltmanns and Oltmanns, 2021).

The extant PD self-informant literature (Cooper et al., 2012, Klonsky et al., 2002, Oltmanns and Oltmanns, 2021) reveals that there are many methodological limitations which should be accounted for in future research, for example, not reporting levels of agreement and not using the preferred method for assessing agreement. As Klonsky and colleagues (2002) point out, many researchers rely on Pearson correlations when addressing self-informant agreement, although intraclass correlations (ICC) are possibly a more appropriate measure because Pearson correlations do not classify mean differences between raters as error. Other methodological restrictions include the methods of selecting an informant, method variance (i.e., data were gathered from each source using different instruments), and relatively small sample sizes. To address some of these limitations, the current study will explore self-informant agreement using ICC and will overcome method variance by using the same instrument for both self and informant perspective while using a large sample size.

SELF- AND INFORMANT-REPORTED PERSONALITY PATHOLOGY

As previously mentioned, another major limitation is that most prior self-informant PD studies focus on adult populations, leaving older adults (> 65 years) understudied. This latter limitation must be seen in a broader context. Generally, older adults still seem to be underrepresented in PD assessment, which is rather curious, as PDs are quite prevalent (Penders et al., 2020) and impactful in older adults. Prevalence rates range from 14.5% in community-dwelling older adults (Reynolds et al., 2015), to 10-33% among psychiatric outpatients (Widiger and Seidlitz, 2002), and up to even 80% among older adult inpatients with chronic and persistent mental disorders (Silberman et al., 1997).

Although clinical and scientific attention in PD research is shifting toward older adults, to date there are only few validated instruments to assess PDs in older adults (Rossi et al., 2014, Penders et al., 2020). Most available instruments were validated in younger age-groups and also did not take into account the specific context of aging. Instruments specifically designed for older adults are also scarce. Although the American Psychological Association (2014) recommends the use of informant information in their general guidelines for working with older adults, informant-report instruments for PDs in older adults have rarely been studied. Including an informant perspective especially among older adults can be of added value, to circumvent the possible deleterious effects of cognitive impairment or sensory deficits which may complicate self-report. Likewise, there is a lack of age-specific instruments that have both a self- and informant version. Only one instrument, the Gerontological Personality disorders Scale (GPS; Van Alphen et al., 2006), includes both a patient version (GPS-pv) and an informant version (GPS-iv). This age-specific screening instrument for PDs is validated for use among outpatient geriatric populations (Van Alphen et al., 2006, Meeuwissen-Van Pol et al., 2020) and among the general older adult population (Penders, Rossi, Metsemakers, Duimel-Peeters & van Alphen, 2016). The GPS also was found to be feasible and acceptable in general practice (Penders et al., 2018).

SELF- AND INFORMANT-REPORTED PERSONALITY PATHOLOGY

To address this gap in late life PD self-informant agreement literature, this study explored agreement issues using the GPS in an exclusively older adult sample. The main research question was:

- What is the agreement between self- and informant-reported ratings?

Various possible differences in agreement were examined, resulting in the following sub-questions:

1. Which perspective, self or informant, reports more personality pathology based on the GPS as PD screener (as reflected in higher GPS scores)?
2. Are the items of the GPS equally accessible for both sources?
3. What is the effect of relational aspects on self-informant agreement?

Based on the literature on self-informant agreement in adults, we hypothesized the following:

(1) self-informant agreement will be moderate in PD assessment in older adults, given a recent meta-analysis mainly focusing on younger adult populations that found an overall self–other agreement on PD symptoms and traits of $r = .37$ (Oltmanns & Oltmanns, 2021); (2) informant-report will reveal higher GPS scores than self-report, revealing more personality pathology as self-ratings on PD measures tend to be lower than informant ratings (Carnovale, Carlson, Quilty & Bagby, 2019); (3) items addressing internal processes will be less accessible for informants, similar to studies in younger adult samples (Vazire, 2010, Oltmanns & Oltmanns, 2021); (4) GPS-iv scores will be higher when the level of congeniality and the quality of the relationship is lower, in line with Leising and colleagues' (2010) 'letter of recommendation effect'.

Methods

Participants

Prior to the current investigation, older adults and informants contributed data to Penders et al. (2016) research on the diagnostic accuracy of the GPS as well as on the feasibility and

acceptability of the GPS (Penders et al., 2018). The research questions and data presented here have not been published elsewhere.

A total of 348 Dutch self-informant dyads were sampled from five general practices in the Netherlands of which 326 dyads were retained for analyses (flow chart available in supplemental material). Older adults, who provided self-report data, were eligible if they were Dutch-speaking, aged 60 years or older, and having a close informant who was willing to participate. Older adults with severe psychiatric disorders as per DSM-5 (APA, 2013) other than a PD (as judged by the general practitioner), having significant cognitive dysfunction (defined as MMSE score ≤ 24 ; Folstein et al. (1975)), and having a limited attention span, intellectual disabilities, or a life expectancy of less than three months were excluded. For informants, no other inclusion or exclusion criteria were formulated other than being Dutch-speaking and willing to participate.

Measures

The Gerontological Personality disorders Scale (GPS). The GPS is a 16-item age-specific screening instrument designed to detect general PD presence in older adult inpatients (Van Alphen et al., 2006). The GPS can also be applied in general practice (Penders et al., 2016) in which the average item correlation of the GPS-pv and GPS -iv were respectively .12 (HAB), .16 (BIO), .10 (total) and .16 (HAB), .15 (BIO), and .12 (total). The test-retest reliability of the GPS-pv was ($r_s = .56$ (HAB), $r_s = .67$ (BIO), $r_s = .66$ (total)) and of the GPS-iv ($r_s = .52$ (HAB), $r_s = .65$ (BIO), $r_s = .68$ (total)). The sensitivity and specificity of the GPS-pv was respectively .83 and .27 with a cutoff score of ≥ 1 . After raising the cutoff score to ≥ 2 , sensitivity dropped to .59 whereas specificity rose to .57. For the informant version, a cutoff score of ≥ 3 maximizes sensitivity (.78) and specificity (.65). Additionally, in a geriatric medicine setting (Meeuwissen-Van Pol et al., 2020), the average item correlation of the GPS-pv total was .14, test-retest reliability was $r_s = 0.68$, and sensitivity and specificity

SELF- AND INFORMANT-REPORTED PERSONALITY PATHOLOGY

were 0.91 and 0.67, respectively. The GPS consists of a self-report version (GPS-pv) and an informant-report version (GPS-iv) which only differs in how the statements are worded, respectively the first person or the third person. Items on the GPS were based on both literature and expert opinions on PDs in older adults (Van Alphen et al., 2003, Van Alphen et al., 2004), as well as on PD case studies (Van Alphen et al., 2001), and the general diagnostic PD criteria of the DSM-IV (American Psychiatric Association, 2000). The 16 items are rated as being either *true* or *false*. The total GPS score can thus vary from 0 to 16. The GPS addresses both habitual behavior (HAB) and biographical information (BIO), by assessing enduring unstable behavior as an essential aspect of PD pathology. HAB (7 items) emphasizes habitual behavior in terms of expression of various PD features and accounts for gerontological themes such as cognitive decline, adapting to physical illness, and loss of relationships. The HAB scale ranges from 0 to 7. BIO (9 items) focuses on previous and current events, behavior, or decisions that are conceptually linked to the presence of PDs, such as suicidality or multiple psychiatric treatments in the past. The BIO scale ranges from 0 to 9. Internal consistency of GPS scale scores was examined using average inter-item correlations (AIC) as it is considered to be superior to Cronbach's α because it measures internal consistency independently of the number of scale items. An AIC above .15 is acceptable (Clark & Watson, 1995). In the present sample, the AIC values were .14 (HAB), .20 (BIO), and .13 (total) for the GPS-pv and .17 (HAB), .19 (BIO), and .15 (total) for the GPS-iv. The AIC of the GPS-iv and the BIO scale of the GPS-pv were above the minimum level of .15 (Clark and Watson, 1995).

Relationship Characteristics. Relational information was collected by asking all participants about the nature of the relationship, as well as by grading both the quality of the relationship and the level of congeniality (1-10).

Procedure

Ethics approval was granted from the Medical Ethical Review Commission of the Maastricht University Medical Center + (MUMC+), the Netherlands (approval no. MEC 09-4-060). Data were collected from 2009 to 2012 during home visits to the participants. Prior to data collection, the older adults, informants, and researchers signed the consent form. Demographic information and relationship characteristics with the informant were collected in an interview with the older adult. Next, the GPS-pv was administered. During this same period, the informants were asked to fill out a form with the same demographic and relational information as well as the GPS-iv in a separate room to prevent response bias.

Statistical analyses

All statistics were performed using SPSS 23.0 (IBM Corp., 2015) or IRTPRO (Cai et al., 2011). All GPS-pv and GPS-iv scales were relatively normally distributed, with none of the skewness or kurtosis values exceeding the recommended cut-offs of greater than 2.0 and 7.0 respectively (Curran et al., 1996), with the exception of GPS-pv HAB, of which the skewness was 2.228 (std. error = .135). As multiple comparisons (i.e. in total 46) were done, Bonferroni corrected p-values (Napierala, 2012) were calculated to control for familywise error (p .0011). For a more detailed overview of the statistical analyses, see appendix 1 published as supplementary material online. The main research question, regarding agreement between self- and informant-reports was examined on both the item and the overall level by exploring the frequencies of item endorsement, chi-square for independence, and by using intraclass correlations (ICC) as they are believed to be a more accurate measure of interrater agreement than Pearson's correlations (Klonsky et al., 2002). ICC values below 0.5 indicate poor reliability, between 0.5 and 0.75 indicate moderate reliability, between 0.75 and 0.9 indicate good reliability, and values above 0.9 indicate excellent reliability (Koo & Li, 2016).

SELF- AND INFORMANT-REPORTED PERSONALITY PATHOLOGY

To determine who reports more PD symptoms, sub question 1, descriptive statistics were calculated (mean, SD) as well as by computing between-group differences between self- and informant-reports on the GPS using independent-samples t-test or Mann-Whitney *U* test. Cohen's *d* was used as a measure of effect size (Cohen, 1988). Sub question 2, whether GPS scales have different item properties among the perspectives of selves and informants, was examined with Differential Item Functioning (DIF) analyses, applying the Item Response Theory (IRT) framework. The potential influence of relational aspects on the GPS-iv scores, sub question 3, was first assessed by comparing mean scores of the informants on the GPS and its subscales per level of the various relational aspects. Next, ANOVAs with post hoc comparisons were computed to assess whether the GPS-iv HAB and total score differed between groups based on the level of the relational aspects. As not all assumptions were met for GPS-iv BIO scale, Kruskal-Wallis tests were used.

Results

Approximately half (46%) of the older adults were male ($n = 149$). The average age was 69.8 years (range = 60-91 years, $SD = 7.4$ years). Over 42% of the informants were male ($n = 138$). The average age was 64.7 years (range = 25-89 years, $SD = 11.5$ years). Of the older adults, almost all (98.8%) were living independently, with the remaining living institutionalized or otherwise. Most (77.3%) were married, the remaining were widowed (12.0%), single (4.0%), or had another relationship status. The older adults mostly choose their partner (78.5%) to be their informant, followed by family members (16.7%) or friends (4.8%). The duration of the relationship ranged from 1-80 years, with a mean of 43.6 years. The older adults rated both the quality of the relationship with the informant and the level of congeniality with an 8.6/10, whereas informants rated both aspects with an 8.7/10. A more detailed overview of the demographic information can be found in the supplemental material.

What is the agreement between self- and informant-reported ratings?

SELF- AND INFORMANT-REPORTED PERSONALITY PATHOLOGY

As can be seen in Table 1, the GPS items HAB 2 and HAB 6 were significantly more frequently endorsed in the informant-report version. Regarding the GPS, results indicated that the self-informant agreement was high for the GPS-BIO scale (ICC =.73) and moderate for the GPS-HAB scale (ICC =.37). The agreement with regard to the GPS total score was also high (ICC =.61).

1. Which perspective, self or informant, reports more personality pathology (as reflected by higher GPS scores)?

Table 2 shows the mean scores on the HAB scale and total score on the GPS-pv and GPS-iv, and the median of the BIO scale. Analyses (T-tests) revealed that the mean score of the informants on the GPS HAB was significantly higher than those of the older adults, with a small effect size. The scores on the GPS BIO and total scores did not differ significantly when the Bonferroni correction was applied.

2. Are the items of the GPS equally accessible for both sources?

To examine the possible presence of DIF in the HAB and BIO subscales of the GPS, a total of 16 (omnibus) separate analyses were conducted evaluating all items one by one. According to Bonferonni corrected omnibus tests, two items were found to be significant. Bonferonni corrected follow-up tests specified that HAB2 ('worry about health'; Omnibus X^2 difference= 34.5, $a^{\text{self-report}}$ 6.36, $b^{\text{self-report}}$.80, $a^{\text{informant-report}}$ 1.07, $b^{\text{informant-report}}$.88) and HAB5 ('afraid of losing others'; Omnibus X^2 difference= 16.4, $a^{\text{self-report}}$ 1.54, $b^{\text{self-report}}$.07, $a^{\text{informant-report}}$.73, $b^{\text{informant-report}}$.87) showed non-uniform DIF; these items were more related to the underlying dimension in the self-report than in informant-reports. An overview of all the scores can be found in the supplemental material.

3. What is the effect of relational aspects on the agreement?

Of the relational aspects, only the level of congeniality might have affected the self-informant agreement. ANOVA and Kruskal-Wallis tests with Bonferroni corrected post-hoc

comparisons revealed that those informants who rated the likability of the older adult with a grade of 6.5 or less reported significantly higher GPS-iv scores on the BIO scale and total scale than those who were more favorable about the older adult's congeniality, with small effect sizes (Table 3).

Discussion

As empirical studies on self-informant-report in the assessment of PDs have mainly focused on (young) adult samples thereby leaving older adults understudied, the current study examined self-informant agreement in PDs and potential differences in ratings in a large older adult sample using the GPS. The self-informant agreement on the GPS in the current data from general practices ranges from moderate (HAB) to good (BIO) with good agreement on the total GPS, which mainly mirrors earlier findings of Van Alphen et al. (2005) who reported on the correspondence of the GPS in an older adult outpatient population. Our findings also partly extend those of Oltmanns and Oltmanns (2021) revealing overall moderate levels of self-informant agreement in adults (mean age = 18.2-59.6 years).

More specifically, this study supports the findings of previous research with samples aged 55-65 years old (Oltmanns et al., 2014, Cooper et al., 2012, Balsis et al., 2018) by showing that informants reported more personality pathology than did the older adults themselves. These results might suggest that the self-report could be biased as a result of limited self-awareness or recognition that one has a (psychiatric) problem, distorted self-perceptions (Ganellen, 2007), or possibly due to the reluctance to disclose problems (Cape and McCulloch, 1999).

The contrast in self-informant agreement between the GPS subscales might also be partly due to differential functioning of the items. The IRT analyses revealed that two GPS HAB items displayed differential item functioning. More specifically, items HAB2 ('worry about health') and HAB5 ('afraid of losing others') were more discriminative in self- over

SELF- AND INFORMANT-REPORTED PERSONALITY PATHOLOGY

informant-reported underlying problems of habitual behavior. Older adults and informants with the same underlying level of reported personality pathology problems in habitual behavior appear to have unequal accessibility to these items, which can result in biased raw scores on the HAB scale. Hence, these differences in endorsement probabilities are in line with our findings regarding lower self-informant agreement on HAB items compared to BIO items. Inspection of the HAB items that show differential item functioning indicates that these items mainly tap into intrapersonal information, whereas the BIO items appear to be more interpersonal and observable. Intrapersonal aspects have a more covert nature and are less observable for informants. This is congruent with previous research addressing both personality and personality pathology which pointed out that information that is low observable and low evaluative are best accessible to the self (Vazire, 2010, Vazire and Carlson, 2011, Oltmanns and Oltmanns, 2021). As informants appear to have a lower probability of answering affirmative on various GPS items, the informant endorsement rates, especially for HAB, are likely to be underestimated leading to heightened levels of self-informant agreement. Hence, the true self-informant agreement might actually be even lower.

Relational aspects also seem to have an influence on informant ratings. The more informants liked their older adult, the lower their GPS scores were, which is similar to the findings from Leising et al. (2010), yet only small effects were found. Furthermore, congeniality may have affected the magnitude of GPS scores, given that it might be a reflection of a romantic bond, as most older adults chose their spouse to be their informant. The informants' scores were still higher than the self-reported scores, which might have affected the degree of self-informant agreement.

Strengths of the current study are that it is one of the first to address PD self-informant agreement in older adults as well as to examine several explanatory variables. By taking these variables into account, the study is able to obtain a more complete and holistic picture of the

factors that may play a role in self-informant agreement in the assessment of PDs. Also, in comparison to earlier research (Klonsky et al., 2002) the sample size in the current study was large.

Despite these strengths, some limitations should also be mentioned. Although we accounted for the small number of GPS items by using AIC to assess internal consistency (Clark and Watson, 1995), most AIC values barely exceeded the minimum level of consistency. This is likely to be a consequence of the intention of the GPS, while adhering to experts recommendations (Van Alphen et al, 2003), on being a very short, feasible, and relatively quick screening instrument on the one hand, and the abstract and multifaceted nature of the PD construct on the other hand. This is especially true for the HAB scale, as these items relate to various traits relevant for PDs, which are not necessarily highly correlated with each other. The low internal consistency is likely to have affected the self-informant agreement, as it attenuates the maximum correlation that would be possible between two scores (Furr and Bacharach, 2008).

The older adults in the current study nominated their informants, and they generally chose those persons with whom they had a close and good relationship and who they thought would be willing to do them a favor to participate. Previous research (i.e. Leising et al., 2010) pointed out this could lead to unintended biases and we concur that it would therefore be worthwhile to use other inclusion strategies in future studies in which older adults do not choose the informants. However, the current selection procedure was mainly based on practical considerations, as it is known that recruitment of older adults in research is often difficult (Gregson et al., 1997) and including both older adults and non-self-selected informants might be even more challenging.

As the present data set was collected 10 years ago, and the findings might not fully reflect current characteristics and trends within the older population, further replication is

needed. The data set is however based on the DSM-5 definition of PDs which is still the leading classification system (APA, 2013). Furthermore, in future research the diagnostic accuracy of the GPS should be further addressed using an reference criterion that isn't solely informant-based (as the HAP is), as this might have negatively impacted the results in favoring the informant's perspective. Potential reference instruments could include the Assessment of the DSM-IV Personality Disorders (ADP-IV; Schotte and de Doncker, 1994), the Severity Indices of Personality Problems-Short Form (SIPP-SF; Verheul et al, 2008) or the Levels of Personality Functioning Scale- Brief Form 2.0 (LPFS-BF 2.0; Weekers, Hutsebaut, and Kamphuis, 2018) as these measures have proved to be suitable for older adults (Rossi, Debast, and van Alphen, 2017, Debast, Rossi, and van Alphen, 2018, Stone, Segal, and Noel, 2021). As the GPS merely addresses the presence or absence of the general criteria of PDs and does not specify PD type, inclusion of a maladaptive trait-based instrument, such as the Personality Inventory for DSM-5 (PID-5; Markon et al., 2013) might be considered as well. Moreover, in an almost completed study, the GPS' nomological network was investigated by examining how the GPS relates to personality, psychopathology, as well as to the Hierarchical Taxonomy of Psychopathology (HiTOP; Kotov et al., 2017), an empirical-driven classification system.

Since this study is one of the first to address self-informant PD agreement in older adults, it is based on exploratory analyses. As these are valuable initial steps in understanding research questions in a rather new field, future research should replicate the findings by preregistering the study on the Open Science Framework and testing research questions using more confirmatory techniques.

To learn more about PD self-informant-report and agreement in older adults, further research should extend the current findings by performing multiple cross-validation studies (in other countries). It would also be relevant to explore relational aspects more thoroughly,

SELF- AND INFORMANT-REPORTED PERSONALITY PATHOLOGY

by using other than older adult-nominated selection procedures for including informants, as well as by being more attentive to include larger numbers of various relationship types to permit more in-depth analysis. We hope that the current study stimulates more inquiry into self-informant-report issues in PD assessment in older adults.

Description of authors' roles:

K. Penders co-designed the study and statistical design, collected the data, carried out the statistical analyses (except for IRT) and wrote the manuscript with input of all authors. I. Debast was responsible for carrying out the IRT analysis and writing the results from this analysis. G. Rossi co-designed the statistical design, verified the statistical analyses, provided critical feedback and help shape the research. D. Segal carefully edited the manuscript for language issues and additionally contributed by providing his view on the manuscript. I. Peeters discussed the results and commented on the manuscript. J. Metsemakers discussed the results and provided critical feedback and helped shape the research, and manuscript. B. van Alphen conceived the main conceptual idea and co-designed the study, and critically reviewed the manuscript.

References

- American Psychiatric Association** (2000). *Diagnostic and statistical manual of mental disorders*, Washington DC, Author.
- American Psychological Association** (2014). Guidelines for psychological practice with older adults. *American Psychologist*, 69, 34-65.
- APA** (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*, Arlington, VA, American Psychiatric Publishing.
- Balsis, S., Loehle-Conger, E., Busch, A. J., Ungredda, T. & Oltmanns, T. F.** (2018). Self and informant report across the borderline personality disorders spectrum. *Personality Disorders: Theory, Research, and Treatment*, 9, 429-436.
- Cai, L., Du Toit, S. & Thissen, D.** (2011). Irtpro: Flexible, multidimensional, multiple categorical irt modeling. Chicago: IL: Scientific Software International.
- Cape, J. & Mcculloch, Y.** (1999). Patients' reasons for not presenting emotional problems in general practice consultations. *British Journal of General Practice*, 49, 875-879.
- Carnovale, M., Carlson, E. N., Quilty, L. C., & Bagby, R. M.** (2019). Discrepancies in self- and informant-reports of personality pathology: Examining the DSM-5 Section III Trait Model. *Personality Disorders: Theory, Research, and Treatment*, 10(5), 456-467
- Clark, L. A. & Watson, D.** (1995). Constructing validity. Basic issues in objective scale development. *Psychological Assessment*, 7, 309-319.
- Cohen, J.** (1988). *Statistical power analysis for the behavioral sciences*
Hillsdale: NJ: Lawrence Erlbaum.
- Cooper, L. D., Balsis, S. M. A. & Oltmanns, T. F.** (2012). Self- and informant-reported perspectives on symptoms of narcissistic personality disorder. *Personality Disorders: Theory, Research, and Treatment*, 3, 140-154.

- Cooper, L. D., Balsis, S. M. A. & Oltmanns, T. F.** (2014). A longitudinal analysis of personality disorder dimensions and personality traits in a community sample of older adults: Perspectives from selves and informants. *Journal of Personality Disorders*, 28, 151-165.
- Curran, P. J., West, S. G. & Finch, J. F.** (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, 1, 16-29.
- 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*, 1-14
- Folstein, M. F., Folstein, S. E. & Mchugh, P. R.** (1975). Mini-mental state. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189-198.
- Furr, R. M. & Bacharach, V. R.** (2008). *Psychometrics: An introduction*, Thousand Oaks, CA, Sage Publications.
- Ganellen, R. J.** (2007). Assessing normal and abnormal personality functioning: Strengths and weaknesses of self-report, observer, and performance-based methods. *Journal of Personality Assessment*, 89, 30-40.
- Gregson, B. A., Smith, M., Lecouturier, J., Rousseau, N., Rodgers, H. & Bond, J.** (1997). Issues of recruitment and maintaining high response rates in a longitudinal study of older hospital patients in England-pathways to care study. *Journal of Epidemiology and Community Health*, 51, 541-548.
- Griens, A. M., Jonker, K., Spinhoven, P. & Blom, M. B.** (2002). The influence of depressive state features on trait measurement. *Journal of Affective Disorders*, 70, 95-99.

SELF- AND INFORMANT-REPORTED PERSONALITY PATHOLOGY

- IBM Corp.** (2015). IBM spss statistics for windows. 23.0 ed. Armonk, New York: IBM Corp.
- 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, 300-311.
- Koo TK, Li M.Y.** (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2):155–163.
- Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., Brown, T. A., Carpenter, W. T., Caspi, A., Clark, L. A., Eaton, N. R., Forbes, M. K., Forbush, K. T., Goldberg, D., Hasin, D., Hyman, S. E., Ivanova, M. Y., Lynam, D. R., Markon, K., Miller, J. D., Moffit, T. E., Morey, L. C., Mullins-Sweatt, Ormel, J., Patrick, C. J., Regier, D. A., Rescorla, L., Ruggero, D. A., Samuel, D. B., Selbom, M., Simms, L. J., Skodol, A. E., Slade, T., South, S. C., Tackett, J. L., Waldman, I. D., Waszczuk, M. A., Widiger, T., Wright, A. G. C., & Zimmerman, M.** (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A Dimensional Alternative to Traditional Nosologies. *Journal of Abnormal Psychology*, 126(4), 454-477.
- Leising, D., Erbs, J. & Fritz, U.** (2010). The letter of recommendation effect in informant ratings of personality. *Journal of Personality and Social Psychology*, 98, 668-682.
- Lieberman, L., Gorka, S. M., Huggings, A. A., Katz, A. C., Sarapas, C. & Shankman, S. A.** (2016). Agreement between self- and informant-reported ratings of personality traits. The moderating effects of major depressive and/or panic disorder. *Journal of Nervous and Mental Disease*, 204, 306-313.

SELF- AND INFORMANT-REPORTED PERSONALITY PATHOLOGY

Markon, K. E., Quilty, L. C., Bagby, R. M., & Krueger, R. F. (2013), The development and psychometric properties of an informant-report form of the personality inventory for DSM-5 (PID-5). *Assessment*, 20(3), 370-383.

Meeuwissen-Van Pol, E., Rossi, G., De Weerd-Spaetgens, C. & Van Alphen, S. P. J. (2020). Screening of personality disorders in a geriatric medicine population. *European Geriatric Medicine*, 11, 289-295.

Mosterman, R. M. & Hendriks, A. a. J. (2011). Self-other disagreement in personality assessment: Significance and prognostic value. *Clinical Psychology and Psychotherapy*, 18, 159-171.

Napierala, M. A. (2012). What is the Bonferroni correction? *Aaos Now*, 40-41.

Oltmanns, J. R. & Oltmanns, T. F. (2021). Self-other agreement on ratings of personality disorder symptoms and traits: Three meta-analyses. *In: Letzring, T. D. & Spain, J. S. (eds.) The handbook of accurate personality judgment: Theory and empirical findings*. Oxford Press University.

Oltmanns, T. F., Rodrigues, M. M., Weinstein, Y. & Gleason, M. E. J. (2014). Prevalence of personality disorders at midlife in a community sample: Disorders and symptoms reflected in interview, self, and informant reports. *Journal of Psychopathology and Behavioral Assessment*, 36, 177-188.

Penders, K. A. P., Peeters, I. G. P., Metsemakers, J. F. M. & Van Alphen, S. P. J. (2020). Personality disorders in older adults: A review of epidemiology, assessment, and treatment. *Current Psychiatry Reports*, 22, 1-14.

Penders, K. A. P., Rossi, G., Metsemakers, J. F. M., Duimel-Peters, I. G. P. & Van Alphen, S. P. J. (2016). Diagnostic accuracy of the Gerontological Personality disorder Scale (GPS) in dutch general practice. *Aging & Mental Health*, 20, 318-328.

Penders, K. A. P., Van Zadelhoff, E., Rossi, G., Duimel-Peeters, I. G. P., Van Alphen, S.

P. J. & Metsemakers, J. F. M. (2018). Feasibility and acceptability of the Gerontological Personality disorders Scale (GPS) in general practice: A mixed methods study. *Journal of Personality Assessment*, 1-10.

Reynolds, K., Pietrzak, R. H., El-Gabalawy, R., Mackenzie, C. S. & Sareen, J. (2015).

Prevalence of psychiatric disorders in u.S. Older adults; findings from a nationally representative survey. *World Psychiatry*, 14, 74-81.

Rossi, G., Debast, I., & van Alphen, S.P.J. (2017). Measuring personality functioning in older adults; construct validity of the Severity Indices of Personality Problems-Short Form (SIPP-SF). *Aging & Mental Health*, 21, 703-711.

Rossi, G., Van Den Broeck, J., Dierckx, E., Segal, D. & Van Alphen, S. P. J. (2014).

Personality assessment among older adults: The value of personality questionnaires unraveled. *Aging & Mental Health*, 18, 936-940.

Samuel, D. B. (2015). A review of the agreement between clinicians' personality disorder diagnoses and those from other methods and sources. *Clinical Psychology; Science and Practice*, 22, 1-19.

Santor, D. A., Bagby, R. M. & Joffe, R. T. (1997). Evaluating stability and change in personality and depression. *Journal of Personality and Social Psychology*, 73, 1354-1362.

Schotte, C.K.W., & de Doncker, D. (1994). ADP-IV vragenlijst [ADP-IV questionnaire]. Antwerp, Belgium: Universitair Ziekenhuis Antwerpen.

Silberman, C. S., Roth, L., Segal, D. L. & Burns, W. (1997). Relationship between the millon clinical multi-axial inventory-II and coolidge axis-II inventory in chronically mentally ill older adults. A pilot study. *Journal of Clinical Psychology*, 53, 559-566.

SELF- AND INFORMANT-REPORTED PERSONALITY PATHOLOGY

- Stone, L. E., Segal, D. L., & Noel, O. R.** (2021). Psychometric evaluation of the Levels of Personality Functioning Scale—Brief Form 2.0 among older adults. *Personality Disorders: Theory, Research, and Treatment*, 12(6), 526–533.
- Van Alphen, S. P. J., Engelen, G. J. J. A., Kuin, Y. & Derksen, J. J. L.** (2004). Persoonlijkheidsstoornissen bij ouderen. Een overzicht. [personality disorders in older adults. An overview.]. *Tijdschrift voor Psychiatrie*, 46, 145-156.
- Van Alphen, S. P. J., Engelen, G. J. J. A., Kuin, Y., Derksen, J. J. L. & Van Der Staak, C. P. F.** (2001). Persoonlijkheidsstoornissen bij ouderen: Drie casestudies. [personality disorders in older adults: Three case-studies]. *Tijdschrift voor Gerontologie en Geriatrie*, 32, 245-251.
- Van Alphen, S. P. J., Engelen, G. J. J. A., Kuin, Y., Hoijtink, H. J. A. & Derksen, J. L. L.** (2006). A preliminary study of the diagnostic accuracy of the gerontological personality disorders scale (gps). *International Journal of Geriatric Psychiatry*, 21, 862-868.
- Van Alphen, S. P. J., Engelen, G. J. J. A., Kuin, Y., Van Der Staak, C. P. F. & Derksen, J. J. L.** (2003). Een delphi-onderzoek naar persoonlijkheidsdiagnostiek in de (a)ggz bij ouderen. [a delphi-study on diagnosis of personality disorders in older adults in mental healthcare]. *Tijdschrift voor Gerontologie en Geriatrie*, 34, 208-214.
- Van Alphen, S. P. J., Rettig, A. M., Engelen, G. J. J. A., Kuin, Y. & Derksen, J. J. L.** (2005). Patient-informantenovereenkomst op de gerontologische persoonlijkheidsstoornissen schaal [patient- informant agreement on the gerontological personality disorders scale]. *Tijdschrift voor Psychiatrie*, 47, 613-617.
- Vazire, S.** (2010). Who knows what about a person? The self-other knowledge asymmetry (soka) model. *Journal of Personality and Social Psychology*, 98, 281-300.

SELF- AND INFORMANT-REPORTED PERSONALITY PATHOLOGY

- Vazire, S. & Carlson, E. N.** (2011). Others sometimes know us better than we know ourselves. *Current Directions in Psychological Science*, 20, 104-108.
- Verheul, R., Andrea, H., Berghout, C., Dolan, C.C., Busschbach, J.J.V., Van der Kroft, P.J.A., Bateman, A.W., & Fonagy, P.** (2008). Severity Indices of Personality Problems (SIPP-118): Development, factor structure, reliability, and validity. *Psychological Assessment*, 20, 23-34.
- Watson, D., Hubbard, B. & Wiese, D.** (2000). Self-other agreement in personality and affectivity: The role of acquaintanceship, trait visibility, and assumed similarity. *Journal of Personality and Social Psychology*, 78, 546-558.
- Weekers, L. C., Hutsebaut, J., & Kamphuis, J. H.** (2018). The Level of Personality Functioning Scale-Brief Form 2.0: Update of a brief instrument for assessing level of personality functioning. *Personality and Mental Health*, 13(1), 3-14.
- Widiger, T. & Seidlitz, L.** (2002). Personality, psychopathology, and aging. *Journal of Research in Personality*, 36, 335-362.

SELF- AND INFORMANT-REPORTED PERSONALITY PATHOLOGY

	Self-reported 'Yes'	Informant-reported 'Yes'	Chi-square
GPS HAB subscale			
HAB 1 I don't like growing older because I become less attractive	24	27	.085
HAB 2 I often worry about my health	72	135	27.2**
HAB 3 I'm often concerned about my memory	72	79	.310
HAB 4 I hope that others solve my problems	29	49	5.26*
HAB 5 I'm often afraid of losing those who care for me, such as family members of my partner	157	140	1.58
HAB 6 I'm often taken advantage of by others	7	44	27.6**
HAB 7 I find it difficult to fend for myself	48	54	.291
GPS BIO subscale			
BIO 1 In my life I've been to see the doctor often for many vague physical complaints	52	70	2.91
BIO 2 I have sometimes said to my family or friends that I don't want to live any longer	14	18	.296
BIO 3 In the past I've been admitted to a psychiatric institution because of nerves	16	19	.121
BIO 4 At important times of my life, I experienced high levels of nervousness, stress or sadness	66	95	6.47*
BIO 5 In the past I've already had treatment from a psychiatrist or psychologist	51	40	1.28
BIO 6 I have sometimes tried to end my life	9	5	.657
BIO 7 At the most I've only had 1 acquaintance or friend in my life	12	29	6.63*
BIO 8 In my life I've not been very interested in sexual contact	39	49	1.06
BIO 9 In the past I've often taken tranquilizers and/or sleeping pills	26	33	.671

Table 1: Self- and Informant-Reported Frequencies of Item Endorsements and Agreement

Note. * $p < .05$, ** $< .01$, bold indicates that the chi-square is significant at Bonferroni corrected alpha level (in total 46 comparisons; $p .0011$)

Table 2. Differences and Levels of Agreement of Self-Report and Informant-Report Scores on the GPS

	Self-report	Informant-report
GPS HAB subscale		
<i>M (SD)</i>	1.25 (1.3)	1.62 (1.5)
<i>d (CI)</i>	.264 (.046 - .482)	
<i>t (634)</i>	-3.328**	
GPS BIO subscale		
<i>Md (n)</i>	.00 (326)	1.00 (326)
<i>r</i>	.087	
<i>U</i>	58036	
<i>z</i>	2.22*	
GPS Total scale		
<i>M (SD)</i>	2.13 (2.2)	2.72 (2.5)
<i>d (CI)</i>	.251 (.033 - .469)	
<i>t (637)</i>	-3.166**	

Note. T-tests were used to compare the means scores of the self- and informant report versions of the GPS HAB scale and GPS Total. Mann-Whitney *U* tests were used to calculate differences between self-report and informant-report with respect to the GPS BIO scale. The effect size for the HAB scale and Total score, Cohen's *d*, were calculated by subtracting the means and dividing the result by the pooled standard deviation. The effect size for the BIO scale, is calculated using the following formula; $r = z / \sqrt{N}$. Cohen's criteria (1988) were used to judge the effect size, with .1=small, .3=medium, .5=large. * $p < .05$; ** $p < .01$, bold indicates the t-test of Mann-Whitney *U* test is significant at Bonferroni corrected alpha level (in total 46 comparisons; $p .0011$)

SELF- AND INFORMANT-REPORTED PERSONALITY PATHOLOGY

Table 3. Means, (Medians) Standard Deviations, of the GPS-iv Separated by Levels of Rated Quality of Relationship and the Congeniality Older Adult and Relationship type.

		<i>N</i>	<i>M (SD)</i>	ANOVA	
				(<i>df.</i> , <i>df.</i>) <i>F</i>	<i>D</i> ² .
GPS-iv HAB scale					
Rated quality (grade)					
	< 6.5	10	2.60 (1.8)	(2, 321) = 1.68	0.01
	7-8.5	138	1.69 (1.6)		
	9-10	176	1.70 (1.5)		
	Overall	324	1.77 (1.5)		
	Missings	2	-		
Rated congeniality (grade)					
	< 6.5	12	3.08 (1.9)	(2, 222) = 6.42	0.05
	7-8.5	131	1.76 (1.6)		
	9-10	180	1.43 (1.3)		
	Overall	323	1.63 (1.5)		
	Missings	3	-		
Relationship type					
	Romantic	256	1.64 (1.5)	(2, 323) = 1.64	0.01
	Family	54	1.35 (1.5)		
	Friends	16	2.25 (2.2)		
	Overall	326	1.62 (1.5)		
GPS-iv BIO scale					
Rated quality (grade) [^]					
	< 6.5	10	<i>Md</i> 2.00	<i>K-W</i> (2, 324) = 9.89	0.04
	7-8.5	138	1.00		
	9-10	176	.00		
	Overall	324	1.00		
	Missings	2	-		
Rated congeniality (grade) [^]					
	< 6.5	12	<i>Md</i> 2.00	<i>K-W</i> (2, 323) = 15.77 <6.5 scores sig. higher than both other groups	0.05
	7-8.5	131	1.00		
	9-10	180	.00		
	Overall	323	1.00		
	Missings	3	-		
Relationship type					
	Romantic	256	1.01 (1.4)	(2, 323) = 4.65	0.03
	Family	54	1.22 (1.6)		
	Friends	16	2.13 (2.0)		
	Overall	326	1.10 (1.5)		
GPS-iv Total scale					
Rated quality (grade)					
	< 6.5	10	5.10 (4.0)	(2, 323) = 5.64	0.03
	7-8.5	138	2.88 (2.6)		
	9-10	176	2.48 (2.3)		
	Overall	324	2.73 (2.5)		
	Missings	2	-		
Rated congeniality (grade)					
	< 6.5	12	5.75 (3.5)	(2, 320) = 10.99 <6.5 scores sig. higher than both other groups	0.06
	7-8.5	131	2.98 (2.5)		
	9-10	180	2.35 (2.3)		
	Overall	323	2.73 (2.5)		
	Missings	3	-		
Relationship type					
	Romantic	256	2.64 (2.4)	<i>W</i> (2, 323) = 2.42	0.01
	Family	54	2.57 (2.7)		
	Friends	16	4.38 (3.7)		
	Overall	326	2.72 (2.5)		

SELF- AND INFORMANT-REPORTED PERSONALITY PATHOLOGY

Note: bold indicates that the chi-square is significant at Bonferroni corrected alpha level ($p < .0011$ given the total of 46 comparisons); ^ Assumptions are violated; Kruskal-Wallis tests were done instead of ANOVA. Eta-squared (η^2) was used as the effect size, following the criteria according to Cohen (1988) .02=small, .13=medium, .26=large.