

Can personality disorder be accurately assessed in older age? A differential item functioning analysis of ICD-11 inventories

Gina Rossi¹  | Sebastiaan P. J. van Alphen^{1,2,3} | Arjan C. Videler^{4,5} | Carmen Diaz-Batanero^{6,7}

¹Personality and Psychopathology Research Group (PEPS), Department of Psychology, Vrije Universiteit Brussel (VUB), Brussels, Belgium

²Clinical Centre of Excellence for Older Adults with Personality Disorders, Mondriaan Mental Health Centre, Heerlen-Maastricht, The Netherlands

³Department of Medical and Clinical Psychology, Tilburg University, Tilburg, The Netherlands

⁴PersonaCura, Clinical Centre of Excellence for Personality and Developmental Disorders in Older Adults, Tilburg, The Netherlands

⁵Tranzo, Scientific Centre for Care and Wellbeing of the Tilburg School of Social and Behavioral Sciences of Tilburg University, Tilburg, The Netherlands

⁶Department of Clinical and Experimental Psychology, University of Huelva, Huelva, Spain

⁷Research Center for Natural Resources, Health and the Environment, University of Huelva, Huelva, Spain

Correspondence

Gina Rossi.
Email: Gina.Rossi@vub.be

Abstract

Objectives: The contested categorical personality disorder (PD) criteria are not well suited to inform PD diagnoses in older adults. Yet, the classification of PDs is undergoing a critical transition phase with a paradigm shift to a dimensional approach for diagnosing PDs. No special attention was given to the expression of PDs in older age when the dimensional ICD-11 model was developed. Given that PDs are highly prevalent in older adults, there is an urgent need to examine if ICD-11 related instruments are able to adequately assess for PDs in older adults.

Methods: The age-neutrality of ICD-11 measures was examined in a sample of 208 Dutch community-dwelling adults ($N = 208$, M age = 54.96, $SD = 21.65$), matched on sex into 104 younger (age range 18–64) and 104 older (age range 65–93) adults. An instrument is considered not to be age-neutral if a collective large level of differential item functioning (DIF) exists in a group of items of an instrument (i.e., 25% or more with DIF). We therefore set out to detect possible DIF in the following ICD-11 self-report measures: the Standardized Assessment of Severity of Personality Disorder (SASPD), the Personality Inventory for ICD-11 (PiCD), and the Borderline Pattern Scale (BPS).

Results: DIF analyses using a non-parametric odds ratio approach demonstrated that SASPD, PiCD, and BPS were age-neutral with less than 25% of items showing DIF. Yet, impact of DIF at scale level, examined by way of differential test functioning (DTF), indicated a DTF effect on the SASPD total score.

Conclusions: These results of age-neutrality of the PiCD and BPS are promising for measuring ICD-11 traits and the borderline pattern. Yet, the age-neutral measurement of PD severity requires further research. With a rapidly aging population, its accurate assessment across the entire adult life span, including older age, is a prerequisite for an adequate detection of PDs.

KEYWORDS

accurate assessment, age-neutrality, BPS, differential item functioning, ICD-11, older adults, personality disorder, PiCD, SASPD, self-report

Key points

- Personality disorders (PDs) in older adults are prevalent, therefore accurate assessment instruments are needed.

- The classification and assessment of PDs is shifting toward a dimensional approach, given the critiques on categorical approaches.
- The current study demonstrates that the dimensional ICD-11 model of PDs and related assessment instruments allow for accurate assessment of maladaptive traits and the borderline pattern in older adults by corroborating the age-neutrality of the PiCD and borderline pattern scale instruments.
- Accurate assessment of PD severity across the adult life span requires further research.

1 | INTRODUCTION

The contested categorical personality disorder (PD) criteria, which are still the standard in the current fifth edition of the Diagnostic and Statistical Manual (DSM-5),¹ are not well suited to inform PD diagnoses in older adults. General shortcomings,^{2–4} such as large heterogeneity within PD categories, arbitrary thresholds, high comorbidity, and the high number of other PDs also hold in older age groups. Categorical criteria were developed in younger age groups and the lack of attention to age-related changes in behavior and interpersonal functioning are remarkable.^{5–9} Older adults might, for example, no longer display physical aggression because of frailty or prefer solitary activities due to mobility problems, not because of underlying personality pathology (such as antisocial and schizoid PD). These general and age-specific shortcomings imply that categorical criteria should ideally not be used to assess and diagnose PDs in later-life context.

Yet, the classification of PDs is currently undergoing a critical transition phase with a paradigm shift to a dimensional approach of diagnosing PDs. According to meta-analyses dimensional approaches have greater clinical utility than categorical approaches.^{10–12} Also, there is no empirical evidence for PDs being categorical; on contrary, there is vast evidence supporting dimensional models, including dimensions such as severity and maladaptive traits.^{2,13–18} The recently released 11th edition of the International Classification of Diseases (ICD-11)¹⁹ therefore launched a dimensional model of PDs (which we will label the ICD-11 PD model hereafter), where PD presence is evaluated by a general severity rating of impairments in self and/or interpersonal functioning. Optionally, the PD style can then be further described by five trait domains (Negative Affectivity, Detachment, Dissociality, Disinhibition, Anankastia) and a Borderline pattern.

Unfortunately, the ICD-11 PD model¹⁹ paid no specific attention to the expression of personality pathology in older age, although the prevalence of PDs in older adults is comparable to or even higher than that in younger adult populations. The National Epidemiological Survey on Alcohol and Related Conditions (NESARC), one of the largest national representative studies on the prevalence of mental illness in the United States, making lifetime assessment of PDs across two waves,²⁰ found a prevalence rate of 9.1% for PDs in community samples. Two studies used these NESARC data to estimate the prevalence in older adults. Based on NESARC wave 1, a prevalence of 10.7% was found in community-dwelling older adults aged 50 or more (yet not all PDs were included, namely borderline,

schizotypal and narcissistic PD were not assessed),²¹ and in wave 2 (which assessed all PDs) the prevalence was 14.5%.²² On one hand, dimensional models such as ICD-11 might hold promise for PD assessment in older adults. By measuring the degree of severity of personality functioning as the core of PD presence and providing quantitative trait qualifiers to further describe individual patterns of expression of personality dysfunction, dimensional approaches allow a more fine-grained assessment than qualitative distinct PD categories.^{5,23,24} On the other hand, a systematic review²⁵ on the measurement invariance of personality measures capturing normative (e.g., adaptive traits) as well as pathological personality constructs (e.g., PD) found 9 studies that analyzed differential item functioning (DIF; i.e., different score patterns²⁶) across age groups to examine failure of age-neutrality. Mixed results (reported by ranges of 10% item-DIF) were found: five studies showed less than 10% items with DIF, one 10%–20%, two 20%–30% and one 30%–40%. Generally, a test is considered age-neutral if less than 25% of items show DIF.²⁶ Given the mixed findings, it is important to examine whether the items of measures using personality constructs are not age-biased.

Research into the age-neutrality of ICD-11 model¹⁹ measures is currently lacking, yet there is some evidence for the conceptually similar DSM-5¹ Alternative Model of PDs (AMPD). Neither model gave specific attention to the gerontological context during their development, but both the ICD-11 and the AMPD conceptualize PD by level of severity and maladaptive traits. The AMPD also considers self and interpersonal functioning to evaluate PD presence, and additionally requires the presence of a maladaptive trait (4 traits overlap with the ICD-11 traits, namely Negative Affectivity, Detachment, Antagonism, and Disinhibition, yet AMPD does not include Anankastia but a Psychoticism trait domain). The limited evidence for instruments measuring PDs from the AMPD framework has provided mixed results. Age-neutrality was examined by exploring item level DIF.²⁶ Additionally, impact at scale level was evaluated by the additive effect of DIF across all items of a scale, which is named differential test functioning (DTF²⁷). The Severity Indices of Personality Functioning—Short Form (SIPP-SF; Short form of the SIPP-118,²⁸ online available at www.deviersprong.nl) is a 60 item measure of personality dysfunction with four scales aligned with the AMPD severity.^{29,30} The SIPP-SF appeared age-neutral as omnibus test³¹: 94% of the items did not show DIF. In other words, only 6% of items showed DIF (which is well below the 25% threshold for lack of age-neutrality²⁷). Most scales related to self and interpersonal functioning (respectively self-control, identity integration, and social

concordance) showed no measurement bias, yet there was DTF for relational functioning. The Level of Personality Functioning Scale Brief Form 2.0 (LPFS-BF 2.0)³² was developed to assess self and interpersonal functioning with 12 items, as conceptualized by the AMPD. Although only 8% of DIF items was found, the total LPFS-BF 2.0 score showed DTF, yet the self and interpersonal subscales showed no DTF at scale level.³³ AMPD's maladaptive traits can be measured using the Personality Inventory for DSM-5 (PID-5) measures. The original PID-5 version³⁴ with 25 facets, belonging to five trait domains, was age-neutral as a whole (only 15% of items had DIF). Impact on scale level was limited to 4 facets (Withdrawal, Attention Seeking, Rigid Perfectionism, and Unusual Beliefs), so 21 out of 25 facets were age-neutral³⁵ (but is with 220 items rather lengthy for older adults). However, a study³¹ on the brief form measuring only the domain level with 25 items, the PID-5-BF,³⁶ lacked age-neutrality at omnibus level, 25% of the items showed DIF, and this impacted all 5 domains (all showed DTF). On contrary, the Personality Inventory for DSM-5 Modified + (PID-5-BF + M)³⁷ assessing both the AMPD and the ICD-11 maladaptive personality traits by 36 items based on AMPD facets related to this traits, was 94% DIF free on item level and no DTF was found for traits at scale level.³³ It is clear that results are partly hopeful, but that age-invariance cannot be taken for granted, thus information on age-neutrality of the ICD-measures is urgently needed. Also, when DIF was found in the above-mentioned studies, it could be explained by item content not taking into account the unique gerontological context. For example, items reflecting withdrawal might be more endorsed by older adults compared to younger adults. Yet, this is probably not due to personality pathology but can be ascribed to situations that are more common at older age such as dealing with loss experiences or social isolation caused by physical deterioration.^{35,38}

Up-to-now ICD-11 measurement studies focused on the development and validation of instruments. Among the available measures, the Standardized Assessment of Severity of Personality Disorder (SASPD)³⁹ rates the impact of a personality-related problem on social and interpersonal functioning, as well as on risk of harm to self and others, which corresponds to the initial ICD-11 severity proposal. SASPD scores showed high test-retest reliability (intraclass coefficient = 0.93) and good predictive validity to classify 110 patients into severity of PD defined by clinical raters. The area under the receiver operating characteristic curve (AUC) was 0.86 for mild and 0.84 for moderate PD severity (AUC values can be interpreted⁴⁰ as 90–100 = excellent; 80–90 = good; 70–80 = fair; 60–70 = poor predictive validity). The SASPD had strong convergence with PD severity measures (Inventory of Personality Organization [IPO]⁴¹ and Level of Personality Functioning Scale [LPFS-SR]⁴²), and logical relationships with 5-factor model measures in an Amazon.com's Mechanical Turk (MTurk) sample ($N = 192$) who were currently or had been in mental health treatment.⁴³ This MTurk study was also the initial validation study of the Borderline Pattern Scale (BPS).⁴³ The four borderline components measured by the scale (Affective Instability, Maladaptive Self-Functioning, Maladaptive Interpersonal-Functioning and

Maladaptive Regulation Strategies) were internally consistent. The BPS demonstrated convergent validity with validated measures of borderline PD and measures of general PD severity (IPO, LFPS, SASPD), and logical relationships with five-factor model measures and trait domains of the Personality Inventory for ICD-11 (PiCD).⁴⁴ The PiCD original validation study⁴⁴ used two MTurk samples (with $N = 278$ and 306) and could demonstrate coherent associations with dimensional measures of general and maladaptive personality and provided evidence for a four-factor structural model. Independent validation studies^{45–52} generally further corroborated reliability of the scales and structural and convergent validity of the SASPD, PiCD, and BPS in student, community, and clinical samples. More recently, the ICD-11 Personality Disorder Severity Scale (PDS-ICD-11)⁵³ was developed to provide an up-to-date unidimensional measure of ICD-11 severity, with items capturing self and interpersonal dysfunctioning. The initial study⁵³ demonstrated its unidimensionality, convergence with other PD severity measures and psychosocial impairment measures and diagnostic validity in a community and a mental health sample.

Specifically in older adults (M age = 69.8 years, $SD = 2.8$), the factor structure of the PiCD and the relation of the PiCD with self and informant-reported criteria measures (dissatisfaction with life, mental health problems, relationship dissatisfaction, general health problems, and cognitive problems) were corroborated, as well as self-other agreement with an informant-report version of the PiCD, the Informant-Personality Inventory for ICD-11 (the IPiC).⁵⁴

Summarized, studies in older adults remain scarce and did not focus on age-neutrality of ICD-11 measures. Given the mixed age-neutrality findings with measures of the conceptually similar AMPD model, we set out to explore possible DIF in the ICD-11 self-report measures that were available in Dutch at the start of the study: the SASPD,³⁹ the Personality Inventory for ICD-11 (PiCD)⁴⁴ and the BPS.⁴³ Data was collected in community samples. As the above mentioned NESARC prevalence rates²⁰ indicate, PDs are a public health issue and burden in the community.^{53,55} Following the guidelines for psychological practice with older adults,⁵⁶ we set the start of later life at age of 65 in order to compare younger (≤ 64) with older (≥ 65) adults.

2 | METHODS

2.1 | Participants and procedure

A non-probabilistic convenience sample of adults was collected (through social media, hobby clubs, etc.). After providing informed consent, 289 participants completed the questionnaires online ($n = 221$) or on paper ($n = 68$). To ensure data integrity, 47 participants were excluded using the following criteria: an inconsistent answering strategy (score ≥ 17 on the PID-5³⁴ inconsistency scale^{34,53,57}), no more than 25% missing items for one of the questionnaires, and diabetes, surgery or chemotherapy within the last 3 months (given these conditions may have impact on cognitive

functioning^{58–61}). This resulted in 138 younger and 104 older adults. Next, given sex differences can be expected for PDs,⁶² matching on sex was done (with a case-control matching function). A final sample of 104 younger (age range 18–64, *M* age = 36.82, *SD* = 15.38) and 104 older (age range 65–93, *M* age = 73.10, *SD* = 6.41) Dutch community-dwelling adults (*N* = 208, *M* age = 54.96, *SD* = 21.65) was collected this way (with 50% males in both subgroups).

2.2 | ICD-11 self-report inventories

2.2.1 | Standardized assessment of severity of personality disorder

The SASPD³⁹ is a nine item self-report instrument to assess PD severity by level of disability in areas of personality and psychosocial functioning. Four items have an interpersonal focus (i.e., friendships; and being with, trusting, and caring about other people) and five items are more intrapsychic/behavioral (i.e., two items concerning affect and/or emotion regulation and three items assessing acting on impulse, being organized, and self-reliance, respectively). Items are rated on an ordinal level (absent = 0, mild = 1, moderate = 2, severe = 3). It is a brief and good predictor for mild or moderate ICD-11 PD severity, with high test-retest reliability.³⁹

2.2.2 | Personality inventory for ICD-11 (PiCD)

The PiCD⁴⁴ is a 60 item self-report instrument assessing the five broad ICD-11 maladaptive trait domains. The five 12-item scales with the following response format for all items: 1 = strongly disagree with the statement, 2 = disagree, 3 = neutral, 4 = agree, and 5 = strongly agree. Support was provided for the theorized four factor solution, with three unipolar factors (Negative Affectivity, Detachment, Dissociality) and one bipolar factor (Disinhibition-Anankastia), and convergent validity was demonstrated for the PiCD scales with measures of general and maladaptive personality traits.⁴⁴

2.2.3 | Borderline pattern scale

The BPS⁴³ is a 12 item self-report instrument assessing four components (each by 3 items) of the borderline pattern. Items are rated on a five-point scale (1 = strongly disagree with the statement, 2 = disagree, 3 = neutral, 4 = agree, and 5 = strongly agree). The BPS showed good convergent validity with well-validated measures of borderline PD and was internally consistent.⁴³

2.3 | Statistical analyses

Similar to previous AMPD studies,^{31,33,35} DIF analyses were conducted using a non-parametric odds ratio approach for polytomous

items with DIFAS 5.0.⁶³ An item shows DIF when younger and older adults with a similar position on the underlying construct of interest do not have the same probability of endorsing the item and was examined with three statistics: the Mantel–Haenszel statistic,⁶⁴ the Liu-Agresti Cumulative Common Log-Odds Ratio,⁶⁵ and the Cox's Noncentrality Parameter Estimator (Cox's B).⁶⁶ All three statistics address item-level invariance and measure the overall omnibus effect across all score levels.⁶⁷ The Mantel–Haenszel statistic is a chi-square distribution with one degree of freedom with a critical value of 3.84 for a Type 1 error rate of 0.05.⁶⁴ L-A LOR and Cox's B provide an effect size with large DIF indicated by $|L-A LOR| > 0.64$ ²⁷ and $|Cox's B| > 0.40$.⁶⁶ Impact of DIF at scale level was examined by way of DTF and the effect variance is considered large for $v^2 > 0.14$.²⁷ Total scale scores were used as stratifying variables with a stratum size 5 to avoid empty cells. To reduce Type I errors critical values were Bonferonni corrected⁶⁸ for the number of statistical tests (i.e., number of items in the scale and number of statistics used). DIF and DTF analyses were conducted for the SASPD, the PiCD trait domain, and the BPS component scales.

Mean scores across age groups were compared with a *t* test and Cohen *d* effect sizes.⁶⁹

3 | RESULTS

Table 1 displays the items showing large DIF according to Bonferonni corrected critical values. For the SASPD, one out of nine items displayed DIF, for the PiCD five out of 60, and for the BPS, one item out of 12. All tests remained below the 25% criterion for lack of age-neutrality with DIF percentages of respectively 11% for the SASPD, 8% for the PiCD, and 8% for the BPS. Table 2 reports the effect variance v^2 . SASPD total scale score was impacted given the large effect at scale level.

Means for all scales across age groups were compared (see Table 3). However, we will not interpret the results and effect size of the SASPD total scale score, given the large DTF effect found for this scale. If a scale shows DTF, scale scores cannot be used to accurately compare subgroups on presence of the measured construct.²⁷ Older adults scored significantly lower with a large effect on the Disinhibition trait domain of the PiCD, and Affective Instability and Maladaptive Regulation Strategies of the BPS, and with a medium effect on the Negative Affectivity trait domain of the PiCD, and Maladaptive Self Functioning of the BPS.

4 | DISCUSSION

In terms of item level, DIF analyses results are promising: all tools (SASPD, PiCD, and BPS) were age-neutral. With a rapidly aging population, age-neutrality of psychological instruments is becoming ever more important. Further, no DTF scale effects were found for the PiCD maladaptive traits scales and components of the BPS. We can conclude that accurate assessment of maladaptive traits and

TABLE 1 Items showing large differential item functioning according to Bonferonni corrected critical values.

SASPD items displaying DIF against older adults (older < younger)					
Critical Bonferonni corrected values: Mantel–Haenszel 9.69, L-A LOR 0.93, Cox's B 0.58					
Total scale	Item	Content	Mantel–Haenszel	L-A LOR	Cox's B
Worrying (one option needs to be selected)	6	In general, I am not a worrier. I sometimes get worried about things that others don't. I am generally a worrier. Constant worrying stops me from doing things I need to do	16.97	1.25	0.84
PiCD items displaying DIF against older adults (older < younger)					
Critical Bonferonni corrected values: Mantel–Haenszel 10.22, L-A LOR 0.96, Cox's B 0.60					
Domain	Item	Content	Mantel–Haenszel	L-A LOR	Cox's B
Detachment	3	I prefer to stay away from other people	17.44	1.25	0.74
Dissocial	4	My anger has gotten me into fights	21.80	2.11	1.01
Dissocial	59	I have successfully deceived and manipulated persons	11.12	1.13	0.64
PiCD items displaying DIF against younger adults (older > younger)					
Critical Bonferonni corrected values: Mantel–Haenszel 10.22, L-A LOR 0.96, Cox's B 0.60					
Domain	Item	Content	Mantel–Haenszel	L-A LOR	Cox's B
Negative affectivity	26	Changes in my mood are unrelated to what is happening in my life	13.50	–1.07	–0.55
Detachment	23	I feel pretty much the same all the time	13.09	–1.00	–0.51
Anankastia	60	I love the motto, “think before you act”	9.35	–0.97	–0.62
BPS items displaying DIF against older adults (older < younger)					
Critical Bonferonni corrected values: Mantel–Haenszel 7.69, L-A LOR 0.81, Cox's B 0.49					
Component	Item	Content	Mantel–Haenszel	L-A LOR	Cox's B
Maladaptive regulation strategies	8	“I have no real self-control over what I do”	10.25	1.23	0.85

Note: Positive L-A LOR and Cox's B values indicate DIF in favor of the reference group (i.e. younger adults), and negative values indicate DIF in favor of the focal group (i.e. older adults). For item 60 of the PiCD, the Mantel–Haenszel statistic was below the critical Bonferonni corrected values, yet both L-A LOR and Cox's B indicated DTF with values above the critical Bonferonni corrected value.

TABLE 2 Results of differential test functioning.

Scale	Weighted v^2
SASPD total score	0.36
PiCD negative affectivity	0.16
PiCD detachment	0.32
PiCD disinhibition	0.04
PiCD dissociality	0.15
PiCD anankastia	0.20
BPS affective instability	–0.10
BPS maladaptive self functioning	0.06
BPS maladaptive interpersonal functioning	0.21
BPS maladaptive regulation strategies	0.11

Note: Large Bonferonni corrected effect if weighted $v^2 \geq 0.35$ for the SASPD total score, ≥ 0.34 for the PiCD scales, and ≥ 0.40 for the BPS scales. Large effects are bold.

the borderline pattern is possible across the adult life span, and more specifically also in older adults.

However, at the same time, a DTF effect was found for the total scale of the SASPD. PD severity is a core criterion to indicate PD presence according to ICD-11.¹⁹ Thus, accurate assessment of severity across the entire adult life span, including older age, is a prerequisite for an adequate detection of PDs. On one hand, a possibility would be to replace the item with DIF (item 6 Worrying) by a more age-neutral one, covering the same content. Item 6 captures possible dysfunctional consequences of the ICD-11 trait Negative Affectivity,³⁹ together with items 4 (Temper) and 9 (Lack of Self-reliance) that showed no DIF. On the other hand, the content of this item (i.e., worrying) (and other SASPD items) is no longer represented in the most recent ICD-11 model for PDs. SASPD (based on an early ICD-11 draft) provided an index of complexity by items representing personality related problems from different PD categories.^{39,46,70} The PDS-ICD-11 is with 14 items a short instrument assessing PD

Scale/Instrument	Younger adults M (SD)	Older adults M (SD)	t test p value	Cohen d
Total/SASPD (PD severity)	6.23 (2.84)	4.06 (2.55)	<0.001	0.81
Negative affectivity/PiCD	35.71 (9.14)	30.26 (7.61)	<0.001	0.65
Detachment/PiCD	28.43 (7.34)	27.36 (6.76)	0.272	0.15
Disinhibition/PiCD	28.88 (7.12)	23.59 (5.26)	<0.001	0.85
Dissociality/PiCD	24.70 (7.67)	22.03 (5.28)	0.004	0.41
Anankastia/PiCD	38.02 (6.97)	39.39 (6.92)	0.115	-0.20
Affective instability/BPS	7.20 (2.93)	5.05 (1.85)	<0.001	0.88
Maladaptive self functioning/BPS	6.44 (2.73)	4.94 (1.98)	<0.001	0.63
Maladaptive interpersonal functioning/BPS	6.95 (2.28)	5.88 (2.15)	<0.001	0.48
Maladaptive regulation strategies/BPS	5.96 (2.39)	4.23 (1.35)	<0.001	0.89

Note: Cohen $d = 0.20$ indicates a small effect, $d = 0.50$ a moderate effect, and $d = 0.80$ a large effect. Large effects are bold. Medium effects are cursive.

severity according to the final ICD-11 model, which thus also has the advantage to make systematic screening for PD presence possible in geriatric psychiatry. We therefore advise that future research examines the age-neutrality of the PDS-ICD-11.

Although the PiCD and BPS scales showed no DTF, there were some items with DIF. For example, the PiCD Detachment scale had two DIF items: "I prefer to stay away from other people" was less endorsed by older adults and the item "I feel pretty much the same all the time" was more endorsed by older adults, compared to younger adults with a similar position on Detachment. Yet, the DTF value of 0.32 remained below the critical DTF value of 0.34. This is not unusual, with relatively few DIF items at item-level, the DIF can indeed be canceled out at scale level. DIF findings also corresponded to what can be expected according to the age-group. The item "I have successfully deceived and manipulated persons" showed DIF, with older adults endorsing less frequently in the keyed direction of the scale. This corresponds with the linear decreasing prevalence rate of the DSM deceitfulness criterion across increasing age groups.⁷¹ Other items with DIF measured control over behavior (e.g., "My anger has gotten me into fights") and emotion regulation strategies (e.g., "I feel pretty much the same all the time"), with younger adults endorsing more in the direction of lack of behavioral control and older adults endorsing higher on emotion regulation. The SASPD item with DIF ("worrying") confirmed this latter finding (younger adults endorsed more in the keyed direction of worrying than younger adults). The differences in mean scores further corroborated these findings. Older adults scored lower on Disinhibition, Negative Affectivity, Affective Instability, Maladaptive Regulation Strategies, and Maladaptive Self Functioning (and thus reported to experience these traits less frequently). Feeling more in control and behaving more emotionally stable is a normative change in personality associated with more psychological maturity as one gets older.⁷² Additionally, increasing age goes along with physical age-related impairments, resulting in less possibilities to behave recklessly and impulsively.^{73,74} Also,

TABLE 3 Independent samples t test for means of all scale scores across age groups.

according to socio-emotional selectivity theory,⁷⁵ the shorter time horizon of older adults explains the higher motivation for prioritizing emotional goals. Older adults generally apply a more response-focused emotion regulation when negative life events happen. Studies indeed confirm emotion regulation improves across the adult life span.^{76,77}

A limitation to be noted is that the current results are restricted to self-report. Further research is needed using informant report and interviews by clinicians. Self-informant agreement on PD measures is generally low to moderate,⁷⁸ so findings on age-neutrality of PD measures obtained with self-report cannot be extended to informant-report. The same holds for the agreement between clinicians making PD diagnoses with semi-structured interviews versus self-report, where agreement is also modest.⁷⁹

Furthermore, data were collected in community-dwelling participants, which partly limits generalizability to clinical samples where higher levels of maladaptive personality functioning or traits can be expected to be more frequently endorsed. On the other hand, in our total sample 11.5% indicated to receive psychological treatment at the moment the study was conducted, and 18.3% indicated to have recently received psychological treatment. Our current study could demonstrate the age-neutrality of the PiCD and BPS, and the presence of DIF in the SASPD at item and scale level with an omnibus approach. Nevertheless, clinical samples remain an important focus for further research to detect manifestation of DIF at any particular score level, for example, using graded response models.⁸⁰ Yet, such an approach would also require larger samples.

Finally, definitions of older age can vary, and although the age of 65 is often considered the start of later life, there will still be a large degree of heterogeneity between older adults. Therefore, older adults are sometimes further subdivided into youngest-old (65–74), middle-old (75–84) and oldest-old (≥ 85).⁷⁴ Future studies using smaller older age-ranges might thus reveal even more fine grained differences in item endorsement due to age-related changes.

ACKNOWLEDGMENTS

We would like to thank the following VUB master thesis students for their contribution to the study data collection: Laura Gooskens, Silke Glorie, & Olga Turovskaja. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Data is available on <https://osf.io/5qnm7/>.

ETHICS STATEMENT

Ethical approval was provided by the Medical Ethical Commission of University Hospital Brussel *Universitair Ziekenhuis Brussel (UZB)–Vrije Universiteit Brussel (VUB)*.

INFORMED CONSENT

Written informed consent was obtained from all participants before study enrollment.

ORCID

Gina Rossi  <https://orcid.org/0000-0002-6803-9820>

REFERENCES

- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. text rev.; 2022.
- Hopwood CJ, Kotov R, Krueger RF, et al. The time has come for dimensional personality disorder diagnosis. *Pers Ment Health*. 2018;12(1):82-86. <https://doi.org/10.1002/pmh.1408>
- Widiger T, Trull T. Plate tectonics in the classification of personality disorder: shifting to a dimensional model. *Am Psychol*. 2007;62(2):71-83. <https://doi.org/10.1037/0003-066x.62.2.71>
- Clark LA. Assessment and diagnosis of personality disorder: perennial issues and an emerging reconceptualization. *Annu Rev Psychol*. 2007;58(1):227-257. <https://doi.org/10.1146/annurev.psych.57.102904.190200>
- Rossi G, Videler AC, van Alphen SPJ. Challenges and developments in the assessment of (Mal)adaptive personality and pathological States in older adults. *Assessment*. 2018;25(3):279-284. <https://doi.org/10.1177/1073191116685810>
- Rossi G, Van den Broeck J, Dierckx E, Segal D, Van Alphen SJP. Personality assessment among older adults: the value of personality questionnaires unraveled. *Aging Ment Health*. 2014;18(8):936-940. <https://doi.org/10.1080/13607863.2014.924089>
- van Alphen SPJ, Bolwerk N, Videler AC, et al. Age-related aspects and clinical implications of diagnosis and treatment of personality disorders in older adults. *Clin Gerontol*. 2012;35(1):27-41. <https://doi.org/10.1080/07317115.2011.628368>
- Balsis S, Segal DL, Donahue C. Revising the personality disorder diagnostic criteria for the Diagnostic and Statistical Manual of Mental Disorders-Fifth Edition (DSM-V): consider the later life context. *Am J Orthopsychiatry*. 2009;79(4):452-460. <https://doi.org/10.1037/a0016508>
- Holzer KJ, Vaughn MG, Fearn NE, Loux TM, Mancini MA. Age bias in the criteria for antisocial personality disorder. *J Psychiatr Res*. 2021;137:444-451. <https://doi.org/10.1016/j.jpsychires.2021.03.025>
- Bornstein RF, Natoli AP. *Clinical Utility of Categorical and Dimensional Perspectives on Personality Pathology: A Meta-Analytic Review*. Personality disorders; 2019.
- Milinkovic MS, Tiliopoulos N. A systematic review of the clinical utility of the DSM-5 section III alternative model of personality disorder. *Personal Disord*. 2020;11(6):377-397. <https://doi.org/10.1037/per0000408>
- Tracy M, Tiliopoulos N, Sharpe L, Bach B. The clinical utility of the ICD-11 classification of personality disorders and related traits: a preliminary scoping review. *Aust N Z J Psychiatr*. 2021;55(9):849-862. <https://doi.org/10.1177/00048674211025607>
- Krueger RF, Markon KE. The role of the DSM-5 personality trait model in moving toward a quantitative and empirically based approach to classifying personality and psychopathology. *Annu Rev Clin Psychol*. 2014;10(1):477-501. <https://doi.org/10.1146/annurev-clinpsy-032813-153732>
- Morey LC, Hopwood CJ, Gunderson JG, et al. Comparison of alternative models for personality disorders. *Psychol Med*. 2007;37(7):983-994. <https://doi.org/10.1017/s0033291706009482>
- Widiger TA, Trull TJ. Plate tectonics in the classification of personality disorder: shifting to a dimensional model. *Am Psychol*. 2007;62(2):71-83. <https://doi.org/10.1037/0003-066x.62.2.71>
- Kim YR, Blashfield R, Tyrer P, Hwang ST, Lee HS. Field trial of a putative research algorithm for diagnosing ICD-11 personality disorders in psychiatric patients: 1. Severity of personality disturbance. *Pers Ment Health*. 2014;8(1):67-78. <https://doi.org/10.1002/pmh.1248>
- Kim YR, Tyrer P, Lee HS, et al. Preliminary field trial of a putative research algorithm for diagnosing ICD-11 personality disorders in psychiatric patients: 2. Proposed trait domains. *Pers Ment Health*. 2015;9(4):298-307. <https://doi.org/10.1002/pmh.1305>
- Haslam N, McGrath MJ, Viechtbauer W, Kuppens P. Dimensions over categories: a meta-analysis of taxometric research. *Psychol Med*. 2020;50(9):1418-1432. <https://doi.org/10.1017/s003329172000183x>
- World Health Organization. *International Statistical Classification of Diseases and Related Health Problems*. 11th ed.; 2019. <https://icd.who.int/>
- Trull TJ, Jahng S, Tomko RL, Wood PK, Sher KJ. Revised NESARC personality disorder diagnoses: gender, prevalence, and comorbidity with substance dependence disorders. *J Pers Disord*. 2010;24(4):412-426. <https://doi.org/10.1521/pedi.2010.24.4.412>
- Holzer KJ, Huang J. Physical health-related quality of life among older adults with personality disorders. *Aging Ment Health*. 2019;23(8):1031-1040. <https://doi.org/10.1080/13607863.2018.1468412>
- Reynolds K, Pietrzak RH, El-Gabalawy R, Mackenzie CS, Sareen J. Prevalence of psychiatric disorders in U.S. older adults; findings from a nationally representative survey. *World Psychiatr*. 2015;14(1):74-81. <https://doi.org/10.1002/wps.20193>
- Rossi G, Videler AC, van Alphen SP. A meta-structure for DSM-5 and ICD-11 pathological traits and the differentiation of personality functioning at different trait levels in older adults. *Aging Ment Health*. 2022:1-8.
- Wright AG. Qualitative and quantitative distinctions in personality disorder. *Pers Assess DSM-5*. 2014:115-124.
- Dong Y, Dumas D. Are personality measures valid for different populations? A systematic review of measurement invariance across cultures, gender, and age. *Pers Individ Differ*. 2020;160:109956. <https://doi.org/10.1016/j.paid.2020.109956>
- Holland PW, Wainer H. *Differential Item Functioning*. L. Erlbaum Associates; 1993.
- Penfield RD, Algina J. A generalized DIF effect variance estimator for measuring unsigned differential test functioning in mixed format tests. *J Educ Meas*. 2006;43(4):295-312. <https://doi.org/10.1111/j.1745-3984.2006.00018.x>

28. Verheul R, Andrea H, Berghout CC, et al. Severity Indices of personality problems (SIPP-118): development, factor structure, reliability, and validity. *Psychol Assess*. 2008;20(1):23-34. <https://doi.org/10.1037/1040-3590.20.1.23>
29. Waugh MH, McClain CM, Mariotti EC, et al. Comparative content analysis of self-report scales for level of personality functioning. *J Pers Assess*. 2021;103(2):161-173. <https://doi.org/10.1080/00223891.2019.1705464>
30. Berghuis H, Kamphuis JH, Verheul R. Specific personality traits and general personality dysfunction as predictors of the presence and severity of personality disorders in a clinical sample. *J Pers Assess*. 2014;96(4):410-416. <https://doi.org/10.1080/00223891.2013.834825>
31. Debast I, Rossi G, van Alphen SPJ. Age-neutrality of a brief assessment of the section III alternative model for personality disorders in older adults. *Assessment*. 2018:1-14. <https://doi.org/10.1177/1073191118754706>
32. Weekers LC, Hutsebaut J, Kamphuis JH. The Level of Personality Functioning Scale-Brief Form 2.0: update of a brief instrument for assessing level of personality functioning. *Pers Ment Health*. 2019;13(1):3-14. <https://doi.org/10.1002/pmh.1434>
33. Facon MF, van Alphen SPJ, Dierckx E, Rossi G. Age-neutral Measurement of Personality Functioning and Maladaptive Personality Traits; 2023. (manuscript accepted for publication in *Assessment*).
34. Krueger RF, Derringer J, Markon KE, Watson D, Skodol AE. Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychol Med*. 2012;42(9):1879-1890. <https://doi.org/10.1017/S0033291711002674>
35. Van den Broeck J, Bastiaansen L, Rossi G, Dierckx E, De Clercq B. Age-neutrality of the trait facets proposed for personality disorders in DSM-5: a DIFAS analysis of the PID-5. *J Psychopathol Behav Assess*. 2013;35(4):487-494. <https://doi.org/10.1007/s10862-013-9364-3>
36. American Psychiatric Association. Online Assessment Measures: The Personality Inventory for DSM-5–Brief Form (PID-5-BF). <https://www.psychiatry.org/psychiatrists/practice/dsm/educational-resources/assessment-measures>
37. Bach B, Kerber A, Aluja A, et al. International assessment of DSM-5 and ICD-11 personality disorder traits: toward a common nosology in DSM-5.1. *Psychopathology*. 2020;53(3-4):179-188. <https://doi.org/10.1159/000507589>
38. Van Alphen S, Engelen G, Kuin Y, Derksen J. *The Relevance of a Geriatric Sub-classification of Personality Disorders in the DSM-V*. Wiley Online Library; 2006:205-209.
39. Olajide K, Munjiza J, Moran P, et al. Development and psychometric properties of the standardized assessment of severity of personality disorder (SASPD). *J Pers Disord*. 2018;32(1):44-56. https://doi.org/10.1521/pedi_2017_31_285
40. Safari S, Baratloo A, Elfil M, Negida A. Evidence based emergency medicine; Part 5 receiver operating curve and area under the curve. *Emerge*. 2016;4(2):111-113. Spring.
41. Lenzenweger MF, Clarkin JF, Kernberg OF, Foelsch PA. The Inventory of Personality Organization: psychometric properties, factorial composition, and criterion relations with affect, aggressive dyscontrol, psychosis proneness, and self-domains in a nonclinical sample. *Psychol Assess*. 2001;13(4):577-591. <https://doi.org/10.1037/1040-3590.13.4.577>
42. Morey LC. Development and initial evaluation of a self-report form of the DSM-5 level of personality functioning scale. *Psychol Assess*. 2017;29(10):1302-1308. <https://doi.org/10.1037/pas0000450>
43. Oltmanns JR, Widiger TA. Evaluating the assessment of the ICD-11 personality disorder diagnostic system. *Psychol Assess*. 2019;31(5):674-684. <https://doi.org/10.1037/pas0000693>
44. Oltmanns JR, Widiger TA. A self-report measure for the ICD-11 dimensional trait model proposal: the personality inventory for ICD-11. *Psychol Assess*. 2018;30(2):154-169. <https://doi.org/10.1037/pas0000459>
45. Carnovale M, Sellbom M, Bagby RM. The Personality Inventory for ICD-11: investigating reliability, structural and concurrent validity, and method variance. *Psychol Assess*. 2020;32(1):8-17. <https://doi.org/10.1037/pas0000776>
46. McCabe GA, Widiger TA. A comprehensive comparison of the ICD-11 and DSM-5 section III personality disorder models. *Psychol Assess*. 2020;32(1):72-84. <https://doi.org/10.1037/pas0000772>
47. Gutiérrez F, Aluja A, Ruiz J, et al. Personality disorders in the ICD-11: Spanish validation of the PiCD and the SASPD in a mixed community and clinical sample. *Assessment*. 2021;28(3):759-772. <https://doi.org/10.1177/1073191120936357>
48. Stricker J, Pietrowsky R. Incremental validity of the ICD-11 personality disorder model for explaining psychological distress. *Personal Disord*. 2022;13(2):97-107. <https://doi.org/10.1037/per0000489>
49. Tarescavage AM, Menton WH. Construct validity of the personality inventory for ICD-11 (PiCD): evidence from the MMPI-2-RF and CAT-PD-SF. *Psychol Assess*. 2020;32(9):889-895. <https://doi.org/10.1037/pas0000914>
50. Damovsky F, Zettl M, Zimmermann J, et al. The personality inventory for ICD-11 (PiCD): reliability and validity of the German version in a clinical and non-clinical sample. *Psychother Psychosom Med Psychol*. 2022.
51. Stricker J, Buecker S, Pietrowsky R. Alignment of the personality inventory for ICD-11 with the five-factor model of personality. *Psychol Assess*. 2022;34(7):711-716. <https://doi.org/10.1037/pas0001141>
52. Stricker J, Jakob L, Köhler D, Pietrowsky R. Six-month stability and predictive validity of the personality inventory for ICD-11. *BMC psychol*. 2022;10(1):1-5. <https://doi.org/10.1186/s40359-022-00979-2>
53. Bach B, Brown TA, Mulder RT, Newton-Howes G, Simonsen E, Sellbom M. Development and initial evaluation of the ICD-11 personality disorder severity scale: PDS-ICD-11. *Pers Ment Health*. 2021;15(3):223-236. <https://doi.org/10.1002/pmh.1510>
54. Oltmanns JR, Widiger TA. The self- and informant-personality inventories for ICD-11: agreement, structure, and relations with health, social, and satisfaction variables in older adults. *Psychol Assess*. 2021;33(4):300-310. <https://doi.org/10.1037/pas0000982>
55. Skodol AE. Personality disorders: a burden in the community, neglected in the clinic? *J Clin Psychiatr*. 2015;76(11):14168. <https://doi.org/10.4088/jcp.14com09597>
56. American Psychological Association. Guidelines for psychological practice with older adults. *Am Psychol*. 2014;69(1):34-65. <https://doi.org/10.1037/a0035063>
57. Keeley JW, Webb C, Peterson D, Roussin L, Flanagan EH. Development of a response inconsistency scale for the personality inventory for DSM-5. *J Pers Assess*. 2016;98(4):351-359. <https://doi.org/10.1080/00223891.2016.1158719>
58. Bhat NA, Muliayala KP, Kumar S. Psychological aspects of diabetes. *Eur Med J*. 2020;8(1):297-304.
59. Hodgson KD, Hutchinson AD, Wilson CJ, Nettelbeck T. A meta-analysis of the effects of chemotherapy on cognition in patients with cancer. *Cancer Treat Rev*. 2013;39(3):297-304. <https://doi.org/10.1016/j.ctrv.2012.11.001>
60. Rasmussen LS. Postoperative cognitive dysfunction: incidence and prevention. *Best Pract Res Clin Anaesthesiol*. 2006;20(2):315-330. <https://doi.org/10.1016/j.bpa.2005.10.011>
61. Moller JT, Cluitmans P, Rasmussen LS, et al. Long-term post-operative cognitive dysfunction in the elderly: ISPOCD1 study. *Lancet*. 1998;351(9106):857-861. [https://doi.org/10.1016/s0140-6736\(97\)07382-0](https://doi.org/10.1016/s0140-6736(97)07382-0)

62. Schulte Holthausen B, Habel U. Sex differences in personality disorders. *Curr Psychiatr Rep*. 2018;20(12):1-7. <https://doi.org/10.1007/s11920-018-0975-y>
63. Penfield RD. DIFAS: differential item functioning analysis system. *Appl Psychol Meas*. 2005;29(2):150-151. <https://doi.org/10.1177/0146621603260686>
64. Mantel N. Chi-square tests with one degree of freedom; extensions of the Mantel-Haenszel procedure. *J Am Stat Assoc*. 1963;58(303):690-700. <https://doi.org/10.2307/2282717>
65. Liu I-M, Agresti A. Mantel-Haenszel-type inference for cumulative odds ratios with a stratified ordinal response. *Biometrics*. 1996;52(4):1223-1234. <https://doi.org/10.2307/2532838>
66. Camilli G, Congdon P. Application of a method of estimating DIF for polytomous test items. *J Educ Behav Stat*. 1999;24(4):323-341. <https://doi.org/10.3102/10769986024004323>
67. Gattamorta KA. *A Comparison of Adjacent Categories and Cumulative DSF Effect Estimators*. University of Miami; 2009.
68. Napierala MA. What is the Bonferroni correction? *Aaos Now*. 2012;40-41.
69. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Lawrence Erlbaum Associates, Publishers; 1988.
70. Rek K, Thielmann I, Henkel M, et al. A psychometric evaluation of the Standardized Assessment of Severity of Personality Disorder (SASPD) in nonclinical and clinical German samples. *Psychol Assess*. 2020;32(10):984-990. <https://doi.org/10.1037/pas0000926>
71. Balsis S, Gleason MEJ, Woods CM, Oltmanns TF. An item response theory analysis of DSM-IV personality disorder criteria across younger and older age groups. *Psychol Aging*. 2007;22(1):171-185. <https://doi.org/10.1037/0882-7974.22.1.171>
72. Roberts BW, Walton KE, Viechtbauer W. Patterns of mean-level change in personality traits across the life course: a meta-analysis of longitudinal studies. *Psychol Bull*. 2006;132(1):1-25. <https://doi.org/10.1037/0033-2909.132.1.1>
73. Abrams RC, Bromberg CE. Personality disorders in the elderly: a flagging field of inquiry. *Int J Geriatr Psychiatry*. 2006;21(11):1013-1017. <https://doi.org/10.1002/gps.1614>
74. Segal DL, Coolidge FL, Rosowsky E. *Personality Disorders and Older Adults: Diagnosis, Assessment, and Treatment*. Wiley; 2006.
75. Carstensen LL, Mikels JA, Mather M. Aging and the intersection of cognition, motivation, and emotion. In: Birren JE, Schaie KW, eds. *Handbook of the psychology of aging*. 6th ed. Academic Press; 2006:343-362.
76. Carstensen LL, Turan B, Scheibe S, et al. Emotional experience improves with age: evidence based on over 10 years of experience sampling. *Psychol Aging*. 2011;26(1):21-33. <https://doi.org/10.1037/a0021285>
77. Burr DA, Castrellon JJ, Zald DH, Samanez-Larkin GR. Emotion dynamics across adulthood in everyday life: older adults are more emotionally stable and better at regulating desires. *Emotion*. 2021;21(3):453-464. <https://doi.org/10.1037/emo0000734>
78. Klonsky ED, Oltmanns TF, Turkheimer E. Informant-reports of personality disorder: relation to self-reports and future research directions. *Clin Psychol Sci Pract*. 2002;9(3):300-311. <https://doi.org/10.1093/clipsy/9.3.300>
79. Samuel DB. A review of the agreement between clinicians' personality disorder diagnoses and those from other methods and sources. *Clin Psychol Sci Pract*. 2015;22(1):1-19. <https://doi.org/10.1111/cpsp.12088>
80. Cohen AS, Kim S-H, Baker FB. Detection of differential item functioning in the graded response model. *Appl Psychol Meas*. 1993;17(4):335-350. <https://doi.org/10.1177/014662169301700402>

How to cite this article: Rossi G, van Alphen SPJ, Videler AC, Diaz-Batanero C. Can personality disorder be accurately assessed in older age? A differential item functioning analysis of ICD-11 inventories. *Int J Geriatr Psychiatry*. 2024;e6075. <https://doi.org/10.1002/gps.6075>