International Journal of Methods in Psychiatric Research
Int. J. Methods Psychiatr. Res. 20(3): 135–144 (2011)
Published online 2 August 2011 in Wiley Online Library
(wileyonlinelibrary.com) **DOI:** 10.1002/mpr.342

A NEO-PI-R short form for older adults

BIANCA MOOI,[1] HANNIE C. COMIJS,[2] FILIP DE FRUYT,[3] DINEKE DE RITTER,[4] HANS A. HOEKSTRA[5] & AARTJAN T. F. BEEKMAN[2]

1 Department of Old Age Psychiatry, GGz Centraal, Amersfoort, The Netherlands
2 EMGO Institute, VU University Medical Center, Amsterdam, The Netherlands
3 Department of Developmental, Personality, and Social Psychology, Ghent University, Belgium
4 Department of Old Age Psychiatry, GGZ Breburg, Breda, The Netherlands
5 Department of Differential Psychology, Groningen University, The Netherlands

Key words
personality, older adults, NEO-PI-R, short form, Five-Factor Model

Correspondence
Bianca Mooi, GGz Centraal, Department of Old Age Psychiatry, Postbus 3051, 3800 DB Amersfoort, The Netherlands.
Telephone (+31) (0)33 4609635
Fax (+31) (0)33 4609480
Email: b.mooi@ggzcentraal.nl

Received 26 March 2009;
revised 9 September 2009;
accepted 3 November 2009

Abstract

This article reports on the construction and examination of a 120 item version (NEO-PI-R-SF) of the Dutch/Belgian NEO-Personality Inventory-Revised (NEO-PI-R). The item selection was based on criteria of appropriateness for older adults and psychometric contribution. The factor structure of the NEO-PI-R-SF proved highly equivalent to the parent instrument, indicating concordant validity. Coefficients alpha were generally lower due to the reduction of items. When assessed in a new sample of older adults (N=794), the factor structure of the NEO-PI-R-SF proved replicable, except for the facet scales A1: Trust and A3: Altruism. Coefficients alpha for the domain scales were satisfactory, while the coefficients alpha for the facet scales were marginal to satisfactory and below those found in two adult samples (N=1305 and N=682). It was concluded that the NEO-PI-R-SF may be a time-saving alternative for research when a fine-grained description of personality among older adults is required. *Copyright © 2011 John Wiley & Sons, Ltd.*

Introduction

This article examines the feasibility and validity of an abbreviated version of the Dutch/Belgian NEO-Personality Inventory-Revised (NEO-PI-R) (Hoekstra *et al.*, 2003, 1996) adapted for use with older adults, while preserving its original psychometric properties, structure and validity.

The NEO-PI-R (Costa and McCrae, 1992) has been extensively validated across the life span (McCrae and Costa, 2003; McCrae *et al.*, 1999, 2000), but in most studies individuals over the age of 70 are under-represented (Terracciano *et al.*, 2005; Weiss and Costa, 2005). Although the NEO-PI-R has been proven to be applicable for use with older adults (≥70) who are vital and well educated (Terracciano *et al.* 2005, 2006), its administration can be demanding for older adults with physical or cognitive constraints. The time needed to complete the 240 items for older adults varies up to three times as long compared to reading proficient younger adults (Roepke *et al.*, 2001; Hoekstra *et al.*, 2003). The shorter 60 items NEO-Five Factor Inventory (NEO-FFI; Costa and McCrae, 1992) is usually preferred to reduce the burden of the assessment, but does not provide information on the facet level traits.

Reise and Henson (2000) used a Computerized Adaptive Test (CAT) algorithm to evaluate the number of items that are needed for a satisfactory recovery of the full-scale facet scores of the NEO-PI-R and found that by selecting the best four items per facet scale, it is feasible to reduce the facet scales in half with little loss of precision. McCrae *et al.* (2005) improved the readability of the

NEO-PI-R for young adolescents and middle scholars (the NEO-PI-3), and subsequently used conventional regression analyses for their reduced-in-half version, the NEO-PI-First Half (NEO-PI-FH). They demonstrated that the original structure could be retained, but that there was 10% loss of predictive scale validity and that internal consistency decreased.

Another point of interest when using the NEO-PI-R with older adults is that the items were not specifically formulated or validated for use with older adults. Mroczek *et al.* (1999) pointed to the importance of age-appropriate items, and stated that behavioral manifestations of the same underlying construct are unique at different ages and may be influenced by psychosocial, physical and cognitive changes. At an advanced age one can refer to a wide range of life experiences (Widiger and Seidlitz, 2002). This may be accentuated when items, that are not appropriate to describe behavior at old age, force older adults to refer to earlier life stages (Clarkin *et al.*, 1999), for example: "*I like the excitement of the rollercoaster*" or "*I am a workaholic*".

In the present study we selected four items per facet scale according to criteria of appropriateness and psychometric contribution for use with older adults.

Item reduction to 120 items would be considered successful if the NEO-PI-R-Short Form (NEO-PI-R-SF) factor structure showed equivalence to the full scale Dutch/Belgian NEO-PI-R. Based on the results reported by McCrae and Costa (2007) we expected lower coefficients alpha because of the item reduction and predictive scale validity to decrease by about 10%.

Method

Participants and procedure

For the initial item reduction process we used a mixed age sample of Dutch citizens (Sample A) to prevent the item selection being influenced by ageism. We then examined equivalence of the reduced item set (NEO-PI-R-SF) in an age-homogeneous sample (Sample B). Finally, the NEO-PI-R-SF was administered to a new sample of older adults living in the Netherlands (Sample C) to analyze the psychometric properties.

Sample A and B originally served to standardize the Dutch/Belgian adaptation of the NEO-PI-R. For Sample A the addresses of 4800 inhabitants age >17 were randomly selected from municipal registers in the Netherlands and the Dutch speaking region of Belgium proportional to the number of residents in these communities. Selected persons were sent the NEO-PI-R by post mail along with an explanation of the research and an inquiry form for demographic characteristics. Participation was anonymous and not rewarded. The sample of 682 responders (age 17–83; mean [*M*]=43.7; standard deviation [SD]=15.2; female=51.9%) proved representative for the general population with respect to gender and age, though the educational level was higher in Sample A (Hoekstra *et al.*, 2003).

Sample B (*N*=1305; age 16–43; *M*=23.1; SD=2.6; female=52.9%) consisted of 934 academic and college students who participated voluntarily in a study on the prediction of employment status (see De Fruyt and Mervielde, 1999) and 371 psychology students who were administered the NEO-PI-R during a course on personality psychology. The total response rate for Sample B was 54%.

Sample C consisted of 794 older members (age ≥ 60) of two unions for older adults within the Netherlands, the Protestants-Christelijke Ouderen Bond (PCOB, the Protestant Association of Seniors) and the Algemene Nederlandse Bond voor Ouderen (ANBO, the Dutch Association for Seniors) that have comparable aims and activities. The computer randomly selected addresses of 1750 PCOB and 500 ANBO members in proportion to the geographical regions covered by each union. These numbers were determined by the permission for cooperation of each union and did not reflect the distribution in the population. The members were sent a letter explaining the purpose and the procedure of the research, a questionnaire to obtain demographic information, the NEO-PI-R-SF, and a rating scale on depression. In total, 826 questionnaires were returned (PCOB: 639 [36.5%]; ANBO: 187 [37.4%]). Questionnaires with no more than two missing values, one per facet scale, were considered acceptable, resulting in 794 questionnaires (PCOB: 616 [35.2%]; ANBO: 178 [35.6%]) that were suitable for analysis. Missing values were imputed as the mean item score per facet scale after item-recoding. Demographic characteristics of the PCOB members (age 61–98; *M*=77.2; SD=5.9) from Sample C are indicated in Table 1. We were not able to collect demographic data from the ANBO members due to privacy regulations.

Instrument

The authorized Dutch adaptation of the NEO-PI-R (Hoekstra *et al.*, 1996, 2003) is a 240-item self-report questionnaire. The five major domain scales: Neuroticism (N), Extraversion (E), Openness to experience (O), Agreeableness (A), and Conscientiousness (C) are each composed of six facet scales, designed to measure lower-order traits. For example, "Neuroticism" is composed

Table 1 Characteristics of Sample A ($N = 682$), Sample B ($N = 1305$), and Sample C (PCOB members; $N = 616$)

	Sample A	Sample B	Sample C (PCOB)
Variable	*N* (%)	*N* (%)	*N* (%)
Gender			
Female	350 (51.3)	690 (52.9)	290 (47.1)
Male	325 (47.7)	615 (47.1)	307 (49.8)
Age			
15–30	156 (23.2)	1069 (84.1)	
31–40	151 (22.1)	19 (1.6)	
41–50	144 (21.1)	4 (0.3)	
51–60	107 (15.8)		
61–70	83 (12.1)		81 (13.4)
71–80	31 (4.4)		340 (56.3)
81–90	1 (0.1)		177 (29.3)
91–100			6 (1.0)
Education			
≤10years	206 (30.2)		230 (37.4)
11–15years	204 (29.9)		337 (54.7)
>15years	224 (32.8)	1305 (100)	40 (6.5)
Marital status			
Married	—	—	353 (57.3)
Single	—	—	39 (6.3)
Divorced	—	—	14 (2.3)
Widowed	—	—	195 (31.7)
Cohabited	—	—	6 (1.0)
Habitat			
Belgian	205 (30.1)	1305 (100)	
Dutch	477 (69.9)		616 (100)

Note: Percentages do not sum up to 100% because of missing data.

of the facet scales N1: Anxiety, N2: Hostility, N3: Depression, N4: Self-consciousness, N5: Impulsiveness and N6: Vulnerability. Items are balanced to control for the effect of acquiescence and scored on a five-point Likert scale ranging from "strongly agree" to "strongly disagree".

The psychometric properties and the factor structure of the Dutch/Belgian NEO-PI-R closely resemble those of the American NEO-PI-R (Costa and McCrae, 1992), except for minor deviations for the facet scales E3: Assertiveness and N5: Impulsiveness. Median coefficients alpha across different samples for the domain scales range from 0.92 for N to 0.86 for O, and those for the facet scales from 0.83 for N1: Anxiety to 0.57 for O6: Values (Hoekstra *et al.*, 1996, 2003).

Procedure for item selection

Item reduction was accomplished following a three step process.

(1) Items eliminated by clinician judgment: Three clinicians evaluated the appropriateness for older adults of the 240 NEO-PI-R items based on irrelevance of the content (for example item 172: "I like the excitement of the roller coaster") (Dutch) vocabulary complexity (for example item 230: "I tend to be cynical and skeptical about others' intentions"), and complex/elaborate (Dutch) syntax (for example item 123: "Human needs should always have priority over economic considerations"). Complexity due to negative formulation was ignored in order to maintain scale balance and to control for effects of response acquiescence. Items were omitted if at least two clinicians considered them inappropriate.

(2) Items eliminated based on item/total correlation: Next, Sample A (Hoekstra *et al.*, 2003) was used to further reduce items per facet scale. The item with the poorest item/total correlation was omitted. Then, the item/total correlation was computed again for the remaining items to select the next item for omission until four items were left per facet scale (NEO-PI-R-SF).

(3) Fine-tuning item selection based on the factor structure: The factor structures of the NEO-PI-R-Short Form and the full scale version were analyzed by Principal Components Analysis (PCA) followed by Varimax rotation. Next, the selection of the NEO-PI-R-SF item set was fine-tuned by substitution of items to accomplish optimal replication of the NEO-PI-R factor structure. This procedure will be illustrated by an example. The NEO-PI-R-SF facet scale with the strongest primary loading (O3: Feelings) on a non-intended factor (N: Neuroticism) was selected for item revision. Then we evaluated an inter-item correlation matrix of O3 and the N facet scale that correlated most strongly with O3 (N5: Impulsiveness). The item of O3 that correlated strongest with the items of N5 was substituted for a previously omitted item. Previously judged inappropriate items (under Step 1) were selected for exchange only if none of the psychometrically omitted items (under Step 2) improved the factor structure. After an item was substituted, the NEO-PI-R-SF factor structure was examined again and the complete procedure was repeated until all incongruent primary loadings

were examined. Next, the procedure was repeated for positive and negative secondary loadings ≥0.40, that exceeded the corresponding NEO-PI-R loading (>0.10 difference), until item substitution did not further improve replication of the NEO-PI-R factor structure.

Data-analytic strategy

Replication of the NEO-PI-R factor structure in the truncated NEO-PI-R-SF item set was tested on Sample B by PCA followed by Varimax rotation. To examine the equivalence of the short and the full scale factor structures, Tucker's congruence (TC) coefficients were computed for the domain and facet scales in Samples A and B. TC coefficients ≥ 0.85 indicate fair similarity and TC coefficients ≥ 0.95 imply perfect similarity (Lorenzo-Seva and ten Berge, 2006). TC coefficients ≥ 0.85 were interpreted as acceptable for replication of the factor structure, while TC coefficients below 0.85 indicate different constructs. To test whether item reduction influenced internal consistency of the NEO-PI-R-SF domain and facet scales, the coefficients alpha were converted to 240-item-coefficients alpha using the Spearman Brown formula (Nijdam, 2003). Subsequently, the converted coefficients alpha were compared to those of the NEO-PI-R scales using the alpha test (Hox, 2006). This program tests the difference between two or more coefficients alpha while considering the number of items and the sample sizes. Accordingly, we tested whether the internal consistency per age group was influenced by item

Table 2 Items from the NEO-PI-R selected for omission by at least two clinicians due to irrelevant subjects, complex vocabulary, or complex formulation

	Irrelevant subject	Complex vocabulary	Complex formulation
Neuroticism			
N2 Angry hostility			216
N4 Self-consciousness			166
N6 Vulnerability	26, 56		
Extraversion			
E1 Warmth	182, *212*		
E2 Gregariousness	*187*, 217		
E3 Assertiveness		192	
E4 Activity	17, 167, *197*[a]		
E5 Excitement seeking	*52*, *172*	82	
Openness			
O1 Fantasy			33, *123*
O2 Aesthetics		8, 98	
O4 Actions	108		
O5 Ideas		23, 53, *113*	173
O6 Values			28, 88, 148,[a] 178
Altruism			
A1 Trust		*4*	
A2 Straightforwardness		39	
A3 Altruism		14[a]	
A4 Compliance	*19*	49	
A6 Tender mindedness		59[a]	89, 149
Conscientiousness			
C1 Competence		*215*	
C2 Order	190		
C3 Dutifulness	*15*,[a] *225*	*165*	
C4 Achievement striving	*140*	*230*	*50*, 80
C5 Self-discipline	*85*, 145	*175*	

Note: Items selected by three clinicians are shown in italic typeface.
[a]Items that eventually were retained for their psychometric contributions.

Int. J. Methods Psychiatr. Res. 20(3): 135–144 (2011). DOI: 10.1002/mpr

reduction or age-related response homogeneity in two subsamples of Sample A (age <60 and age ≥ 60). Finally, we examined the psychometric properties of the NEO-PI-R-SF when actually assessed within a general population of older adults (Sample C).

Results

Item selection

First, items were judged on their appropriateness for use with older adults. Table 2 shows the 48 items that were judged inappropriate by at least two clinicians; of these, 18 items were selected by all three clinicians (proportional agreement >0.78).

Next, another 72 items were omitted based on the poorest item-total correlation per facet scale, resulting in 120 remaining items equally divided over 30 facet scales. Finally, items that contributed poorly to the factor structure were substituted according to a fine tuning item selection procedure (see Method section). Five replacement items were judged inappropriate for older adults by the clinicians, but had to be retained for their psychometric contribution: item 14: "Some people think I am selfish and self-centered" (*Complex* Dutch language), item 15: "I try to fulfill all tasks that are assigned to me conscientiously" (*Irrelevance*), item 59: "I am hard-headed and tough-minded in my attitudes" (*Complex* Dutch language), item 148: "I believe that it's better to stick to your own principles than to be open-minded" (*Complex* Dutch language) and item 197: "I have a hurried life" (*Irrelevance*). Item substitution was carried out until optimal replication the NEO-PI-R-SF and the full scale factor structure was achieved.

Equivalence of the factor structure

Equivalence of the NEO-PI-R-SF factor structure to the full scale is highly indicative that construct validity of the full scale is maintained. TC coefficients for the domain scales varied from 0.95 for O to 0.98 for N, which implied similarity (TC ≥ 0.95; see Lorenzo-Seva and ten Berge, 2006) for the domain scales. For six facet scales TC coefficients fell within the range 0.85–0.94, indicating fair similarity, while the remaining facet scales could all be considered similar (TC ≥ 0.95). However, it should be noted that the results represent an upper bound of equivalence because analyses are conducted on the same sample. As in the original NEO-PI-R factor structure in Dutch/Belgian samples, E3: Assertiveness primarily loaded on N. For N5: Impulsiveness NEO-PI-R-SF loadings were comparable for N (0.42) and E (0.40), whereas the full scale N5 loaded primarily on E (0.50) and secondarily on N (0.40). The total variance explained was reduced by 8.7% to 51.5% for the NEO-PI-R-SF, which is still satisfactory.

Next, replication of the factor structure was examined within Sample B ($N=1305$). Figure 1 shows that Exploratory Factor Analysis revealed a five-factor structure based on Catell's scree test (Catell, 1966). Again, the total variance explained was sufficient (51.3%) and the NEO-PI-R factor structure was retained. Table 3 shows that all but two TC coefficients implied similarity (≥ 0.95). TC coefficients for O5: Ideas (0.88) and A5: Modesty (0.94) indicated fair similarity.

Internal consistency

As expected, internal consistency coefficients were generally lower for the NEO-PI-R-SF than for the NEO-PI-R. Internal consistency depends on the specificity of the construct measured and thus the spectrum of items (Saucier, 1998), the number of items, and response homogeneity of the population assessed (Nijdam, 2003). To examine whether the decrease in coefficients alpha was caused by the reduction of items, we converted the NEO-PI-R-SF alphas into full scale alphas using the Spearman–Brown formula. All of the converted coefficients alpha were equal to or exceeded those of the NEO-PI-R in Sample A, indicating that the decrease in internal consistency could be fully attributed to the item

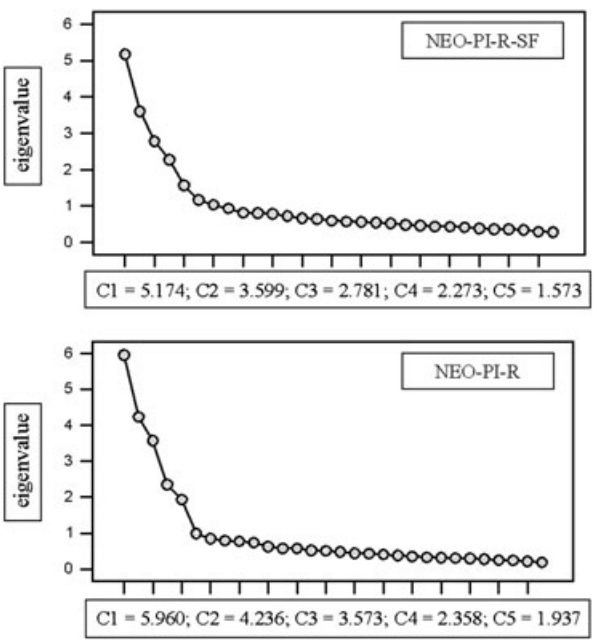

Figure 1 Exploratory Factor Analysis of facet scores for the NEO-PI-R-SF and the NEO-PI-R ($N = 1305$).

Table 3 Tucker's congruence (TC) coefficients and coefficients alpha for the NEO-PI-R-SF (NEO-SF) and the NEO-PI-R (NEO) for Sample A (N=682) and Sample B (N=1305)

	Sample A			Sample B		
		Coefficient alpha			Coefficient alpha	
	TC	NEO-SF	NEO	TC	NEO-SF	NEO
Neuroticism	0.98	0.88	0.92	0.98	0.86	0.92
Extraversion	0.97	0.84	0.88	0.97	0.86	0.90
Openness	0.95	0.78	0.88	0.96	0.77	0.88
Altruism	0.97	0.78	0.86	0.98	0.81	0.89
Conscientiousness	0.97	0.85	0.89	0.98	0.88	0.92
N1 Anxiety	1.00	0.77	0.83	0.99	0.75	0.84
N2 Angry hostility	1.00	0.61	0.73	0.97	0.61	0.75
N3 Depression	0.99	0.70	0.80	0.98	0.73	0.82
N4 Self-consciousness	0.96	0.71	0.76	1.00	0.39	0.62
N5 Impulsiveness	0.96	0.58	0.70	0.98	0.64	0.74
N6 Vulnerability	1.00	0.75	0.79	0.99	0.73	0.80
E1 Warmth	0.95	0.63	0.69	0.95	0.69	0.72
E2 Gregariousness	0.98	0.68	0.77	0.99	0.74	0.80
E3 Assertiveness	0.97	0.78	0.82	0.99	0.75	0.84
E4 Activity	0.93	0.57	0.68	0.99	0.57	0.70
E5 Excitement seeking	0.98	0.51	0.67	0.95	0.44	0.61
E6 Positive emotions	0.99	0.72	0.77	0.99	0.75	0.80
O1 Fantasy	0.99	0.70	0.79	0.96	0.72	0.83
O2 Aesthetics	0.98	0.70	0.76	0.99	0.73	0.78
O3 Feelings	0.97	0.65	0.70	0.98	0.72	0.76
O4 Actions	0.95	0.47	0.68	0.98	0.44	0.65
O5 Ideas	0.96	0.64	0.70	0.88	0.65	0.73
O6 Values	0.97	0.51	0.64	1.00	0.47	0.58
A1 Trust	0.95	0.77	0.76	0.98	0.78	0.80
A2 Straightforwardness	0.99	0.65	0.70	0.99	0.73	0.77
A3 Altruism	0.97	0.57	0.66	0.98	0.62	0.71
A4 Compliance	0.94	0.59	0.68	0.98	0.55	0.67
A5 Modesty	0.94	0.70	0.72	0.94	0.78	0.79
A6 Tender-mindedness	0.94	0.47	0.57	0.96	0.49	0.61
C1 Competence	0.96	0.58	0.65	0.96	0.38	0.66
C2 Order	0.98	0.66	0.64	0.98	0.68	0.72
C3 Dutifulness	0.93	0.53	0.60	0.99	0.65	0.72
C4 Achievement striving	0.98	0.64	0.76	0.98	0.70	0.80
C5 Self-discipline	0.96	0.65	0.71	0.97	0.78	0.82
C6 Deliberation	0.94	0.75	0.76	0.99	0.78	0.80

Note: Tucker's congruence coefficients; $0.88 \leq TC \leq 0.94$ implies fair similarity; $TC \geq 0.95$ implies perfect similarity (Lorenzo-Seva and ten Berge, 2006).

reduction. Table 3 shows that coefficients alpha for the NEO-PI-R-SF were moderate to satisfactory, but were marginal for N5: Impulsiveness, E4: Activity, E5: Excitement seeking, O4: Actions, O6: Values, A3: Altruism, A6: Tender-mindedness, and C3: Dutifulness. Internal consistency remained stable at the domain level for N (0.92 and 0.88), E (0.88 and 0.84), and C (0.89 and 0.85), but was somewhat lower for O (0.88 and 0.78) and A (0.86 and 0.78) for the NEO-PI-R and the NEO-PI-R-SF respectively.

In Sample B the notable decline of the NEO-PI-R-SF coefficients alpha for C1: Competence and N4:

Int. J. Methods Psychiatr. Res. 20(3): 135–144 (2011). DOI: 10.1002/mpr

Self-consciousness could be attributed to response homogeneity for item 35 ("I have most things in my life well organized") and item 46 ("I seldom feel timid around company") respectively. Both items performed poorly in both versions (item 35: item/total correlation: NEO-PI-R-SF= 0.03; NEO-PI-R=0.09; M=3.75; SD=0.91; item 46: item/total correlation: NEO-PI-R-SF=−0.17; NEO-PI-R=−0.18; M= 3.51; SD=0.97) indicating that these items discriminated insufficiently when applied in Sample B. Again, the NEO-PI-R-SF coefficients alpha for E5: Excitement seeking, O4: Actions, O6: Values, and A6: Tender-mindedness were poor.

To examine whether reliability of the NEO-PI-R-SF was influenced by item reduction or age-related response homogeneity, Sample A was divided into two sub-samples of 122 older adults (age 60–83; M=67.1; SD=5.5; female = 40.2%) and 551 younger adults (age 17–59; M=38.6; SD = 11.4; female=54.5%). According to the scheme presented by Figure 2, differences in coefficients alpha were evaluated across age groups for the NEO-PI-R (I) and the NEO-PI-R-SF (II), and across both instruments per age group: age < 60 (III) and age ≥ 60 (IV). Significant differences in coefficients alpha across instruments (III and IV) were considered to be a result of item reduction. Significant differences across age groups for the NEO-PI-R (I) were considered to be a result of age-related response homogeneity, while those for the NEO-PI-R-SF (II) were interpreted as a combined effect of item reduction and age-related response homogeneity.

For N3: Depression coefficients alpha in the subsample of older adults were below those found in the group of younger adults. Differences reached significance for both instruments (NEO-PI-R: p=0.02 and NEO-PI-R-SF: p=0.00), but coefficients alpha did not differ across instruments per age group (<60: p=0.06 and ≥60: p= 0.40). Item inspection revealed that item 131 ("I usually blame myself for mistakes") performed poorly with older adults because of response homogeneity (item/total correlation: NEO-PI-R-SF=0.10; NEO-PI-R=0.17; M= 3.59; SD=0.84). For E1: Warmth and E4: Activity only differences in NEO-PI-R-SF coefficients alpha across both age groups (II) were significant (E1: p=0.01 and E4: p= 0.03), indicating that response homogeneity among older adults is expressed more emphatic in the short E1 and E4 facet scales. For the remaining NEO-PI-R-SF coefficients alpha that were marginal (α<0.65), no significant differences were found and these were considered within the normal range of item reduction.

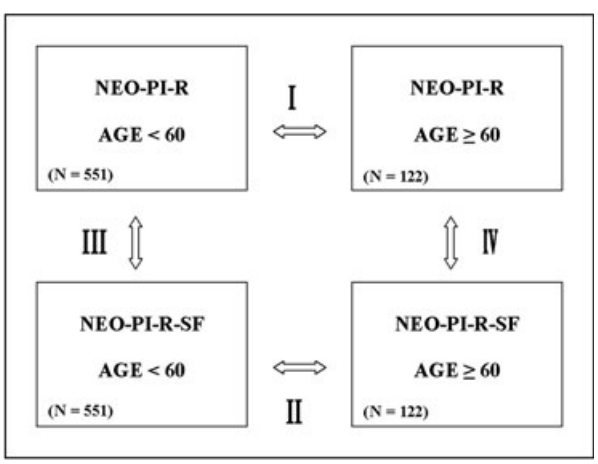

Figure 2 Scheme of analysis of item compilation and response homogeneity.

Replication in a new older adult sample

Finally, we examined the psychometric properties of the NEO-PI-R-SF when actually assessed with older adults (Sample C). To maximize similarity to the original factor structure, we extracted five factors in a PCA with Varimax rotation. The results of the factor analysis (Table 4) show that the full scale factor structure was well retained. All but two facet scales had primary loadings on the intended factor. A1: Trust primarily loaded on O (0.38) and A3: Altruism on E (0.48). Most secondary loadings reported in Table 4 were consistent with the original factor structure. Predictive scale validity was 49.5%, which is slightly below the variance explained in Samples A and B. Internal consistency was sufficient to high for the domain scales and moderate to satisfactory for the facet scales. For O6: Values and A6: Tender mindedness the coefficients alpha were unacceptably low. Generally, most coefficients alpha were below those found for Samples A and B, but were in line with the coefficients alpha found for older adults within Sample A.

Discussion

The primary objective of this study was to construe a reduced NEO-PI-R item-set for use with older adults, while retaining its psychometric properties, structure, and validity. We tested internal convergent and divergent validity of the NEO-PI-R-SF facet scales by factor analysis. The NEO-PI-R-SF factor structure proved highly equivalent to the NEO-PI-R factor structure and replicable in adult samples, suggesting concordant construct validity. In accordance with the results McCrae and Costa (2007) reported on their short version of the NEO-PI-3,

Table 4 Principal Components Analysis (PCA) with Varimax rotation and coefficients alpha for the NEO-PI-R-SF in Sample C (*N*=794)

	Neuroticism, N	Extraversion, E	Openness, O	Altruism, A	Conscientiousness, C	Coefficient alpha
N1 Anxiety	*0.77*	−0.16	−0.05	0.03	−0.09	0.72
N2 Angry hostility	*0.65*	−0.03	−0.16	−0.40	−0.02	0.45
N3 Depression	*0.76*	−0.14	0.07	0.03	−0.18	0.60
N4 Self-consciousness	*0.66*	−0.30	−0.02	0.05	−0.14	0.67
N5 Impulsiveness	*0.45*	−0.03	0.12	−0.45	−0.10	0.54
N6 Vulnerability	*0.55*	−0.10	−0.21	0.13	−0.46	0.74
E1 Warmth	−0.19	*0.66*	0.19	0.04	0.15	0.54
E2 Gregariousness	−0.29	*0.67*	0.06	−0.06	−0.18	0.55
E3 Assertiveness	−0.38	*0.39*	0.04	−0.32	0.15	0.78
E4 Activity	−0.17	*0.53*	0.14	−0.37	0.35	0.54
E5 Excitement seeking	0.06	*0.32*	0.21	−0.51	−0.05	0.47
E6 Positive emotions	−0.27	*0.50*	0.27	0.02	0.22	0.66
O1 Fantasy	−0.01	0.18	*0.62*	−0.15	−0.09	0.52
O2 Aesthetics	0.11	0.12	*0.59*	0.06	0.07	0.58
O3 Feelings	0.38	0.25	*0.52*	0.07	0.20	0.55
O4 Actions	−0.29	0.13	*0.43*	−0.28	−0.24	0.43
O5 Ideas	−0.14	−0.06	*0.47*	−0.12	0.40	0.56
O6 Values	−0.20	−0.04	*0.53*	0.01	−0.27	0.35
A1 Trust	−0.09	0.20	0.38	0.22	0.27	0.66
A2 Straightforwardness	−0.06	−0.03	0.03	*0.65*	0.14	0.46
A3 Altruism	0.09	0.48	0.05	*0.34*	0.34	0.60
A4 Compliance	−0.13	−0.06	0.10	*0.62*	−0.02	0.45
A5 Modesty	0.21	0.14	−0.16	*0.64*	−0.03	0.61
A6 Tender-mindedness	0.27	0.38	0.24	*0.48*	0.01	0.35
C1 Competence	−0.27	0.05	0.15	−0.33	*0.62*	0.54
C2 Order	−0.20	0.10	−0.16	0.28	*0.56*	0.55
C3 Dutifulness	0.08	0.16	−0.05	0.28	*0.62*	0.59
C4 Achievement striving	−0.09	0.27	0.06	−0.32	*0.64*	0.55
C5 Self-discipline	−0.29	0.32	−0.13	0.05	*0.60*	0.61
C6 Deliberation	−0.07	−0.14	0.02	0.16	*0.67*	0.67
Coefficient alpha	0.86	0.81	0.71	0.74	0.83	

Note: Loadings of facet scales >0.30 on the intended factor are shown in italic typeface.

predictive scale validity decreased about 9%. Except for deviant loadings for A1: Trust and A3: Altruism, the factor structure proved well replicable in a new sample of older adults. This is consistent with studies that found some abnormalities in the NEO-PI-R factor structure in older adult samples (Roepke *et al.*, 2001; Small *et al.*, 2003; Weiss and Costa, 2005). These irregularities may indicate that some personality traits become more fluid late in life and that stability of personality, as amply demonstrated with the NEO-PI-R, is veiling underlying variability in behavioral manifestations of traits across the life span (Mooi *et al.*, 2006). As yet the NEO-PI-R-SF facet scales A1: Trust and A3: Altruism of the NEO-PI-R-SF should be cautiously used in older adults.

Reciprocal dependency between internal consistency and content range is implicit in a heterogenic construct such as personality (Saucier, 1998). In accordance with our assumption, internal consistency was lower for the NEO-PI-R-SF due to item reduction. Coefficients alpha remained satisfactory for the NEO-PI-R-SF domain scales and most of the facet scales, but were marginal for eight of the facet scales in the adult samples. Within Sample A internal consistency for older and younger adults proved comparable. Although coefficients alpha for the domain scales were sufficient in the new older adult sample (Sample C), these were marginal to poor for 15 facet scales, and unacceptable for two facet scales. Considering the differences in age between the older adult samples

Int. J. Methods Psychiatr. Res. 20(3): 135–144 (2011). DOI: 10.1002/mpr

(Sample C: age 61–98; $M=77.2$; SD=5.9; sub-Sample A: age 60–83; $M=67.1$; SD=5.5), the decrease in internal consistency in Sample C may indicate that aging is associated with an increase of response homogeneity in later life. Nevertheless, internal consistency of short versions should meet the regular standards (Smith *et al.*, 2000) and use of facet scales with marginal coefficients alpha should be considered very carefully. In this perspective test–retest reliability may be an useful additive indicator to determine reliability of the NEO-PI-R-SF and its applicability with older adults.

Since abbreviating the NEO-PI-R implies loss of content range and impacts upon psychometric qualities, use of the full length NEO-PI-R is always preferred if feasible. However, the NEO-PI-R-SF seems a welcome alternative for use with less vigorous populations, and may provide researchers an opportunity to examine personality and adaptation to life events in late life at the facet level. Although the item reduction in this study met the relevant conditions for short-form development as stated by Smith *et al.* (2000), it also had some methodological limitations.

First, relying on clinicians to judge the appropriateness of items does not guarantee that the omitted items would indeed perform poorly or would adversely affect the psychometric properties of the scales. According to the procedure followed by McCrae *et al.* (2005), making the NEO-PI-R more readable for younger respondents, an alternative method is to rely on item judgments of older adults. Our assumption, that some diagnostic expertise is necessary to judge the relevance of items adequately, was confirmed prior to this study, when we asked several older adults to judge items of the NEO-PI-R. They consistently omitted items that were susceptible for ageism (for example all items of E4: Activity), which is not desirable for an instrument that is construed to measure development of personality throughout the adult life span. Another alternative method for item selection, relying on the psychometric contribution of items in a NEO-PI-R sample of solely older adults, would be costly, time consuming and also bears the risk to be determined by ageism. Second, Sample C may not be optimal representative for the general population of older adults in the Netherlands because of selective non-response. For example, it is assumable that responders are generally more vital and younger, and have less cognitive restraints than non-responders. Therefore, applicability of the NEO-PI-R-SF must be evaluated carefully when used with older adults, especially those with cognitive or physical constraints (see APA, 2004). It is not plausible that selective non-response have impacted on the replication of the NEO-PI-R factor structure in Sample C, for older adults who were not capable of administering the NEO-PI-R-SF would neither be able to administer the NEO-PI-R. Future research should clarify for which subgroups of older adults the NEO-PI-R-SF is preferable to the NEO-PI-R (Smith *et al.*, 2000), and what the minimal assessment capacities necessary for its administration are.

For researchers the NEO-PI-R-SF seems an alternative instrument to examine personality and adaptation at the facet level in older adults when a fine-grained description of personality is warranted and use of the NEO-PI-R is not feasible.

Acknowledgements

This study was supported, in part, by Grant 5.0.1.111104 from the Christian Association for the Care of the Mentally Ill (Vereniging Christelijke Verzorging Geestelijke Zieken: VCVGZ) in the Netherlands. The authors are most grateful to two Dutch associations for older adults, the Protestant Association of Seniors (Protestants Christelijke Ouderen Bond: PCOB) and the Dutch Association for Seniors (Algemene Nederlandse Bond voor Ouderen: ANBO), for facilitating the data collection among older adults. They are thankful to Dr Robert R. McCrae of the National Institute on Aging, Baltimore, MD, for his valuable and constructive comments on an earlier version of the manuscript, and to Professor Dr Jos M.F. ten Berge of the University of Groningen, the Netherlands for computing the Tucker's congruence coefficients.

Declaration of interest statement

H.A. Hoekstra and F. De Fruyt both receive royalties from the Dutch adaptation of the NEO-PI-R.

References

American Psychological Association (APA) (2004) Guidelines for Psychological Practice with Older Adults, vol. **59**(4), pp. 236–260, Washington, DC, APA. DOI: 10.1037/0003-066X.59.4.236

Catell R.B. (1966) The scree test for number of factors. *Multivariate Behavioral Research*, **1**(2), 245–276.

Clarkin J.F., Spielman L.A., Klausner E. (1999) Conceptual overview of personality disorders in the elderly. In Rosowsky E., Abrams R.C., Zweig R.A. (eds) *Personality Disorders in Older Adults*, pp. 3–15, London: L.E.A. Publishers.

Costa P.T., McCrae R.R. (1992) Revised Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI): Professional Manual, Odessa, FL, Psychological Assessment Resources.

De Fruyt F., Mervielde I. (1999) RIASEC types and Big Five traits as predictors of employment status and nature of employment. *Personnel Psychology*, **52**(3), 701–727. DOI: 10.1111/j.1744-6570.1999.tb00177.x

Hoekstra H.A., De Fruyt F., Ormel J. (2003) NEO-PI-R/NEO-FFI: Big Five Personality Inventory-Manual, Lisse, Swetz & Zeitlinger (in Dutch).

Hoekstra H.A., Ormel J., De Fruyt F. (1996) NEO-PI-R/NEO-FFI: Big Five Personality Inventory-Manual, Lisse, Swetz & Zeitlinger (in Dutch).

Hox J. (2006) http://www.geocities.com/joophox/papers/papers [1 July 2006].

Lorenzo-Seva U., ten Berge J.F. (2006) Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology*, **2**(2), 57–64. DOI: 10.1027/1614-2241.2.2.57

McCrae R.R., Costa P.T. Jr (2003) Personality in Adulthood, New York, The Guilford Press.

McCrae R.R., Costa P.T. Jr (2007) Brief versions of the NEO-PI-3. *Journal of Individual Differences*, **28**(3), 116–128. DOI: 10.1027/1614-0001.28.3.116

McCrae R.R., Costa P.T. Jr, Lima M.P., Simões A., Ostendorf F., Angleitner A., Marušić I., Bratko D., Caprara G.V., Barbaranelli C., Chae J.-H., Piedmont R.L. (1999) Age differences in personality across the adult life span: Parallels in five cultures. *Developmental Psychology*, **35**(2), 466–477. DOI: 10.1037/0012-1649.35.2.466

McCrae R.R., Costa P.T. Jr, Martin T.A. (2005) The NEO-PI-3: A more readable revised NEO Personality Inventory. *Journal of Personality Assessment*, **84**(3), 261–170. DOI: 10.1207/s15327752jpa8403_05

McCrae R.R., Costa P.T. Jr, Ostendorf F., Angleitner A., Hřebíčková M., Avia M.D., Sanz J., Sánchez-Bernardos M.L., Kusdil M.E., Woodfield R., Saunders P.R., Smith P.B. (2000) Nature over nurture: temperament, personality, and life span development. *Journal of Personality and Social Psychology*, **78**(1), 173–186.

Mooi B., Comijs H.C., Beekman A.T.F., Kerkhof A.J.F.M. (2006) Stability of personality in late life. *Tijdschrift voor Gerontologie en Geriatrie*, **37**(4), 136–141 (in Dutch).

Mroczek D.K., Hurt S.W., Berman W.H. (1999) Conceptual and methodological issues in the assessment of personality disorders in older adults. In Rosowsky E., Abrams R.C., Zweig R.A. (eds) *Personality Disorders in Older Adults*, pp. 135–150, London: L.E.A. Publishers.

Mroczek D.K., Spiro A. (2003) Modeling intraindividual change in personality traits: findings from the normative aging study. *Journal of Gerontology: Psychological Sciences*, **58B**(3), 153–165.

Nijdam A.D. (2003) Statistics for Research: Descriptive Techniques. Groningen, Wolters-Noordhoff bv (in Dutch).

Reise S.P., Henson J.M. (2000) Computerization and adaptive administration of the NEO PI-R. *Assessment*, **7**(4), 347–364. DOI: 10.1177/107319110000700404

Roepke S., McAdams L.A., Lindamer L.A., Patterson T.L., Jeste D.V. (2001) Personality profiles among normal aged individuals as measured by the NEO-PI-R. *Aging & Mental Health*, **5**(2), 159–164. DOI: 10.1080/13607860120038339

Saucier G. (1998) Replicable item-cluster subcomponents in the NEO Five-Factor Inventory. *Journal of Personality Assessment*, **70**(2), 263–276. DOI: 10.1207/s15327752 jpa7002_6

Small B.J., Hertzog C., Hultsch D.F., Dixon R.A. (2003) Stability and change in adult personality over 6years: Findings from the Victoria Longitudinal Study. *Journal of Gerontology: Psychological Sciences*, **58B**(3), 166–176.

Smith G.T., McCarthy D.M. Anderson K.G. (2000) On the sins of short-form development. *Psychological Assessment*, **12**(1), 102–111. DOI: 10.1037/1040-3590.12.1.102

Terracciano A., McCrae R.R., Brant L.J., Costa P.T. Jr (2005) Hierarchical linear modeling analyses of the NEO-PI-R scales in the Baltimore Longitudinal Study of Aging. *Psychology and Aging*, **20**(3), 493–506. DOI: 10.1037/0882-7974.20.3.493

Terracciano A., McCrae R.R., Costa P.T. Jr (2006) Longitudinal trajectories in Guilford-Zimmerman temperament survey data: Results from the Baltimore longitudinal study of aging. *Journal of Gerontology: Psychological Sciences*, **61B**(2), 108–116.

Weiss A., Costa P.T. Jr (2005) Domain and facet personality predictors of all-cause mortality among medicare patients aged 65 to 100. *Psychosomatic Medicine*, **67**(5), 724–733. DOI: 10.1097/01.psy.0000174995. 96183.9b

Widiger T.A., Seidlitz L. (2002) Personality, psychopathology, and aging. *Journal of Research in Personality*, **36**(4), 335–362. DOI: 10.1016/S0092-6566.2802.2900013-2

Int. J. Methods Psychiatr. Res. 20(3): 135–144 (2011). DOI: 10.1002/mpr