To know personality is to measure it

Introducing a Dutch brief form of the Multidimensional Personality Questionnaire

Eigenhuis, A.

Chapter 2

Development and validation of the Dutch brief form of the Multidimensional Personality Questionnaire

Annemarie Eigenhuis

Jan H. Kamphuis

Arjen Noordhof

An excerpt of this chapter is published as:
Eigenhuis, A., Kamphuis, J. H., & Noordhof, A. (2013). Development and validation of the Dutch brief form of the Multidimensional Personality Questionnaire (MPQ-BF-NL). *Assessment*, *20*(5), 565–575. http://doi.org/10.1177/1073191112444920

Abstract

This study describes the development and psychometric properties of the Dutch brief form of the Multidimensional Personality Questionnaire (MPQ-BF-NL). Representative samples from the Netherlands (*N* = 1,055) and the U.S. (*N* = 1,153), and a Dutch student sample (*N* = 987) were used for development, cross- and external validation, respectively. Our strategy for item selection and scale validation replicated the development of the U.S. brief form (MPQ-BF; Patrick et al., 2002). Internal consistencies were generally good, and comparable to the U.S. version, as were correlations with the U.S. full length scales and higher-order structure. Moreover, convergent and divergent patterns were consistent with prediction, with Positive Emotionality related to social and activating behavior; Negative Emotionality to anxiety; and Constraint to reversed impulsivity and externalizing behaviors. In sum, the MPQ-BF-NL provides the Dutch-Flemish language area with a personality inventory well-suited for both psychopathology research and clinical practice, and offers new opportunities for fundamental and cross-cultural studies on personality.

Introduction

The Multidimensional Personality Questionnaire (MPQ; Tellegen, 1982; Tellegen & Waller, 2008) is a comprehensive hierarchical measure of normal personality variation. It has evolved over many cycles of an exploratory test construction process (Tellegen & Waller, 2008). Hallmark of the exploratory approach to test construction is its bi-directional process in formulating and demarcating concepts: data collection serves the purpose of examining and possibly changing the trait constructs that initially guided the test construction effort. The MPQ provides coverage of a range of traits encompassing the domains of temperament, interpersonal and imaginative style, and behavioral regulation.

Strengths of the MPQ include its interconnection with personality theories as well as with recent conceptualizations in psychopathology research. Another strength of the MPQ is the inclusion of validity scales such as True Response Inconsistency (TRIN) and Variable Response Inconsistency (VRIN) that help identify various response sets. These response sets may invalidate personality inferences, and the ability to make statistically informed judgments on this issue has proven useful in both clinical settings for individual case evaluations (Baer & Miller, 2002; Clegg, Fremouw, Horacek, Cole, & Schwartz, 2010), and in research settings (Graham, Watts, & Timbrooks, 1991; Morasco, Gfeller, & Elder, 2007).

Trait realism and the exploratory approach to test-construction

The MPQ follows from a realist perspective on the nature of traits, i.e., traits exist independently of their definition and make an important difference in life. Adopting this realist trait perspective, one can define a trait as an inferred organismic (psychological, psychobiological) structure underlying an extended family of behavioral dispositions. These dispositions are not meant to be viewed as generalized action tendencies, but as inclinations to behave in certain ways in a set of trait-relevant situations (Tellegen, 1991).

The exploratory approach, guiding the development of the MPQ, takes advantage of the powers of both deduction and induction, by formulating hypotheses about traits, and testing the psychometric consequences of these hypotheses empirically. In the process of developing the MPQ, pre-existing theories of personality have inspired the postulation of multiple specific traits. On several occasions the empirical results did not match the predictions based on these hypothetical constructs (Tellegen & Waller, 2008). Such negative results have been used to revise and improve subsequent hypotheses. Therefore, the MPQ is thoroughly related to falsifiable theories of personality, and its development has

significantly contributed to the empirical testing and revision of such theories. Of course, there is no fundamental reason why this process would have stopped with the latest version of the instrument. For example, cross-cultural research as presented in this paper may contribute to further refinements regarding the universality of the proposed traits.

The structure of the MPQ

The full length U.S. version of the MPQ consists of 276 binary scale items that are meant to measure 11 primary trait scales that coalesce into theoretically meaningful higher-order factors. The higher-order factor explaining Wellbeing (WB), Social Potency (SP), Achievement (AC) and Social Closeness (SC) can be labeled Positive Emotionality (PEM), while Stress Reaction (SR), Aggression (AG) and Alienation (AL) are indicative of Negative Emotionality (NEM). Constraint (CON) is the higher-order factor for Control (CO), Harm Avoidance (HA), and Traditionalism (TR). Absorption (AB) cannot be satisfyingly allocated to any of the three higher-order factors, and consequently has a 'status aparte'. PEM and NEM are aligned to the two factors of affect, Positive Affectivity (PA) and Negative Affectivity (NA; Tellegen, Watson, & Clark, 1999), and are hypothesized to represent individual differences in basic psychobiological systems (Depue, 1996; Depue & Lenzenweger, 2001).

The MPQ and psychopathology

MPQ constructs have been used to develop hypotheses to explain the higher- and lower-order structure of comorbid DSM-IV axis-I and axis-II syndromes. For example, individual differences in NEM have been shown to be related to a broad range of internalizing conditions, while CON is related to externalizing conditions (Krueger et al., 2001). Furthermore, it has been found that low PEM is specifically related to a subset of disorders, including major depressive disorder and social anxiety (Sellbom et al., 2008a). Moreover, the MPQ model has also been profitably used to understand within-syndrome heterogeneity. For example, Miller used the higher-order dimensions to derive subtypes of Post Traumatic Stress Disorder (PTSD), where high NEM predicted PTSD in general, and where low PEM and low CON differentiated between internalizing and externalizing forms of the syndrome respectively (Miller, 2003). Further, McGue and others showed that although high NEM in combination with low CON is indicative of alcohol abuse (McGue et al., 1997) the lower the score of a person on CON, the more likely it is that the person also abuses other drugs (McGue et al., 1999). In general, the MPQ primary trait scales were

slightly superior to other widely used scales in terms of explained variance in psychopathology (Grucza & Goldberg, 2007).

International research on the MPQ-constructs

To our knowledge, only two studies have been published on translations of the MPQ. Ben-Porath and others translated the full version of the MPQ to Hebrew and compared the pattern of loadings in U.S. and Israeli samples (Ben-Porath, Almagor, Hoffman-Chemi, & Tellegen, 1995). Factor loading patterns were quite similar across nations, as were the internal consistencies of the comprising scales. Johnson and others used a German translation of the MPQ to examine cultural differences in personality with the help of Differential Item Functioning (DIF) in an Item Response Theory (IRT) framework. Most differences in mean scale scores found between the U.S. and the German samples turned out to be due to differences in item-difficulties (Johnson, Spinath, Krueger, Angleitner, & Riemann, 2008).

Below, the development and characteristics of a Dutch version of the brief form of the MPQ, the MPQ Brief Form Netherlands (MPQ-BF-NL) are described. Item selection procedures replicated the development of the brief version of the U.S. MPQ, the MPQ Brief Form (MPQ-BF; Patrick et al., 2002), taking a Dutch translation of the full length MPQ as the departure point. 12 items per scale were selected, and including 3 extra items needed for determining scores on validity scales, the resulting instrument consists of 135 items. Developing a brief version has important advantages. From a clinical utility perspective a brief version of 135 items is viewed as preferable since it would take only 15 minutes to complete, instead of the double amount of time for the full version of 276 items. In view of the strong psychometric performance of the U.S. short form, we were hopeful we could develop a useful Dutch version. From a research perspective, it is possible to create efficient proxy measures from the MPQ for various relevant constructs in psychopathology. In the US, proxy measures have been developed for Borderline Personality Disorder (BPD; Bornovalova, Hicks, Patrick, Iacono, & McGue, 2011), and for psychopathy (Walton, Roberts, Krueger, Blonigen, & Hicks, 2008).

Method

Description of samples

Dutch normative sample

For item-selection and normative purposes, a representative Dutch sample consisting of 1060 participants was used. These participants took part in a panel survey conducted by Flycatcher, a full service online research company. The sample was stratified by gender, age, educational level, and county, in order to be representative for the adult Dutch population (Centraal Bureau voor de statistiek, 2009). Five participants were removed who met criteria for exclusion on the Dutch (newly developed) VRIN and/or TRIN scales (see below). The final sample (N = 1,055) consisted of 506 (48%) men and 549 women with a mean age of 45.83 (SD = 15.74).

U.S. cross-validation sample

To test the psychometric performance of the newly developed MPQ-BF-NL in a U.S. sample, we employed 1,153 valid records from participants from the Minnesota Twin registry. The sample consisted of 458 (40%) men and 693 women with a mean age of 35.97 (SD = 7.51).

Dutch student external validation sample

For predicting external criteria with the MPQ-BF-NL, an independent sample of 1,004 psychology students was used. These students took part in a mass testing session in partial fulfillment of their first year course requirements. Of the resulting protocols, 15 exceeded the cut-offs for the VRIN/TRIN validity scales. The final sample (N = 987) consisted of 240 (24%) men and 745 women with a mean age of 20.92 (SD = 5.23).

Across samples, protocols were deemed invalid when, within sample, any of the following three criteria were met: a) a response pattern that was excessively inconsistent with respect to item content (i.e., score on VRIN > 3 SD above mean score) b) a response pattern that was excessively polarized toward either true or false irrespective of item content (i.e., score on TRIN > |3.21| SD above mean score); c) a response pattern that was both inconsistent and polarized in direction (i.e., score is 2 SD above the mean for VRIN, and |2.28| > SD from the mean for TRIN). Of note, cut-off values were selected to yield the same percentage of rejected protocols for VRIN and TRIN.

Scale development and item selection procedures

Primary trait scales

For selecting the 12 psychometric optimal items for each scale, excepting Wellbeing and Stress Reaction (see below) two criteria were used: (1) item-scale correlations that were corrected for item overlap and (2) Loadings on the relevant components from a Principal Components Analysis (PCA) with Varimax rotation on all items except Wellbeing and Stress Reaction. PCAs were carried out using the 'psych' package in R 2.12.0 (Revelle, 2010). PCA with Varimax rotation was conducted for replication purposes[1]. In order to guard against selecting items with less than optimal psychometric properties for persons with a particular gender or age, protocols from the normative sample were divided into four gender/age groups. The groups consisted of a) men ≤ 40 years (n = 183), b) women ≤ 40 years (n = 238), c) men > 40 years (n = 327), d) women > 40 years (n = 307). For each of the scales, mean criterion-parameter values across gender/age groups were computed, and the 12 items obtaining the highest mean score for each of the two criteria were selected. Discrepancies were resolved in favor of the mean corrected item-total correlation.

After these rounds of preliminary item selection, we checked the adequacy of content cluster representation of the primary trait scales. The content clusters came into being by clustering the items of the MPQ with the help of a sorting task carried out by several individuals (Tellegen & Waller, 2008). The content clusters are meant to give a description of the breadth of content coverage of the different primary trait scales, but are not used for particular psychometric characteristics. At least one item of each content cluster was to be present in the scale. When a lack of content representation dictated item substitution, the least performing item of the overrepresented cluster was substituted for the best performing item of the underrepresented content cluster.

[1] To check whether our item selection choices were robust, we also examined loadings from an Exploratory Structural Equation Model (ESEM), using Delta parameterization for categorical indicators with Target rotation (Marsh et al., 2009). For these analyses Mplus 6.1 was used (Muthén & Muthén, 2012). With ESEM it is possible to estimate the full loading matrix within a confirmatory framework. Correlations between relevant loadings on components from the PCA and factors from ESEM ranged from .81 for Social Closeness to .99 for Social Potency, with a mean of .95. Item selection choices were identical across analyses.

Wellbeing and Stress Reaction scales

As our point of departure we opted for the two minimally correlated short forms of Wellbeing (11 items) and Stress Reaction (14 items) specifically developed by Tellegen (1982). Preliminary analyses of these short scales also yielded favorable results in our samples. In addition to a relatively small inter-correlation, the items comprising these two scales showed relatively large item-scale correlations and PCA loadings. To arrive at 12 items for each scale, we augmented the short Wellbeing scale with one item and dropped two items from the short Stress Reaction scale. One item of the resulting Stress Reaction scale was substituted by another item. Our main consideration in adding and substituting items was to minimize the correlation between these two primary trait scales while ensuring content cluster representation and thus breadth of content.

Broad trait factors based on the brief scales

PCAs of the primary trait scales have yielded three or alternatively four factors (Tellegen & Waller, 2008). As mentioned above, the three-factor solution consists of PEM, NEM and CON, while the four-factor solution splits PEM into its Agentic and Communal subcomponents. For the present paper, we focused on the three-factor solution (details on the four-factor solution are available from the first author). To compute broad trait scores for the MPQ-BF-NL, sum scores of the standardized full length relevant primary trait scales (for example PEM = Wellbeing + Social Potency + Achievement + Social Closeness) were predicted from all the Dutch brief scales excepting Absorption. Unstandardized regression weights are used to develop a weighted sum score. In this manner, broad trait scale scores can be computed on the basis of the MPQ-BF-NL.

Development of validity scales

Response inconsistency indices

The VRIN (Variable Response Inconsistency) and TRIN (True Response Inconsistency) scales provide for detection of two distinct invalid response patterns. Both scales are comprised of item pairs. VRIN item pairs are content matched and are keyed either True-False (TF) or False-True (FT), or both. The total VRIN score equals the total number of keyed item-pair responses, that is, the total number of scored inconsistencies. TRIN consists of items pairs, opposite in content that are keyed TT or FF or both. The total TRIN score equals the total number TT item-pair responses minus the total number of FF item-pair responses. 'Nay saying' is indicated when highly negative TRIN scores are obtained, 'Yay

saying' when highly positive scores are obtained, and medium level scores suggest consistent responding.

For an item pair to be included in either the VRIN or TRIN scale, items do not only have to be content matched. A conceptually inconsistent combination of responses to the two items should also be empirically inconsistent. A response pattern is not considered to be empirically inconsistent when the occurrence of the pattern in the norm group is not relatively rare. More specifically, when items differ in psychometric difficulty, it is conceivable that one specific opposite pattern is not empirically indicative of inconsistent responding, while the other opposite pattern is (Tellegen, 1988; Tellegen & Ben-Porath, 2008). Inspection of the Dutch VRIN and TRIN item pair frequency tables indicated that the ordinal ranking of the difficulties of the composing item pairs was not identical to those of the U.S.. Therefore, we specifically developed Dutch VRIN and TRIN scales that were optimal to the observed item difficulty patterns. To yield a sufficient number of item pairs, we also selected items from the 276-item version of the MPQ. The final Dutch VRIN and TRIN scales consist of 21 and 16 item pairs, respectively (details on the construction are available from the authors on request).

Unlikely Virtues

The Unlikely Virtues (UV) scale is a 14-item index of social desirability. It taps impression management or other deception instead of self-deception (Paulhus & Reid, 1991), and the composing items reflect uncommon virtuousness or deny even minor shortcomings and frailties. The UV scale was translated as is: i.e. no shortening of the scale was attempted. The MPQ-BF-NL can be administered with or without this scale.

Psychometric properties of the MPQ-BF-NL: analytic strategy

To evaluate the performance of the MPQ-BF-NL, we assessed several aspects of overall instrument function. Using the Minnesota Twin Registry data set and the Dutch normative sample, we conducted the following analyses. First, we computed internal consistencies for each of the primary trait scales in both the Dutch normative sample and the Minnesota Twin Registry data set. Correlations between the MPQ-BF-NL scales and their full-scale U.S. counterparts, and the MPQ-BF scales in the Minnesota Twin Registry data set were also determined. In addition, the higher-order structure was examined by inspecting three-factor solutions of Exploratory Factor Analyses (EFAs) for categorical indicators with Varimax rotation of primary trait scales of the MPQ-BF-NL in both the Dutch and the

Minnesota Twin Registry sample, using Mplus 6.1 (Muthén & Muthén, 2012). Finally, convergent and divergent patterns were examined using external correlates of the MPQ-BF-NL in the Dutch student sample with self-reported personality and mental health indices. In our analyses, the U.S. full length MPQ served as a benchmark to evaluate the performance of the MPQ-BF-NL.

Results

Reliability of the MPQ-BF-NL primary trait scales

Table 2.1 shows the internal consistencies (Cronbach's alphas) for the MPQ-BF-NL primary trait scales in the Dutch normative sample, and the Minnesota Twin Registry data set as well as those for the full length MPQ in the Minnesota Twin Registry sample. As can be seen, alpha coefficients of the MPQ-BF-NL ranged from .70 (Traditionalism) to .87 (Positive Emotionality) in the normative sample. The internal consistencies for most primary trait scales were satisfactory to good (.75-.87), while moderate internal consistencies were observed for Traditionalism (.70), Harm Avoidance (.72) and Aggression (.73) in the normative sample. The observed internal consistencies were generally comparable to those reported for the MPQ-BF (Patrick et al., 2002): differences in Cronbach's alpha were less than or equal to |.04|, except for Traditionalism (difference of .08).

Relations between MPQ-BF-NL, full length U.S. MPQ scales and MPQ-BF

We examined correlations between MPQ-BF-NL and the full length MPQ Scales in the Minnesota Twin Registry sample (see Table 2.1). Correlations were uniformly high, ranging from .94-.96 for both the sets of primary trait scales, and broad trait scales. Correlations between MPQ-BF-NL scale scores and MPQ-BF scale scores ranged from .86 for Achievement to .99 for Wellbeing, with a mean of .94. These numbers support the notion that the MPQ-NL-BF constructs approximate the original U.S. MPQ constructs.

Table 2.1. Internal consistency estimates for the MPQ(-BF-NL) scales in the normative Dutch and the U.S. twin samples, and Pearson correlations between MPQ-BF-NL and full length MPQ scales in the U.S. twin sample

	Normative Dutch (N = 1,055)	U.S. twins (N = 1,153)		
Scale	α MPQ-BF-NL	α MPQ-BF-NL	α MPQ full length	r MPQ-BF-NL & MPQ full length
Primary trait				
Wellbeing	.80	.81	.89	.94
Social Potency	.84	.84	.90	.95
Achievement	.76	.75	.80	.95
Social Closeness	.81	.83	.86	.95
Stress Reaction	.84	.82	.89	.95
Aggression	.73	.72	.78	.95
Alienation	.82	.81	.86	.96
Control	.75	.75	.80	.92
Harm Avoidance	.72	.78	.86	.93
Traditionalism	.70	.75	.82	.92
Absorption	.80	.77	.89	.93
Broad trait[a]				
PEM	.87	.87	.91	.96
NEM	.85	.85	.91	.95
CON	.75	.76	.87	.94

Note. MPQ-BF-NL = Dutch brief form of the Multidimensional Personality Questionnaire; MPQ full length = U.S. current version; PEM = Positive Emotionality; NEM = Negative Emotionality; CON = Constraint.
[a]alphas for broad trait scales are based on scales formed by evenly weighted items.

Higher-order factor structure of the MPQ-BF-NL

In Table 2.2 correlations between the different primary trait scales of the MPQ-BF-NL in the normative Dutch and Minnesota Twin Registry samples can be found. On these correlations, EFAs with Varimax rotation were conducted. For reference, an EFA was also carried out in the Minnesota Twin Registry sample with the full (U.S.) version of the MPQ (see Table 2.3). The patterns of loadings were comparable for the MPQ-BF-NL in the normative sample, and for the full length MPQ in the Minnesota Twin Registry data set. Wellbeing, Social Potency and Achievement showed strong loadings on the factor recognizable as PEM (range .35-.86), although Social Potency loaded equally strong on CON for the MPQ-BF-NL in the Dutch normative sample (PEM: .43, CON: -.45), and had a secondary loading on CON for the MPQ in the Minnesota Twin Registry sample (PEM: .50, CON: -.32). Although Social Closeness is also hypothesized to be most strongly associated with PEM, the scale loaded more strongly on the factor recognizable as NEM for the MPQ-BF-NL in the Dutch normative sample (NEM: -.41, PEM: .17), and equally strong on PEM and NEM for the full length MPQ in the Minnesota Twin Registry sample (NEM: -.25, PEM: .27). Stress Reaction and Alienation loaded, as expected, most strongly on NEM (range .56-.82), while Aggression had its primary loading on CON for the MPQ-BF-NL in the normative sample (CON: -.40; NEM: .30)

and a strong secondary loading on CON for the MPQ in the Minnesota Twin Registry data set (NEM: .42; CON: -.36). Control, Harm Avoidance and Traditionalism had, as expected, primary loadings on CON (range: .32-.71). Secondary loadings were present for both versions and samples for Wellbeing on NEM (MPQ-BF-NL: -.51; MPQ: -.39).

Relationships with other psychological measures

We conducted multiple regression analyses to test associations of MPQ primary and higher-order traits with putatively related external variables. Two domains of external variables were examined: personality and psychopathology.

For personality, we included the Ten Item Personality Inventory (TIPI; Gosling, Rentfrow, & Swann, 2003; Eigenhuis & Kamphuis, 2011), the Dutch Personality Questionnaire, which is a Dutch adaptation of the California Personality Inventory (DPQ; Gough, 1987; Luteijn, Starren, & Dijk, 1985), the Behavioral Inhibition System/Behavioral Activation System (BIS BAS; Carver & White, 1994; Franken, Muris, & Rassin, 2005), and the Barrett's Impulsivity Scale -11 (BIS-11; Patton, Stanford, & Barratt, 1995; Lijffijt & Barratt, 2005).

Psychopathology indices were the Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961; Bouman, Luteijn, Albersnagel, & Ploeg, 1985), the trait anxiety scale from the State-Trait Anxiety Inventory (STAY-DY; Spielberger, Gorsuch, & Luschene, 1970; Ploeg, Defares, & Spielberger, 1980), the trait anger scale from the State-Trait Anger Expression Inventory (STAXI-2; Spielberger, 1980; Ploeg et al., 1980), the Penn State Worry Questionnaire (PSWQ; Meyer, 1990; Kerkhof et al., 2000), and the Ruminative Response Scale (RRS; Nolen-Hoeksema & Morrow, 1991; Raes, Hermans, & Eelen, 2003). Also, three self-report questions about behaviors related to psychopathology were answered, pertaining to: (1) number of alcoholic beverages per week, to be answered on a 7-point scale; (2) regular marihuana use (Yes/No); and (3) use of anti-depressant, barbiturate, benzodiazepine or lithium medication (Yes/No).

Table 2.4 displays Cronbach's alphas of the scales in our sample, and standardized beta weights and multiple correlations for each of the regression equations. For binary dependent variables, we used logistic regression analyses and displayed the relative odds ratios, together with the square root of Nagelkerke R^2.

Table 2.2. Correlations between the primary trait scales for the MPQ-BF-NL in the normative Dutch and the U.S. twin samples

	MPQ-BF-NL primary trait scale										
	WB	SP	AC	SC	SR	AG	AL	CO	HA	TR	AB
Wellbeing		**.36**	**.28**	**.27**	-.27	-.09	-.18	.06	-.11	-.01	.30
Social Potency	**.29**		**.35**	**.20**	-.17	.18	-.11	.00	-.26	-.18	.16
Achievement	**.24**	**.28**		**-.05**	.03	.05	.06	.14	-.13	.07	.19
Social Closeness	**.31**	**.22**	**.12**		-.13	-.13	-.18	-.01	.14	.03	-.01
Stress Reaction	-.41	-.14	.04	-.22		**.25**	**.39**	-.13	.21	.12	.28
Aggression	-.12	.22	.00	-.16	**.20**		**.30**	-.21	-.19	-.01	.10
Alienation	-.25	-.05	.13	-.27	**.43**	**.24**		-.17	.06	.16	.17
Control	.01	-.12	.24	-.08	-.07	-.18	-.11		**.16**	**.08**	-.01
Harm Avoidance	-.10	-.23	-.14	.03	.06	-.16	-.06	**.18**		**.21**	-.13
Traditionalism	.03	-.15	.08	.01	.06	-.12	.10	**.15**	**.24**		-.02
Absorption	.16	.06	.08	-.11	.29	-.02	.21	-.08	-.15	.03	

Note. Below the diagonal scale correlations in the normative Dutch sample are presented; above the diagonal one finds scale correlations in the U.S. twin sample. Correlations ≥ .20 are underlined; correlations between theoretically related scales are presented in bold. WB = Wellbeing; SP = Social Potency; AC = Achievement.; SC = Social Closeness; SR = Stress Reaction; AG = Aggression; AL = Alienation; CO = Control; HA = Harm Avoidance; TR = Traditionalism; AB = Absorption.

Table 2.3. Factor loadings of primary trait scales on the first three factors from Exploratory Factor Analysis with Varimax rotation of the MPQ-BF-NL in the normative Dutch and the U.S. twin samples, and Exploratory Factor Analysis of the full length MPQ in the U.S. twin sample

	Normative Dutch (*N* =1,055) MPQ-BF-NL			U.S. twins (*N* = 1,153) MPQ-BF-NL			MPQ full length		
	Eigenvalue			Eigenvalue			Eigenvalue		
	1.45	2.20	1.81	2.16	1.43	1.89	2.24	1.14	2.11
Scale	PEM	NEM	CON	PEM	NEM	CON	PEM	NEM	CON
Wellbeing	**.35**	-.51	-.11	**.70**	-.39	-.08	**.73**	-.39	.05
Social Potency	**.43**	-.19	-.45	**.48**	-.06	-.34	**.50**	-.05	-.32
Achievement	**.86**	.06	.17	**.43**	.04	-.06	**.37**	.03	-.08
Social Closeness	**.17**	-.41	-.06	**.19**	-.33	.10	**.27**	-.25	.28
Stress Reaction	-.01	**.72**	.02	.05	**.65**	.49	-.14	**.82**	.23
Aggression	.06	**.30**	-.40	.12	**.49**	-.16	.10	**.42**	-.36
Alienation	.12	**.64**	-.07	.03	**.51**	.19	-.04	**.56**	-.06
Control	.17	-.08	**.49**	.01	-.27	**.13**	-.08	-.22	**.32**
Harm Avoidance	-.23	-.01	**.40**	-.18	-.17	**.66**	-.17	.01	**.71**
Traditionalism	.04	.07	**.34**	.00	.00	**.31**	-.02	.07	**.36**
Absorption	.14	.24	-.12	.54	.26	.10	.49	.39	-.05

Note. Loadings ≥ .30 are underlined; expected primary loadings are presented in bold (see Tellegen & Waller, 2008). MPQ-BF-NL = Dutch brief form of the Multidimensional Personality Questionnaire; MPQ full length = U.S. current version; PEM = Positive Emotionality; NEM = Negative Emotionality; CON = Constraint.

Table 2.4. Cronbach's alphas of external variables, beta weights and multiple correlations for prediction of the scales using the primary and higher-order trait scales of the MPQ-BF-NL within the Dutch student sample (*N* = 987)

		MPQ-BF-NL primary trait scale												MPQ-BF-NL higher-order			
Scale	*α*	WB	SP	AC	SC	SR	AG	AL	CO	HA	TR	AB	*R*	PEM	NEM	CON	*R*
Personality																	
TIPI (*n* = 475)																	
Extraversion	.77	.15	.34	.00	.36	.02	-.05	.03	-.12	.05	.03	-.01	.67	.59	.01	-.10	.59
Openness	.37	.08	.20	.05	.04	-.02	-.13	-.06	-.01	-.16	-.07	.16	.42	.28	-.06	-.18	.36
Conscientiousness	.49	.10	.07	.11	-.06	-.07	-.05	.02	.47	.03	.08	-.07	.57	.11	-.07	.45	.47
Neuroticism	.56	-.14	-.05	.06	.04	.53	-.01	-.05	-.10	.06	-.03	-.03	.59	.09	.40	.08	.45
Agreeableness	.12	.11	-.11	-.14	.13	-.18	-.22	.04	.04	.02	.03	.15	.41	.01	-.29	.03	.29
DPQ (*n* = 967)																	
Inadequacy	.82	-.24	-.03	-.05	-.04	.49	.05	.15	-.10	-.02	.06	.09	.73	-.19	.61	-.01	.69
Social Inadequacy	.84	-.09	-.37	-.04	-.33	.14	.06	.03	.06	-.03	.01	-.05	.67	-.54	.14	.11	.61
Rigidity	.74	.00	.06	.08	-.01	.17	.06	.09	.33	.07	.28	.28	.59	.12	.28	.47	.55
Discontentedness	.75	-.08	.06	.01	-.16	.17	.22	.25	-.04	-.02	.13	.01	.60	-.04	.53	-.01	.55
Self-satisfiedness	.65	.04	.04	-.04	-.35	-.13	.19	.12	-.03	-.09	.07	-.06	.48	-.13	.14	-.12	.24
Dominance	.69	.10	.61	.02	.06	-.03	.11	.02	-.03	-.04	.06	.03	.72	.60	.18	-.16	.59
Self-Esteem	.68	.37	.09	.29	.03	-.20	-.06	-.01	.04	-.01	.00	-.09	.64	.50	-.20	.07	.59
BIS BAS (*n* = 476)																	
BAS Drive	.74	.14	.18	.36	.07	-.05	.03	.09	.12	.04	-.02	.02	.54	.49	.13	.15	.49
BAS Fun Seeking	.50	.20	.05	.08	.18	-.01	.06	.07	-.04	-.15	-.05	.11	.42	.35	.09	-.21	.39
BAS Reward Responsiveness	.67	.09	.05	.10	.18	.11	.16	-.08	.05	-.01	.03	.02	.32	.29	.14	.04	.27
BIS	.84	-.06	-.13	.09	.09	.51	.01	-.16	.17	.17	-.04	.00	.66	-.06	.27	.35	.48
Barratt's Impulsivity-11 (*n* = 458)																	
Attentional Impulsiveness	.66	-.05	-.00	-.25	-.03	.27	.09	-.06	-.22	-.01	-.02	.10	.47	-.16	.21	-.24	.38
Motor Impulsiveness	.65	.13	.02	-.15	.11	.04	.00	.01	-.48	-.02	-.06	.03	.61	.11	.00	-.46	.47
Nonplanning Impulsiveness	.90	-.05	-.11	-.29	.10	.17	-.01	.09	-.43	-.04	-.05	-.04	.65	-.22	.09	-.41	.49

Table 2.4. (continued)

		MPQ-BF-NL primary trait scale												MPQ-BF-NL higher-order			
Scale	α	WB	SP	AC	SC	SR	AG	AL	CO	HA	TR	AB	*R*	PEM	NEM	CON	*R*
Psychopathology																	
BDI (*n* = 972)																	
Depression	.82	-.31	.02	.06	-.05	.32	.01	.17	-.04	-.01	.02	.02	.62	-.17	.48	.04	.56
STAI-DY (*n* = 948)																	
Trait Anxiety	.90	-.28	-.01	.07	-.01	.55	-.02	.11	-.08	-.01	-.04	.03	.74	-.16	.59	.03	.66
STAXI-2 (*n* = 948)																	
Trait Anger	.83	-.05	.11	.05	-.03	.29	.37	.04	-.13	.07	-.01	-.02	.57	.08	.53	-.10	.52
PSWQ (*n* = 461)																	
Worry	.93	-.12	-.08	.13	.08	.66	.00	-.03	.15	.05	-.03	-.01	.78	-.04	.51	.29	.63
RRS (*n* = 475)																	
Rumination	.91	-.25	.04	.00	-.06	.42	.00	.01	-.02	-.02	-.04	.11	.57	-.16	.42	.01	.50
Alcohol (*n* = 967)	-	.13	.01	-.11	.15	.02	.08	.01	-.19	-.11	-.03	-.03	.41	.13	.00	-.29	.32
Marihuana[a] (*n* = 969)	-	.94	1.0	.90	.96	1.0	1.1	1.0	.89	.91	.99	1.2	.37	.98	1.0	.93	.32
Psychoactive medication[a] (*n* = 969)	-	.91	.98	1.1	.82	1.6	.68	1.0	.69	1.3	1.2	1.0	.58	.98	1.1	1.0	.28

Note. Standardized betas are underlined at α = .05, Bonferroni corrected: $p < .004$ for primary trait scales and $p < .02$ for higher-order traits. MPQ-BF-NL = Dutch brief form of the Multidimensional Personality Questionnaire; WB = Wellbeing; SP = Social Potency; AC = Achievement; SC = Social Closeness; SR = Stress Reaction; AG = Aggression; AL = Alienation; CO = Control; HA = Harm Avoidance; TR = Traditionalism; AB = Absorption; PEM = Positive Emotionality; NEM = Negative Emotionality; CON = Constraint; TIPI = Ten Item Personality Inventory; DPQ = Dutch Personality Questionnaire; BIS BAS = Behavioral Inhibition System Behavioral Activation System; Barrett's Impulsivity – 11 = Barrett Impulsivity Scale – 11; BDI = Beck Depression Inventory; STAY-DY = State-Trait Anxiety Inventory; STAXI-2 = State-Trait Anger Expression Inventory; PSWQ = Penn State Worry Questionnaire; RRS = Ruminative Response Scale.
[a]Displayed values are relative odds ratios from logistic regression analyses. *R* is the square root of Nagelkerke R^2.

Personality

First, we inspected the pattern of associations for findings that were at odds with the theory underlying the MPQ or with earlier findings in the United States. None of the findings, reported in Table 2.4, were at odds with what we expected. MPQ scales were strongly associated with scales from other personality inventories, and patterns of convergence and divergence were consistent. For example, DPQ-Inadequacy was most strongly predicted by MPQ NEM, which seems to be due to a strong association between DPQ Inadequacy and Stress Reaction. DPQ Dominance was best predicted by MPQ PEM and Social Potency, and DPQ Social Inadequacy by a combination of Social Potency (negative) and Social Closeness (negative). The same pattern was evident on the Big Five TIPI: Primary associations were observed between TIPI Neuroticism and MPQ NEM which was due to a strong association of Neuroticism and Stress Reaction. TIPI Conscientiousness was best predicted by MPQ Control and TIPI Extraversion by the combination of the MPQ Social Potency and Social Closeness scales.

Also as expected, BAS drive was most strongly related to the agentic aspect of PEM (Social Potency and Achievement), while BAS fun seeking was most strongly related to the Communal aspect of PEM (Social Closeness and Wellbeing). BIS was related to NEM and CON about equally, but from the primary traits, Stress Reaction was clearly the dominant predictor. For all three of Barrett's impulsivity facets, MPQ Control and Achievement predicted scores best.

Psychopathology

As the MPQ is intimately connected with theories in psychopathology, it was important to test whether associations with disorders were as expected. General support was found for a strong association of traits with psychological problems, as MPQ scales predicted a great deal of variance in measures of psychopathology. More specifically, one would expect the following pattern of associations: (1) NEM scales should be associated with a very broad range of psychopathology; most strongly with anxiety- and aggression-related problems; (2) PEM scales should be inversely associated with feelings of apathy and anhedonia; (3) CON scales should be the most significant (inverse) predictors of externalizing psychopathology. These patterns have been found in research in the United States (Arseneault, Moffitt, Caspi, Taylor, & Silva, 2000; DiLalla, Gottesman, Carey, & Vogler, 1993; Krueger, Caspi, & Moffitt, 2000; Miller, Greif, & Smith, 2003; Watson, Clark, & Carey, 1988) and were indeed replicated in our samples.

NEM indicated psychopathology in general. Scales specifically associated with Stress Reaction were the State-Trait Anxiety Inventory, the Ruminative Response Scale, and the Penn State Worry Questionnaire. The State-Trait Anger Expression Inventory was most related to Aggression and secondarily to Stress Reaction. Furthermore, our most important marker of anhedonia, the Beck Depression Inventory, was indeed most strongly associated with PEM Wellbeing of all psychopathology scales.

The substance use markers, marking problems in acting out and impulse control, were specifically related to CON-subscales. Specifically, low MPQ Control predicted consistently more substance use. Interestingly, the communal aspect of PEM was positively associated with alcohol use, whereas Achievement (a PEM Agency subscale) was negatively associated. Although low Control was also indicative for marijuana use, no effect for the PEM-scales was found.

Discussion

The main objective of the present study was to develop a Dutch brief form of the MPQ. Central to this aim was to derive the same set of constructs as those that are measured by the U.S. version of the MPQ. Primary tests of similarity of constructs are that a) they hang together similarly, and that b) they relate in similar fashion to external constructs. By and large, our findings are commensurate with those obtained for the U.S. MPQ-BF (Patrick et al., 2002; Tellegen & Waller, 2008).

To test for undesirable construct drift as a result of abbreviation, we used the full length U.S. version as a point of reference, and examined the performance of our item selection in terms of internal consistency and higher-order structure. First, few differences emerged in terms of internal consistencies between the U.S. MPQ-BF and the MPQ-BF-NL. Lower internal consistencies were observed for Traditionalism, Harm Avoidance and Aggression, raising the possibility of cross-cultural differences in structure. Other explanations center on the idea that item difficulties differ between the U.S. and the Netherlands. Future structural investigations, along with external validation studies such as those reported here, will elucidate these issues. Second, the correlations between the primary trait scales of the MPQ-BF-NL and their full length U.S. version were high, and commensurate with those obtained for the MPQ-BF (both mean r = .94).

Overall, the EFA of the primary trait scales approximates the higher-order structure in the U.S. rather well. As observed in the U.S., a number of cross-loadings emerged for the primary trait scales, e.g. Aggression (negative on CON), Social Potency (negative on CON), Wellbeing (negative on NEM) and Social Closeness (negative on NEM). These cross-loadings are interpretable departures from strict simple structure, and are consistent with earlier findings (Tellegen & Waller, 2008).

The pattern of correlations with theoretically relevant external correlates was consistent with expectations. For example, in the trait domain, convergent patterns were observed for BIS and BAS scales with NEM and PEM respectively, while impulsivity measures converged with CON. With respect to psychopathology, of the primary trait scales, the core affective traits of the MPQ (i.e. Stress Reaction and Wellbeing) consistently showed the strongest associations with indices associated with the emotional disorders/internalizing psychopathology (i.e. anxiety, depression, use of psychotropic drugs). In contrast, substance use, typically conceived of as predominantly externalizing psychopathology was most strongly related to primary trait scales belonging to Constraint, particularly the Control scale. Some of the more specific findings point to the heuristic potential of the set of MPQ traits. For example, both alcohol use and marijuana use were both predicted by higher-order affective impulsivity (a combination of high PEM and low CON). Predictive patterns diverged however in that alcohol use was also predicted by the communal aspects of PEM and lower Achievement, while PEM scales were not associated with marijuana use.

Our strategy was to derive an item set that would yield optimal construct replication. As a consequence, the item selection was not identical, though highly similar, to the U.S. MPQ-BF. This strategy has limitations; specifically it complicates studies on measurement invariance between versions. However, both the MPQ-BF and the MPQ-BF-NL are nested in the full length version, which makes direct comparison possible when data on the full length version is available.

The results presented here should be considered as preliminary evidence for the comparability of scores derived from the MPQ-BF in the U.S. and the MPQ-BF-NL in the Netherlands. However, our analyses cannot provide conclusive evidence about measurement invariance of U.S. and Dutch brief versions of the MPQ, or about the structural equivalence of responses of individuals with different genders etc. Currently, we are investigating measurement invariance across genders and language versions, and are

empirically mapping DSM-IV/V personality disorders to the PEM, NEM, CON structure of the MPQ. Moreover, associations between the psychometric properties of the MPQ-BF-NL and brain anatomy are examined.

Based on the observed consistency in higher-order structure and pattern of meaningful extra test correlates, we are confident that the MPQ-BF-NL scales were successful in capturing their target constructs. The MPQ-BF-NL provides the Dutch-Flemish language area with a personality inventory that has promising characteristics for use in both research and clinical practice. For the international research community the instrument offers new opportunities for fundamental and cross-cultural studies on personality.