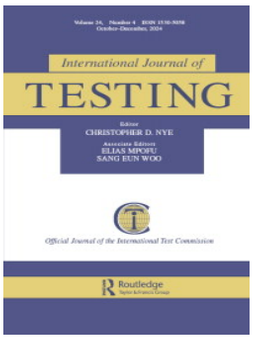

International Journal of Testing

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/hijt20

Examining the construct validity of the MIDUS version of the Multidimensional Personality Questionnaire (MPQ)

Khalid ALMamari

To cite this article: Khalid ALMamari (2024) Examining the construct validity of the MIDUS version of the Multidimensional Personality Questionnaire (MPQ), International Journal of Testing, 24:4, 321-355, DOI: 10.1080/15305058.2024.2364176

To link to this article: https://doi.org/10.1080/15305058.2024.2364176

INTERNATIONAL JOURNAL OF TESTING
2024, VOL. 24, NO. 4, 321–355
https://doi.org/10.1080/15305058.2024.2364176

Routledge
Taylor & Francis Group

Examining the construct validity of the MIDUS version of the Multidimensional Personality Questionnaire (MPQ)

Khalid ALMamari

Military Technological College, Muscat, Oman

ABSTRACT

The Multidimensional Personality Questionnaire (MPQ) measures a wide range of personality traits associated with affect and temperament. However, the lengthy administration time may have hindered its widespread use in personality research. The National Survey of Midlife Development in the United States (MIDUS) has adapted a short version of the MPQ that consists of only 35 items. However, the psychometric properties of this abbreviated version have not been thoroughly examined beyond what is presented in the survey's documentation. This study aimed to investigate the construct validity of the MPQ-35 at both the item and scale levels using exploratory factor analysis (EFA), confirmatory factor analysis (CFA), and exploratory structural equation modeling (ESEM), measurement invariance by gender and age, and external correlates with several personality and psychological constructs. The results show that the MPQ-35 exhibits promising properties, performing well and accurately recapturing the original MPQ structure, with ten lower-order traits and three higher-order broad factors.

KEYWORDS

Construct validity; psychometric properties; personality assessment; MPQ; ESEM

Examining the internal and external construct validity of the MIDUS version of the Multidimensional Personality Questionnaire

The Multidimensional Personality Questionnaire (MPQ; Tellegen, 1982) is a self-report inventory designed to assess a range of personality traits

CONTACT Khalid ALMamari khalid.s.almamari@gmail.com The Military Technological College, Muscat, Oman.

across multiple scales. The MPQ is aligned with Tellegen's model of personality (Church, 1994), which emphasizes both positive and negative affect and provides a comprehensive model for understanding individual differences (Tellegen & Waller, 2008). The MPQ consists of 11 primary trait scales, organized into three higher-order factors: positive emotionality (PEM), negative emotionality (NEM), and constraint (CON) (Marquardt et at., 2021; Patrick et al., 2002, 2013; Tellegen, 1982). The PEM factor is comprised of well-being, social potency, achievement, and social closeness scales, whereas the NEM factor encompasses stress reaction, alienation, and aggression scales. The CON factor includes harm avoidance, control, and traditionalism scales. The eleventh trait, absorption, reflects the tendency for imaginative, artistic, and self-absorbing experiences and is seen as a distinct personality construct (Tellegen, 1982). PEM and NEM are considered complementary dimensions of emotional temperament, with PEM reflecting tendencies toward positive emotions and NEM toward negative emotions. In contrast, CON represents a boundary on behavioral expression (Miller et al., 2003). While the three higher-order factor solution is commonly proposed to underlie the structure of the MPQ, there is also evidence for a four-factor solution, where the PEM factor is subdivided into agentic and communal PEM (Church & Burke, 1994; Javdani et al., 2014; Tellegen & Waller, 2008). Agentic PEM is characterized by high levels of social potency and achievement, while communal PEM is characterized by high levels of well-being and social closeness. Table 1 provides a brief description of the primary traits and the "big" four factors.

The 11 scales of the MPQ were carefully developed through a rigorous process involving item development, data collection, and factor analysis of extensive item pools (Tellegen & Waller, 2008). The application of exploratory test construction[1] techniques has further refined these scales, making them more "'psychologically coherent" and "robust" compared to other methods (Waller et al., 2016). Additionally, recent uses of the recaptured scale technique[2] have demonstrated that the MPQ scales are highly

[1]The "exploratory test construction" technique involves a detailed process of creating psychological scales, such as those in the MPQ. This method is characterized by its thorough and iterative approach, including multiple rounds of item writing, data collection, and factor-analytic scale refinement from over inclusive item pools. It aims to produce scales that are psychologically coherent and robust, enhancing the scale's validity and reliability. This technique is distinguished by its emphasis on exploratory factor analysis as a core component of scale development + (see Tellegen & Waller, 2008).

[2]The "Recaptured Scale Technique" (RST) is a method for testing the structural robustness of personality scales. It involves embedding scale items into a larger item pool and using factor analysis to determine if the original scale is distinguishable within this mix. This technique assesses a scale's coherence and stability across different settings, demonstrating its structural integrity and distinctiveness + (see Bouchard & Waller, 2017; Waller et al., 2016).

Table 1. Descriptions of MPQ Primary Trait Scales, MPQ higher-order factors, and MPQ-35 indicators with number of items.

	MPQ Scale	Description of a high scorer	MPQ-35's indicators (no of items)
Primary traits	Well-being	Feelings of contentment, satisfaction, and positive affect	Livening, adventurous, optimistic (3)
	Social potency	Assertiveness, dominance, and desire for leadership roles	Reluctant, persuasive, decisive (4)
	Achievement	Drive for success, ambition, and high performance	Persistent, ambitious, diligent, perfectionist (4)
	Social closeness	Strong need for social interaction and close relationships with others	Sociable, companion-seeking, warm-hearted, extroverted (4)
	Stress reaction	Tendency to experience psychological distress in response to stressors	Mood-sensitive, tension-prone, easily-irritated (3)
	Aggression	Tendency to act out in a hostile or aggressive manner	Physically-aggressive, hurtful-tongued, vengeful, violent-inclined (4)
	Alienation	Feeling disconnected from others and society	Exploited, hindered, belittled (3)
	Control	Desire for order and control over one's environment	Deliberative, reflective, cautious (3)
	Traditionalism	Strong attachment to tradition, customs, and established norms	Anti-censorship, morally-strict, tradition-valuing (3)
	Harm avoidance	Fear and avoidance of negative outcomes and risky situations	Earthquake-curious, tightrope-adventurous, rapids-averse, loose-lions-averse (4)
	Absorption	Tendency to become deeply engrossed in experiences, ideas, and fantasies	N/A
Higher-order factors	Agentic PEM	Positive emotions associated with success, achievement, and dominance	Well-being, social potency, achievement, social closeness
	Communal PEM	Positive emotions associated with close relationships and social connectedness	
	NEM	Tendency to experience negative emotions such as anxiety, depression, and anger	Stress reaction, alienation, aggression
	CON	Tendency to inhibit impulses and adhere to social norms and rules	Harm avoidance, control, traditionalism

Note. PEM: positive emotionality; NEM: negative emotionality; CON: constraint
Adapted from Tellegen and Waller (2008) and Ryff et al. (2021)

resilient and can be successfully recaptured when their items are included in a large pool of personality and psychopathology items (Bouchard & Waller, 2017; Waller et al., 2016).

Tellegen's (1982) original 300-item MPQ laid the foundation for multiple MPQ forms, including the currently maintained 276-item form (Tellegen, 2011), a simplified 157-item form (Patrick et al., 2013), a brief 155-item form (Patrick et al., 2002), and an abbreviated 55-item research form (Marquardt et al., 2021). The MPQ's adaptability and usefulness are exemplified by its successful adaptation and translation into different languages to expand its use across cultures. Some of the translations include Hebrew (Ben-Porath et al., 1995), German (Johnson et al., 2008), and Dutch (Eigenhuis et al., 2013). The MPQ has also been administered on a wide range of samples, including adolescents (Patrick et al., 2013), university students (Harkness et al., 1995), midlife individuals (Ryff et al., 2019, 2021), and clinical populations, such as those with substance-use disorders (McGue et al., 1997), eating disorders (Fulkerson et al., 1999), and psychiatric disorders (DiLalla et al., 1993). Across diverse samples and cultures, the MPQ has consistently demonstrated robust psychometric properties with unique characteristics that distinguish it from other personality inventories (Patrick et al., 2013).

In the National Survey of Midlife Development in the United States (MIDUS-II 2003–2004 and MIDUS-III 2013–2014; Ryff et al., 2019, 2021), a 35-item MPQ form was administered, making it the shortest version of all MPQs. This short form, henceforth referred to as MPQ-35, consists of subsets of items from the original MPQ to briefly measure 10 of the 11 primary trait scales, excluding the absorption scale. The content covered by each scale is shown in Table 1. Abbreviated MPQ forms have several advantages. As noted by Patrick et al. (2002), the availability of a short form increases the use of the MPQ in research, especially given the relative length of the existing MPQ forms, which can take a long time to administer. It also increases the possibility of including MPQ scales in large-sample research surveys, such as longitudinal projects and epidemiological investigations, where multiple measures are administered. Additionally, a short version of the MPQ would assist research that aims to relate MPQ primary traits with other personality models and constructs, such as the Big Five factors.

This practicality and ease of use have contributed to the rise in popularity of short personality inventories in recent years (Spörrle & Bekk, 2014; Yarkoni, 2010). For instance, the Big Five traits model boasts several brief versions, including the Single-Item Measures of Personality (SIMP; Woods & Hampson, 2005), the Mini-International Personality Item Pool (Mini-IPIP6; Milojev et al., 2013), the Big Five Inventory 10 and 2 (BFI-10 and BFI-2; Rammstedt et al., 2018), and the Ten-Item Personality Inventory (TIPI;

Gosling et al., 2003). Similarly, other personality models have introduced brief versions of their comprehensive inventories, such as the Brief HEXACO Inventory (BHI; De Vries, 2013), the 24-Item Questionnaire Big Six scales (24-QB6; Thalmayer et al., 2011), the Dirty Dozen (Jonason & Webster, 2010), and the Short Dark Triad (Jones & Paulhus, 2014).

While the MPQ-35 has been successfully used in two MIDUS waves (MIDUS-II and MIDUS-III), and is likely to be used in future MIDUS surveys, its internal and external construct validity remains unaddressed. It is crucial to investigate the psychometric properties of the ten scales that make up the MPQ-35, as it would help establish the reliability and validity of the instrument. The increasing use of the open data provided by MIDUS and the numerous research studies conducted using this data further emphasize the significance of validating the MPQ-35. This validation is essential to enhance the confidence with which research findings can be interpreted.

Furthermore, the limitations commonly associated with short forms call for a comprehensive evaluation of the MPQ-35's psychometric properties beyond the information provided in MIDUS documentation. A primary concern with short forms is their potential impact on reliability (Kleka & Soroko, 2018a; Smith et al., 2000). By design, shorter scales include fewer items, which may lead to decreased reliability compared to their longer counterparts. This reduction in reliability could affect the precision of measurements, potentially limiting the scale's ability to capture subtle nuances in personality traits. Another important consideration is the item selection process (Eisenbarth et al., 2015; Yarkoni, 2010). If not conducted with care, it may result in the omission of critical facets from the inventory. Therefore, it is crucial to determine whether the selected items adequately represent the intended traits. Lastly, a concern with short forms is the potential lack of depth in assessment (Kleka & Soroko, 2018b; Stanton et al., 2002). Although shorter scales are time-efficient, they may compromise the comprehensive assessment depth provided by longer versions. Longer scales typically feature a wider array of items, allowing for a more detailed exploration of the measured traits.

Therefore, while the MPQ-35 presents an accessible and feasible option for measuring traits within Tellegen's model of personality, it also introduces considerations that necessitate a comprehensive validation effort. This effort should cover crucial psychometric aspects such as reliability, factor structure, and external correlates of the MPQ-35, to determine the extent to which we can rely on its assessment. To achieve this, I conducted a thorough analysis of the MPQ-35 data, examining both item and scale levels using various models, including exploratory factor analysis (EFA), confirmatory factor analysis (CFA), and exploratory structural equation modeling (ESEM). Additionally, I tested the measurement invariance of the MPQ-35 across gender and age

groups to ascertain whether its structure remains equivalent across these demographics. Furthermore, an extensive examination of the MPQ-35's external construct validity was carried out by utilizing a range of personality and psychological constructs as criteria.

Previous MPQ studies have employed both EFA and CFA to validate the factor structure of the MPQ forms (Church, 1994; Donnellan et al., 2012; Eigenhuis et al., 2013; Marquardt et al., 2021; Patrick et al., 2002, 2013; Tellegen & Waller, 2008). However, the use of ESEM in these validations has been limited, despite recent research suggesting its potential suitability and superiority over traditional CFA models in personality models (Sellbom & Tellegen, 2019), including those of the MPQ (Eigenhuis et al., 2017; Gomez et al., 2020). Therefore, I applied ESEM in this study to offer additional insights into the factor structure of the MPQ and to more effectively address potential cross-loadings among items that measure different facets within the model. Prior tests of measurement invariance for the MPQ have shown evidence of equivalency across different groups. Eigenhuis et al. (2017) demonstrated that the measurement invariance test of the MPQ-BF-NL across general and clinical samples revealed strict invariance, with bias in 10% of items, particularly on the achievement scale. This finding suggests a unified structure between normal and pathological personalities, highlighting the MPQ's clinical applicability. In contrast, Eigenhuis et al. (2015) found that the measurement invariance test and differential item functioning (DIF) analysis of the MPQ-BF-NL indicated partial strict invariance between Dutch and U.S. samples, with as many as 40% of items exhibiting DIF. This highlighted cultural differences in key personality traits. Given the brevity of the MPQ-35 and the potential drawbacks associated with such brevity, further examination through a measurement invariance test is warranted. Considering the concern for measurement invariance by gender and age groups in personality inventories, this study tested the MPQ-35 across these variables to ensure its applicability across diverse population segments and to evaluate potential interpretation bias within them.

To further establish the validity of the MPQ-35, it is essential to examine its external construct validity (e.g., Cronbach & Meehl, 1955; Kline, 2013). This step has consistently been critical in validating MPQ forms. The process involves assessing the relationships between the MPQ scales and other relevant constructs, measuring the strength of these associations, and demonstrating their ability to predict intended outcomes. To achieve this, I selected 19 personality and psychological constructs based on their relevance to psychological functioning and well-being, as well as their established application in previous MPQ research. These constructs have also been extensively utilized in personality research for assessing the construct validity of various personality models. Examples

of these targeted constructs include the Big Five traits, psychological well-being, positive and negative affect, life orientation, self-esteem, life satisfaction, and sense of control. Selecting these constructs improves our ability to identify meaningful relationships with the MPQ-35 scales, thus providing evidence for both convergent and discriminant validity.

Present study

This study aimed to evaluate the psychometric properties of the MIDUS 35-item MPQ version and contribute to the expanding literature on the psychometric validation of the MPQ forms. The primary objective was to gather evidence on the internal and external construct validity of the MPQ-35. To assess internal construct validity, I conducted a comprehensive analysis utilizing EFA, CFA, and ESEM, coupled with measurement invariance tests across gender and age groups. These analyses were conducted to ensure the factor structure of the MPQ-35 and its applicability for use across diverse groups within the midlife population. External construct validity was assessed through an in-depth examination of the relationships between the MPQ-35 scales and 19 other personality and psychological constructs. This assessment aimed to establish both convergent and discriminant validity.

Methods

Sample

Data were collected from 3,294 participants in the MIDUS-III conducted in 2013–2014 (Ryff et al., 2019). The sample was 54.9% female and 45.1% male, with a mean age of 63.64 ($SD = 11.35$), ranging from 39 to 93 years old. Among the participants, 24% were ages 55 years and below, 34% were ages 56–66 years, and 42% were ages 67 years and above. Of the participants, 88.7% self-identified as "White," 3.7% as "Black and/or African American," and the remaining participants chose other categories. Full details about the sample, procedure, and variables can be found on the ICPSR website (https://www.icpsr.umich.edu).

Measures

The Multidimensional Personality Questionnaire (MPQ-35)

The MPQ-35 (Ryff et al., 2019; Tellegen, 1982) is the focus of this study and measures 10 personality constructs using a 4-point Likert-scale (1 = *true of you*; 2 = *somewhat true*; 3 = *somewhat false*; 4 = *false*). The internal consistency of the 10 scales ranges from .54 to .73, with most exceeding .65 (refer to Table 2).

The Big Five personality inventory

The Big Five personality inventory assesses the big five domains of personality (neuroticism, extraversion, openness to experience, agreeableness, conscientiousness) and the agency domain, using 31 self-descriptive adjectives selected from various personality scales (Goldberg, 1992; John, 1990; Trapnell & Wiggins, 1990). Participants rate items on a 4-point Likert scale ranging from *a lot* (1) to *not at all* (4). Cronbach's alpha for the six scales ranges from .56 to .81.

Psychological Well-Being Scale (PWB)

The PWB (Ryff & Keyes, 1995) measures six components of positive functioning, including autonomy, environmental mastery, personal growth, positive relations, purpose in life, and self-acceptance. The MIDUS-III version consists of 42 7-point Likert-type items. The Cronbach's alpha for the six scales ranges from .69 to .84.

The Positive and Negative Affect Schedule (PANAS)

The PANAS (Mroczek & Kolarz, 1998) measures positive and negative affect through two separate scales, each with six items. Participants rate items on a 5-point Likert-type scale ranging from *all the time* (1) to none of the time (5). The negative scale has a Cronbach's alpha of .80, and the positive scale has a Cronbach's alpha of .86.

The Life Orientation Test (LOT)

The LOT (Scheier et al., 1994) assesses dispositional optimism and pessimism through two primary scales, each comprising three items rated on a 5-point Likert scale. The Cronbach's alpha for the Optimism and Pessimism scales are .69 and .80, respectively.

Rosenberg Self-Esteem Scale (RSES)

The RSES (Rosenberg, 1965) measures self-esteem using a 7-item version with responses on a 7-point Likert-type scale ranging from *strongly agree* (1) to *strongly disagree* (7). The Cronbach's alpha for the scale items is .76.

Life Satisfaction Scale

The Life Satisfaction Scale (Prenda & Lachman, 2001) is a 6-item scale that assesses life satisfaction on a 10-point scale from *the worst possible* (0) to *the best possible* (10). The Cronbach's alpha for the scale items is .76.

The Sense of Control scale

The Sense of Control scale (Lachman & Weaver, 1998) includes two subscales: personal mastery (4 items) and perceived constraints (8 items). Participants rate items on a 7-point scale ranging from *strongly agree* (1) to *strongly disagree* (7), with the score being the average of the two subscales. The combined score has a Cronbach's alpha of .87.

Analytic plan

The analyses conducted in this study were performed using Mplus 8.10 (Muthén & Muthén, 2023) and Jamovi 2.3.28 (The Jamovi Project, 2021). The descriptive statistics of the MPQ-35 primary trait scales were first analyzed by gender and age groups. I used independent sample *t*-tests and one-way ANOVA to compare means across gender and age groups, respectively, and calculated effect sizes using Cohen's *d* (Cohen, 1988) and Partial eta squared (η^2) (Olejnik & Algina, 2003). I also evaluated the internal consistency of the MPQ-35 scales through Cronbach's alpha (Cronbach, 1951) and McDonald's omega (McDonald, 1999).

Factor analysis at the item level

To explore and test the factor structure of the MPQ-35, I took a multistep approach. I used EFA to identify potential factor solutions from the data, CFA to test the fit of a factor structure to the data, and ESEM to explore any alternative solutions that were not found by EFA or CFA alone. To enhance the evaluation of the factor structure, I divided the total sample into two random halves: one for EFA and the other for CFA. For the ESEM and subsequent analyses, the entire sample was utilized.

EFA and CFA

I used maximum likelihood (ML) estimation method and geomin rotation in the EFA of MPQ-35 items. Geomin rotation was chosen for its suitability in analyzing intercorrelated items within scales and measures that assess multidimensional facets, as well as its capacity to account for correlated factors (Asparouhov & Muthén, 2009; Browne, 2001). This choice is particularly pertinent in the context of personality inventories, such as MPQ-35. To determine the optimal number of factors to retain, I employed a range of techniques, including Kaiser's (1960) mineigen criterion, Cattell's (1966) scree test, Horn's (1965) parallel analysis (HPA), and Velicer's (1976) minimum average partial (MAP) method. The factor solutions recommended by these criteria, along with the theorized 10-factor model, were then forced and evaluated for interpretability, theoretical consistency, and model fit. For a factor to be deemed significant, it needed to exhibit at least two salient

loadings, with factor pattern coefficients equal to or exceeding .30. The most plausible solution identified by EFA was then tested using a correlated CFA model with ML estimation method.

ESEM & measurement invariance

Subsequently, I applied ESEM to the entire sample, also utilizing ML estimation and geomin rotation, as used in the EFA. ESEM offers greater flexibility in detecting and addressing cross-loadings and residual correlations compared to CFA. Ultimately, the best-fitting model for the MPQ-35 data, whether CFA or ESEM, was tested for measurement invariance across gender and age groups to assess its applicability to diverse groups within the target population. I employed several nested models with progressively stricter constraints on factor loadings, intercepts, and residuals, following the guidelines of Chen (2007) and Cheung and Rensvold (2002), to establish configural, metric, scalar, and residual invariance.

Factor analysis at the scale level

To evaluate the MPQ-35's internal construct validity at the higher construct level, I conducted an EFA on the primary trait scores, using the data from the total sample, with ML and geomin rotation. Several techniques were also used to determine the optimal number of factors to retain, including Kaiser's mineigen criterion, Cattell's scree test, Horn's parallel analysis, and Velicer's minimum average partial method. The solutions suggested by these methods were then assessed based on their interpretability, theoretical coherence, and model fit, following the same procedures as in the item-level analysis.

External construct validity

To assess the external construct validity of the MPQ-35, I computed Pearson's correlation coefficients between the 10 primary trait scales of the MPQ-35 and 19 personality and psychological constructs. I selected these constructs with the expectation that they would effectively validate the constructs measured by the MPQ-35. Convergent validity is evidenced by positive correlations between similar constructs, whereas discriminant validity is evidenced by negative correlations between dissimilar constructs (Carmines & Zeller, 1979). The constructs used in this study include the Big Five personality traits, psychological well-being, affect, life orientation, self-esteem, and sense of coherence. The convergent validity of the PEM scales (well-being, social potency, achievement, and social closeness) is expected to be supported by their positive correlations with constructs related to positive personality traits and psychological well-being (e.g., extraversion, conscientiousness, and personal growth). Their

discriminant validity should emerge from negative correlations with constructs associated with negative emotional states (e.g., neuroticism, negative affect). Similarly, the NEM scales (stress reaction, aggression, alienation) are anticipated to correlate positively with constructs indicative of negative emotional states, supporting their convergent validity, while showing negative correlations with positive personality traits and well-being constructs for discriminant validity. For the CON scales (control, traditionalism, harm avoidance), convergent validity is expected to be reflected in positive correlations with constructs emphasizing order and stability (e.g., conscientiousness, sense of control), and discriminant validity may be apparent through negative correlations with constructs linked to psychological flexibility and openness (e.g., openness, agency).

Model fit

I assessed the fit of measurement models using both incremental and absolute fit indices (Brown, 2015; Hu & Bentler, 1999; Kline, 2015). Incremental fit indices included the Comparative Fit index (CFI), while absolute fit indices included a nonsignificant chi-squared test, root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR). Generally, a CFI value greater than .90 suggests acceptable fit, while a value greater than .95 suggests excellent fit to the data. As for RMSEA/SRMR, values less than .08/.08 are typically indicative of a reasonable fit, and values less than .05/.05 reflect a close fit to the data. The chi-squared statistics were reported but were not used for the assessment. This decision was made because chi-squared statistics are sensitive to large sample sizes and tend to yield significant results even with minor discrepancies between the observed and model-implied covariance matrices (e.g., Kline, 2015). The criteria for measurement invariance testing included a change in CFI (ΔCFI) of no more than 0.01, a change in RMSEA (ΔRMSEA) of no more than 0.015, and a change in SRMR (ΔSRMR) of no more than 0.03 (Chen, 2007; Cheung & Rensvold, 2002). While the ΔCFI cutoff of 0.01 is less stringent than the recently recommended criterion of 0.002 (Meade et al., 2008; Somaraju et al., 2022), I considered this cutoff to better align with my overall omnibus tests of measurement equivalence, as opposed to item-specific nonequivalence, which may require a more stringent cutoff value.

Results

Descriptive statistics, internal consistencies, and mean comparisons

Table 2 display descriptive statistics, internal consistency coefficients (Cronbach's alpha and McDonald's omega), and gender and age group

Table 2. Descriptive statistics for the MPQ-35 primary traits for the overall sample, by gender, and by age groups, and mean comparisons.

MPQ Scales	Total sample				Sex						Age groups							
					Male		Female				55 Y & below		56–66 Y		67 Y & Above			
	M	*SD*	*α*	*ω*	*M*	*SD*	*M*	*SD*	*t*	*d*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*F*	η^2
Well-being	9.06	1.80	0.72	0.72	9.34	1.80	9.17	1.97	2.15*	0.09	8.95	2.03	9.22	1.93	9.43	1.78	11.22***	0.01
Social potency	10.12	2.39	0.71	0.74	10.75	2.31	10.07	2.45	6.98***	0.29	10.59	2.38	10.41	2.47	10.23	2.37	4.29*	0.00
Achievement	12.21	2.22	0.66	0.66	12.55	2.12	12.21	2.10	4.03***	0.16	12.54	2.08	12.54	2.12	12.12	2.10	12.92***	0.01
Social closeness	11.62	2.40	0.67	0.67	11.47	2.31	12.08	2.28	−6.60***	−0.26	11.59	2.42	11.66	2.40	12.05	2.16	10.58***	0.01
Stress reaction	5.96	2.21	0.73	0.73	5.81	2.21	5.63	2.23	1.90^{+}	0.08	6.19	2.29	5.63	2.20	5.51	2.16	18.33***	0.02
Aggression	5.38	1.79	0.66	0.69	5.54	1.88	5.00	1.55	7.95***	0.32	5.45	1.89	5.21	1.65	5.14	1.67	6.55**	0.01
Alienation	5.00	1.85	0.66	0.67	4.82	1.77	4.82	1.74	0.02^{ns}	0.01	5.10	1.79	4.88	1.83	4.62	1.64	15.71***	0.01
Control	9.74	1.51	0.61	0.65	9.94	1.49	9.94	1.44	0.00^{ns}	0.01	9.72	1.58	10.00	1.43	10.02	1.42	8.72***	0.01
Traditionalism	8.03	2.19	0.55	0.60	7.70	2.42	8.30	2.36	−5.98***	−0.25	7.71	2.32	7.63	2.50	8.55	2.27	38.80***	0.03
Harm avoidance	12.21	2.81	0.54	0.58	11.70	2.78	13.10	2.44	−13.54***	−0.54	11.89	2.79	12.43	2.68	12.84	2.57	24.63***	0.02

Note. *M*: mean; *SD*: standard deviation; *α:* Cronbach's alpha; *ω:* McDonald's omega; *t*: *t*-statistic that resulted from an independent samples *t*-test; *d*: Cohen's *d* effect size (values of *d* around 0.20, 0.50, and 0.80 are considered to represent small, medium, and large effect sizes, respectively); Y: years; *F*: *F*-statistic that resulted from a One-way ANOVA; η^2: partial eta squared effect size (values of η^2 around 0.01, 0.06, and 0.14 are considered to represent small, medium, and large effect sizes, respectively); the complete results of the One-way ANOVA and independent-sample *t* test are available upon request.

$^{ns}p>.10$; $^{+}p<.10$; $^{*}p<.05$; $^{**}p<.01$; $^{***}p<.001$.

differences for the MPQ-35 primary trait scales. The majority of scales exhibited acceptable internal consistency, except for the harm-avoidance and traditionalism scales, which showed lower estimates (α/ω = .54/.58 and .55/.60, respectively). The highest consistency was achieved by the stress reaction, well-being, and social potency scales (α/ω = .73/.73, .72/.72, and .71/.74, respectively). The independent sample *t*-test showed that males and females differed significantly on six of the 10 scales, with effect sizes ranging from small (0.01) to moderate (0.53). One-way ANOVA revealed significant age-group differences on all scales, with effect sizes ranging from negligible (0.00) to small (0.03). Therefore, conducting a measurement invariance test across gender and age groups is justified to ensure that the scale is measuring the same constructs with the same level of precision across these groups.

Internal construct validity of MPQ-35 at item-level

EFA of MPQ-35 items

EFA was conducted on the first random half of the sample. The factor retention criteria yielded different recommendations for the number of factors to retain. The MAP suggested retaining four factors, while the HPA and eigenvalue greater than 1 method suggested 9 factors. The scree plot's suggestion presented in Figure 1 is not entirely clear. However, the possibility of a four-factor model remains, and a nine-factor model is also likely. The first eigenvalue > 1.0 identified by EFA was 4.81, explaining 13.76% of the total variance. The ninth and 10th eigenvalues were 1.01 (2.89%) and 0.97 (2.77%), respectively. Together, the 10 factors accounted for 57.31% of the total variance.

When imposing the two suggested solutions (4- and 9-factor models), as well as the theorized 10-factor model, the results, as displayed in Table 3 (M1–M3), indicate that the 4-factor solution did not fit the data well. In contrast, both the 9-factor and 10-factor solutions showed acceptable fit, with the 10-factor model demonstrating slightly better fit than the 9-factor model. Reviewing the factor loadings from the two models indicates comparable loading patterns; however, the items within the well-being and achievement scales were grouped into one single factor in the 9-factor solution and divided into two separate factors in the 10-factor solution, aligning with their respective constructs. I chose the 10-factor model for subsequent analyses due to its superior fit and greater alignment with the MPQ theoretical framework. As shown in Table 4, this EFA model suggests salient factor loadings for most MPQ-35 items, with only three items displaying loadings below .30 (SP1, SC3, and TR1), and one item demonstrating cross-loadings on another factor (AC2). Overall, the EFA results provide supportive evidence for the internal construct validity of the MPQ-35, indicating an acceptable 10-factor model.

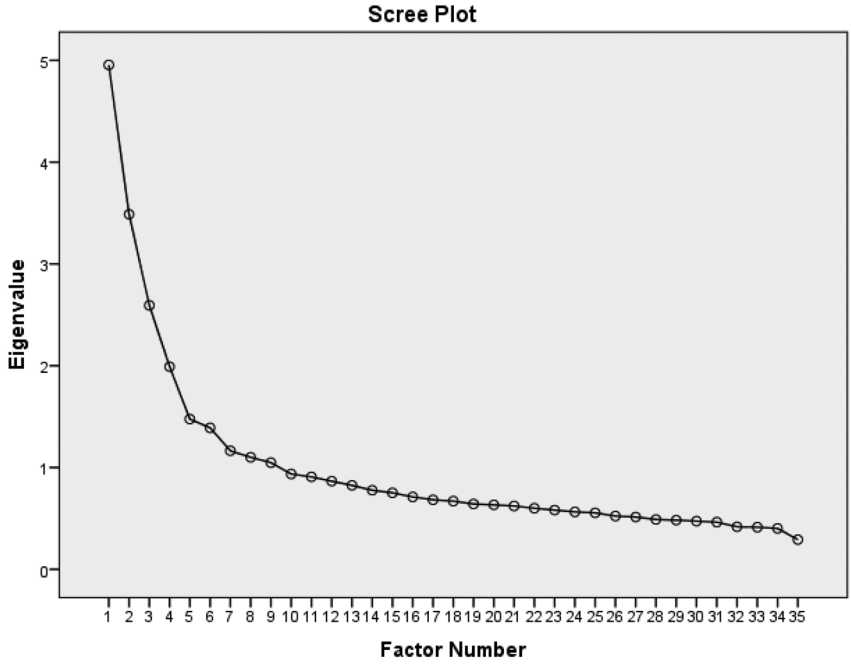

Figure 1. Scree plot of eigenvalues derived from the MPQ-35 item-level data.

CFA of MPQ-35 items

Testing a 10-factor CFA model on the second random half of the sample (Table 3; Model M4) revealed an unacceptable model fit, as indicated by the CFI value of .87. However, the RMSEA and SRMR were within the proposed cutoff values (0.05 for both). The factor loadings of the CFA model are reported in Table 4.

ESEM of MOQ-35 items

Testing a 10-factor ESEM model on the entire sample (Table 3; Model M5) demonstrated an excellent fit to the data, as indicated by a CFI of .98, RMSEA of .03, and SRMR of .01. As presented in Table 4, Factor loadings for most items were salient (> .30), with seven items (SP1, AC4, SC3, AG2, AG3, TR1) showing lower loadings. All cross-loadings were non-salient (< .30), indicating good discriminant validity. Lower factor loadings in ESEM models compared to traditional CFA models are common due to the estimation of unique variances and cross-loadings. The MPQ-35 factor intercorrelations generally align with its theoretical framework, with strong correlations between WB and AC (.49), WB and SP (.42), and SR and AG (.50), and weak correlations between SR and AL (0.01) and between CO and AL (0.01) (Table 4), further supporting its construct validity.

Table 3. Fit statistics for EFA, CFA, and ESEM Models of MPQ-35 with measurement invariance testing by gender and age group.

	Model	χ^2	χ^2 df	CFI	RMSEA	90% CI	SRMR	$\Delta s\chi^2$ (df)	ΔCFI	ΔRMSEA	ΔSRMR
	Item Level										
M1	4-factor EFA	2886.94***	461	.77	.06	[.06–.06]	.04				
M2	9-factor EFA	784.15***	316	.96	.03	[.03–.04]	.02				
M3	10-factor-EFA	564.98***	290	.97	.03	[.02–.03]	.02				
M4	10-factor CFA	1927.72***	515	.87	.05	[.04–.05]	.05				
M5	ESEM 10-factor model	764.44***	290	.98	.03	[.02–.03]	.01				
	Gender Invariance										
M6	Configural	1073.003***	580	.977	.025	[.023–.027]	.015	–	–	–	–
M7	metric (loadings)	1468.843***	830	.970	.024	[.022–.026]	.025	395.840(250) ***	−0.007	.001	.010
M8	scalar (loadings, thresholds)	1687.858***	855	.960	.027	[.025–.029]	.027	219.015(25) ***	−0.01	.003	.002
M9	Latent variance-covariance	1788.356***	910	.958	.027	[.025–.028]	.035	100.498(55) ***	−0.002	.001	.008
M10	Latent means	2256.454***	920	.936	.033	[.031–.034]	.049	468.098(10) ***	−0.022	.006	.014
	Age Invariance										
M11	Configural	1402.601***	870	0.975	0.026	[.023–.028]	0.017	–	–	–	–
M12	metric (loadings)	2062.415***	1370	0.967	0.024	[.022–.026]	0.028	659.814(500) ***	−0.008	−0.002	0.011
M13	scalar (loadings, thresholds)	2278.807***	1420	0.960	0.026	[.024–.028]	0.03	216.392(50) ***	−0.007	0.002	0.002
M14	Latent variance-covariance	2492.227***	1530	0.955	0.026	[.024–.028]	0.041	213.42(110) ***	−0.005	0	0.011
M15	Latent means	2797.566***	1550	0.941	0.03	[.028–.032]	0.047	305.339(20) ***	−0.014	−0.004	0.006
	Scale Level										
M16	2-factor EFA	620.14***	26	0.84	0.09	[.09–.10]	0.06				
M17	3-factor EFA	246.87***	18	0.94	0.07	[.06–.08]	0.03				
M18	4-factor-EFA	36.21***	11	0.99	0.03	[.02–.04]	0.01				

Notes: Model 1 to Model 3 were derived from EFA using the first randomly selected sample; Model 4 was derived from CFA using the second randomly selected sample; Models 5 to 15 were derived from ESEM using the total sample; and Model 16 to Model 18 were derived from EFA at scale score level using the total sample.
***$p < .001$.

Table 4. Factor loadings resulting from 10-factor EFA, CFA, and ESEM Models.

	EFA/ESEM										
	WB	SP	AC	SC	SR	AG	AL	CO	HA	TR	CFA
WB1	.48 |.39										.66
WB2	.64 |.44										.69
WB3	.64 |.58										.73
SP1		**.23 |.20**									.36
SP2		.80 |.79									.78
SP3		.84 |.68									.86
SP4		.43 |.34									.53
AC1			.51 |.37					***.20* |*.17***			.57
AC2			.54 |.39						***.32* |*.26***		.63
AC3			.52 |.39								.55
AC4			.47 |**.27**								.54
SC1				.73 |.60							.69
SC2				.54 |.51							.57
SC3				**.27 |.24**							.52
SC4				.56 |.55							.66
SR1					.47 |.37						.60
SR2					.64 |.57						.69
SR3					.80 |.75						.70
AG1						.71 |.42					.65
AG2						.44 |**.21**					.54
AG3						.38 |**.25**					.50
AG4						.73 |.41					.66
AL1							.71 |.62				.60
AL2							.46 |.41				.61
AL3							.49 |.39				.67
CO1								.63 |.36			.66
CO2								.74 |.51			.74
CO3								.42 |.31			.44
TR1										**.26 |.27**	.34
TR2										.60 |.54	.63
TR3										.76 |.81	.76

(*Continued*)

Table 4. Continued.

	EFA/ESEM										
	WB	SP	AC	SC	SR	AG	AL	CO	HA	TR	CFA
HA1									.56 |.50		.57
HA2									.50 |.47		.59
HA3									.55 |.59		.49
HA4									.37 |.46		.31
EFA Eigenvalues	4.81	3.60	2.50	1.88	1.55	1.37	1.25	1.12	1.01	0.97	
% of Variance	13.75	10.29	7.14	5.37	4.43	3.91	3.56	3.20	2.89	2.77	
	WB	**SP**	**AC**	**SC**	**SR**	**AG**	**AL**	**CO**	**HA**	**TR**	
WB	1	0.40	0.47	0.28	−0.27	−0.02^{ns}	−0.14	0.15	0.14	0.09	
SP	0.42	1	0.38	0.12	−0.05^{ns}	0.07	−0.08	0.04^{ns}	0.20	−0.09	
AC	0.49	0.33	1	0.11	−0.21	−0.05^{ns}	−0.09	0.27	0.08	0.05	
SC	0.32	0.15	0.10	1	−0.20	−0.15	−0.11	0.02^{ns}	−0.10	0.13	
SR	−0.27	−0.04^{ns}	−0.22	−0.22	1	0.48	0.49	0.05^{ns}	0.09	0.12	
AG	−0.06	0.10	−0.08	−0.11	0.43	1	0.45	−0.08	0.24	0.03	
AL	−0.15	−0.05^{ns}	−0.11	−0.07	0.50	0.48	1	0.02^{ns}	0.07^{ns}	0.18	
CO	0.18	0.05	0.28	0.01^{ns}	0.05^{ns}	−0.12	0.00^{ns}	1	−0.09	0.11	
HA	0.15	0.23	0.07	−0.12	0.10	0.26	0.08	−0.09	1	−0.16	
TR	0.07	−0.08	0.08	0.14	0.10	0.01^{ns}	0.16	0.09	−0.21	1	

Notes: The upper section of the table displays factor loadings resulting from EFA, CFA, and ESEM, while the lower section of the table displays the factors' correlation matrices. WB = Well-being; SP = Social Potency; AC = Achievement; SC = Social Closeness; SR = Stress Reactivity; AG = Aggression; AL = Alienation; CO = Control; TR = Traditionalism; HA = Harm Avoidance.

Upper section: The values within each cell are separated by a (|) symbol, denoting the factor loadings obtained from the 10-factor EFA model (left) and the 10-factor ESEM model (right). Only loadings equal to or exceeding .20 is presented for clarity and conciseness. Values that are **bold** indicate factor loading coefficients below .30, while values that are ***bold and italicized*** represent items with cross-loadings equal to or above .20 on factors other than their respective factors.

Lower section: The upper diagonal of the matrix represents factor intercorrelations resulting from the EFA model, while the lower diagonal represents factor intercorrelations resulting from the ESEM model. Correlation coefficients with (^{ns}) symbol indicate nonsignificance, while the remaining coefficients without any sign indicate significance below .05.

$^{ns}p > .05$.

Measurement invariance by gender

Measurement invariance by gender was assessed on the 10-factor ESEM solution of the MPQ-35. Five increasingly restrictive models were compared across male and female genders (M6 to M10 in Table 3). The unconstrained configural model had an acceptable fit ($\chi2 = 1073.003$, $df = 580$, $p < .001$, CFI = .977, RMSEA = .025, SRMR = .015) and, as constraints were imposed, the fit indices for metric, scalar, and strict invariance remained within acceptable thresholds. However, when latent means invariance was imposed, a significant worsening of the model fit was observed (ΔCFI = .022). These results suggest measurement invariance of the MPQ-35 measurement model across genders up to the latent variance-covariance matrix. Such non-equivalence in the latent means, however, does not preclude the comparison of factor means across genders. Instead, it indicates the likelihood of mean differences between men and women.

Measurement invariance by age groups

Five increasingly restrictive models were compared across three age groups (M11 to M15 in Table 3). The unconstrained configural model yielded an acceptable fit for the three age groups ($\chi2 = 1402.601$, $df = 870$, $p < .001$, CFI = .975, RMSEA = .026, SRMR = .017). Imposing constraints on the model resulted in only a slight drop in fit indices for metric, scalar, and strict invariance. However, when latent means invariance was imposed, there was a significant drop in model fit as indicated by a ΔCFI of .014, although ΔRMSEA and ΔSRMR showed only minor changes (.004 and .006, respectively). Similar to the gender invariance test, these results suggest measurement invariance of the MPQ-35 measurement model across age groups up to the latent variance-covariance matrix. Despite this nonequivalence in the latent means, comparisons of factor means between age groups are still feasible, though mean differences between them are likely.

Internal construct validity at scale level

Applying EFA to scale-level data from the entire sample, I found that the factor retention criteria yielded varying recommendations regarding the number of factors to retain. The MAP criterion suggested 2 factors, while the eigenvalue greater than 1 and scree plot (Figure 2) suggested 3 factors. Additionally, the HPA suggested 3 components and 4 factors. The three eigenvalues > 1.0 identified by EFA were: 2.41 (24.12% variance) for the first factor, 1.69 (16.92%) for the second factor, and 1.39

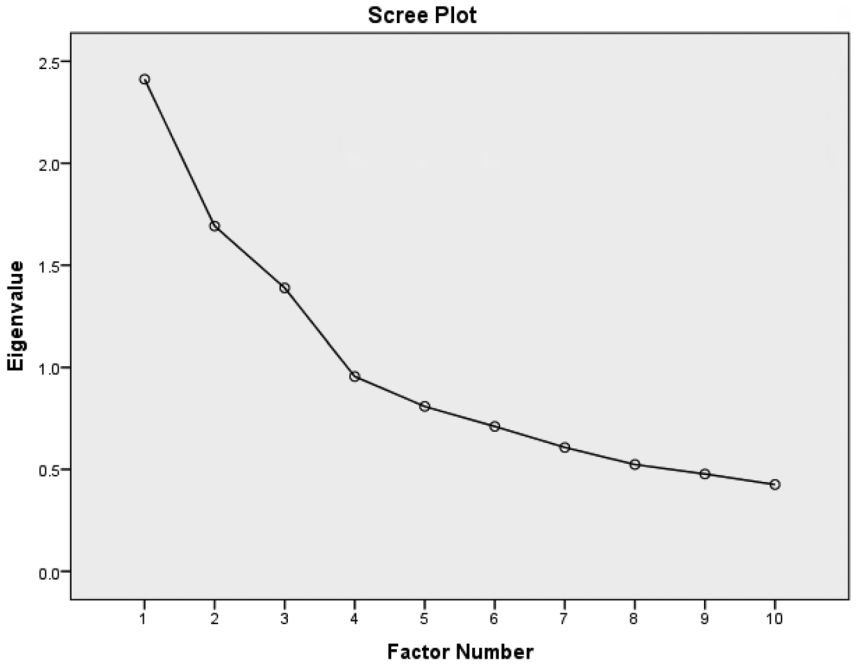

Figure 2. Scree plot of eigenvalues derived from the MPQ-35 scale-level data.

(13.88%) for the third factor, together accounting for 54.92% of the variance.

The subsequent EFA on the three suggested factor solutions (2-, 3-, and 4-factor) yielded mixed results for model fit (Table 3, M16 to M18). While the 2-factor solution had poor model fit, the 3-factor and 4-factor solutions showed relatively stronger support, with acceptable and excellent model fit, respectively. The 3-factor solution presented in Table 5 closely aligned with the theoretical structure of the MPQ-35, featuring three higher-order domains (PEM, NEM, and CON). The 4-factor solution (Table 5) divided the PEM domain into Agentic and communal factors, in line with some prior MPQ EFA studies, and identified the AC scale as the primary indicator for the agentic factor. However, its viability was weakened by the lack of multiple salient indicator loadings. Considering statistical and theoretical aspects, the 3-factor model had stronger support. Intercorrelations among the three factors showed negative correlations between NEM and CON (–0.38) and PEM and NEM (–0.23), and a weak positive correlation between PEM and CON (.03). These results indicate that NEM is negatively associated with CON and that PEM is relatively independent of CON.

Table 5. Factor loadings of MPQ-35 primary trait scales derived from EFA.

	3-factor solution			4-factor solution			
Scale	PEM	NEM	CON	Communal-PEM	Agentic-PEM	NEM	CON
Well-being	0.77	−0.00	0.12	0.65	0.17	−0.03	0.01
Social potency	0.59	−0.01	−0.20	0.49	0.12	0.03	−0.27
Achievement	0.64	0.12	−0.01	0.02	1.04	−0.01	−0.01
Social closeness	0.33	−0.02	0.33	0.52	−0.07	−0.01	0.29
Stress reaction	−0.19	0.64	−0.01	−0.27	0.01	0.60	−0.01
Aggression	0.04	0.61	−0.15	0.05	−0.09	0.67	−0.22
Alienation	−0.03	0.73	0.16	−0.12	0.06	0.65	0.11
Control	0.19	0.13	0.33	0.02	0.26	0.02	0.32
Traditionalism	0.01	0.29	0.54	0.07	0.02	0.19	0.48
Harm avoidance	−0.20	−0.07	0.39	−0.03	−0.10	−0.11	0.38
Eigenvalues	2.41	1.69	1.39				.96
% of explained	24.11%	16.92	13.88				9.55%
PEM	1			1			
PEM				0.40*	1		
NEM	−0.23*	1		−0.20*	−0.01	1	
CON	.03	−0.38*	1	−0.02	−0.14*	−0.11*	1

Notes: EFA = Exploratory factor analysis; underlined values indicate the cross-loading of items above .20 on factors other than their respective factors; PEM = Positive Emotionality; NEM = Negative Emotionality; CON = Constraint (CON); the lower section of the table displays the correlation matrix, which represents the intercorrelations among the factors.

*$p > .05$.

External construct validity of MPQ-35 scales

In this phase of the study, I investigated the external construct validity of the MPQ-35 scales by assessing their associations with 19 personality and psychological constructs using Pearson's correlation coefficient. The results, presented in Table 6, show that all scales within the positive emotionality domain were significantly positively correlated with extraversion, conscientiousness, agreeableness, openness, agency, six psychological well-being constructs, optimism, positive affect, self-esteem, sense of control, and life satisfaction, while being significantly negatively correlated with neuroticism, pessimism, and negative affect. In contrast, the scales in the negative emotionality domain showed significant negative correlations with constructs that were significantly positively associated with the scales in the PEM domain but were positively correlated with constructs that were negatively correlated with the PEM domain scales. The constraint domain scales exhibited a mixed pattern of weak or insignificant relationships with the Big Five factors of personality and the agency scale, psychological well-being constructs, and the other seven constructs. Overall, the associations of the PEM and NEM traits with the examined constructs were much clearer than those of the CON traits. Specifically, there were six large, 37 medium, and 33 small correlations for the PEM traits, and six large, 23 medium, and 28 small correlations for the NEM traits. In contrast, only one medium

Table 6. MPQ-35 traits' correlations with 19 personality and psychological constructs.

	WB	SP	AC	SC	SR	AG	AL	CO	TR	HA
The Big-5 personality inventory										
Extraversion	**.51**	*.45*	*.35*	*.48*	−0.22	−0.09	−0.13	.05*	.05*	−0.00
Conscientiousness	.26	.22	*.45*	.13	−0.24	−0.17	−0.20	*.35*	.03^{+}	.06**
Agreeableness	.26	.14	.21	*.41*	−0.14	−0.20	−0.06**	.19	.15	.16
Openness	*.47*	*.46*	*.49*	.21	−0.17	−0.06**	−0.13	.13	−0.19	−0.15
Neuroticism	*−0.33*	−0.19	−0.15	−0.15	**.63**	*.30*	*.36*	−0.02^{ns}	.01^{ns}	.03^{ns}
Agency	*.39*	**.55**	*.38*	.13	−0.08	.15	−0.09	.01^{ns}	−0.06**	−0.13
Psychological well-being										
Autonomy	*.32*	*.36*	*.31*	.08	*−0.34*	−0.15	*−0.31*	.12	−0.07	−0.03^{ns}
Environmental mastery	*.48*	*.30*	.27	.28	**−0.57**	−0.28	*−0.45*	.11	−0.01^{ns}	.01^{ns}
Personal growth	**.51**	*.38*	*.42*	.28	*−0.39*	−0.29	*−0.35*	.11	−0.12	−0.04^{+}
Positive relations with others	*.43*	.28	.20	**.50**	*−0.42*	*−0.34*	*−0.40*	.12	.06**	.10
Purpose in life	*.48*	*.34*	*.36*	.28	*−0.39*	−0.27	*−0.36*	.18	.00^{ns}	.03^{ns}
Self-acceptance	**.56**	*.38*	*.30*	*.31*	**−0.50**	−0.25	*−0.40*	.12	−0.00^{ns}	.00^{ns}
Positive and negative affect schedule										
Positive affect	*.49*	.28	*.33*	.28	*−0.37*	−0.19	−0.26	.13	.06	−0.01^{ns}
Negative affect	−0.28	−0.12	−0.11	−0.18	**.50**	.29	*.33*	−0.04^{ns}	−0.01^{ns}	−0.03^{ns}
Life orientation test										
Optimism	**.55**	*.32*	*.32*	.28	*−0.36*	−0.20	−0.24	.11	.04*	−0.04*
Pessimism	*−0.34*	−0.25	−0.15	−0.18	*.42*	*.35*	**.51**	−0.05*	.08	−0.03^{ns}
Other subjective well-being										
Self-esteem	*.48*	*.34*	*.30*	.25	**−0.54**	−0.26	*−0.41*	.10	−0.05*	−0.00^{ns}
Life satisfaction	*.35*	.16	.17	.16	*−0.32*	−0.20	*−0.35*	.09	.01^{ns}	.033^{+}
Sense of Control	*.46*	*.32*	*.32*	.20	*−0.48*	−0.24	*−0.46*	.08	−0.11*	−0.06**

Notes: Values without a significance sign are significant at a level below .001. Effect sizes are small for *r* between .10 and .30 (not emphasized), medium for *r* between .30 and .50 (italicized), and large for *r* between .50 and 1.00 (bold), according to Cohen (1988).
$^{ns}p > .10$. $^{+}p < .10$. * $p < .05$. ** $p < .01$.

correlation was observed for the CON domain, with the remainder being either weak or insignificant. These findings suggest that the PEM and NEM traits of the MPQ-35 have good external construct validity, while the CON domain may have less effectiveness in predicting external constructs.

Discussion

The present study aimed to evaluate the psychometric properties of the MPQ-35, a short version of the MPQ adapted specifically for the Midlife in the United States Study (MIDUS) to meet the need for a brief personality measure. Despite its brevity, the MPQ-35 has demonstrated psychometric qualities that are somewhat reflective of those often observed in longer versions of the MPQ, such as the MPQ-BF (Patrick et al., 2002), MPQ-SF (Javdani et al., 2014; Patrick et al., 2013), MPQ-BF-NL (Eigenhuis et al., 2013), and the abbreviated research form (Marquardt et al., 2021). Results from various analyses indicated that the MPQ-35 retains the factor structure of the original MPQ 10 lower-order and three higher-order factor dimensions. These findings indicate that the MPQ-35 potentially achieves a favorable balance between brevity, validity, and reliability, suggesting its practicality for use in various research settings.

MPQ scales' internal consistency

While the majority of MPQ-35 scales demonstrated acceptable reliability coefficients (α and ω > .60), the traditionalism and harm-avoidance scales exhibited lower coefficients. These results highlight the complexity of measuring certain personality constructs with only a few items. The inherent heterogeneity within these constructs, as suggested by previous research on similar complex constructs (e.g., Del Rosario & White, 2005; Romero et al., 2012; Thørrisen et al., 2021), may have contributed to the compromised internal consistency observed in these scales. This pattern might also be linked to the demographics of the study's sample, which predominantly consists of older individuals ages 39 to 93 years, with 88.7% identifying as White. It is plausible that these specific traits are more susceptible to interpretation variations influenced by generational and cultural factors, potentially explaining the greater response variability observed. In contrast, other traits might be less influenced by such demographic factors, resulting in more consistent responses across the sample.

The initial validation study of the MPQ-35 in MIDUS using three large samples also found lower coefficients for these two scales (.56, .57, .64 for traditionalism and .56, .56,.59 for Harm avoidance; Ryff et al., 2021). Despite this limitation, MIDUS researchers have

continued to use the MPQ-35, suggesting that its other strengths outweigh this potential drawback. Future research could explore ways to improve the reliability of the harm avoidance and traditionalism scales in the MPQ-35 while maintaining its brevity, such as replacing some of their items with others, examining test-retest reliability, and considering alternative psychometric evidence such as item-response theory (IRT). Nevertheless, it is important to remember that internal consistency is only one facet of reliability, which itself is only one aspect of validation (American Educational Research Association et al., 2014); thus, a comprehensive examination of various aspects is also essential.

Internal construct validity

Applying multiple methods to evaluate the factor structure of the MPQ-35 was crucial, as each method has unique assumptions and may yield different results. EFA is data-driven, CFA is based on a priori hypotheses, and ESEM integrates both approaches, allowing for a more nuanced exploration of the data. These analyses, conducted at both the item and scale levels, collectively provide deeper insights into the construct validity of the MPQ-35 and strengthen the evidence for its effectiveness in measuring the intended personality traits. The EFA results were found to replicate the theorized factor structure of the MPQ, providing evidence for the MPQ-35's validity in measuring 10 primary traits and three broad personality dimensions. This is consistent with the factor structure underlying Tellegen's model (e.g., Eigenhuis et al., 2013; Johnson et al., 2008) and highlights the MPQ's consistency across various forms, populations, and cultures. However, the CFA with 10 lower-order factors did not show an adequate fit for the MPQ-35 data. This aligns with research indicating that traditional ICM-CFA models might be too restrictive to effectively fit personality test data (Church & Burke, 1994; Hopwood & Donnellan, 2010; Sellbom & Tellegen, 2019). In contrast, an ESEM model with 10 lower-order factors demonstrated a good fit for the data. This underscores the complex and interconnected nature of MPQ-35 traits, which ESEM captures more accurately due to its enhanced ability to account for these intricate trait interconnections (Asparouhov & Muthén, 2009; Marsh et al., 2009, 2010). Several recent studies have demonstrated the advantages of using ESEM over CFA in assessing the factor structure of personality inventories (e.g., Booth & Hughes, 2014; Marsh et al., 2010; Perera et al., 2015). Hopwood and Donnellan (2010) point out the limitations of CFA, attributing them to the complex and multidimensional nature of personality constructs. They suggest ESEM as a more suitable alternative, as it can accommodate both correlated and uncorrelated factors within a single model, thus providing a more accurate representation of complex constructs like personality traits. Therefore, considering the multidimensional nature of the MPQ-35 scales,

ESEM emerges as a promising approach for future studies aiming to explore the factor structure of the MPQs.

Two aspects of the EFA results warrant further attention: the emergence of potential 4- and 5-factor models at the item level, and the emergence of a potential 4-factor model at the scale level. The scree plot from the item-level EFA results indicates that 4- or 5-factor models are conceivable to some extent. This observation may reflect the viability of fewer higher-order constructs in explaining personality traits, akin to the Big Five model, suggesting potential similarities across various personality models. These findings hint at the possibility of a unified model of personality that transcends specific trait differences across various frameworks. Such a model could bridge the existing gaps between models and offer a more holistic approach to understanding personality research.

The scale-level EFA results for the MPQ-35 are in alignment with previous MPQ literature, providing evidence that supports both a 3-factor (PEM, NEM, and CON) and potentially a 4-factor (PEM-Agentic, PEM-Communal, NEM, and CON) structure. However, the viability of the agentic PEM factor in the 4-factor model was invalided by the presence of only one salient indicator, the achievement scale. The limited number of MPQ-35 scale scores available for analysis, which includes only 10 scales, likely contributed to this inadequacy. Consequently, this may favor a unified agentic-communal factor representing Positive Emotionality, as opposed to separate factors for agentic traits (i.e., achievement, well-being, assertiveness, and ambition) and communal traits (i.e., social potency, social closeness, and empathy). This interplay between a single PEM factor and two distinct PEM factors (often called the "big two") is not uncommon and is frequently observed in various personality models (see Abele et al., 2016; Abele & Wojciszke, 2007, 2014; Chen et al., 2019; Rucker et al., 2018; Trapnell & Paulhus, 2012).

Measurement invariance

Our gender invariance testing with the 10-factor ESEM model on the MPQ-35 revealed that both males and females share similar interpretations of the constructs and have a similar conceptual basis when responding to the MPQ-35 items. Higher levels of invariance were observed up to the latent variance-covariance level, suggesting a consistent factor structure for the MPQ-35 across genders. The observation of limited invariance at the level of latent means aligns with prior research, which indicates that gender invariance in personality inventories may not always extend to mean scores (Dong & Dumas, 2020; Fonseca-Pedrero et al., 2011; Gomez et al., 2022; Gustavsson et al., 2008; Picconi et al., 2018). Potential differences in mean scores between males and females may arise

from gender-specific factors, including social desirability, self-perceptions, or response biases (Cross & Madson, 1997; Eagly & Steffen, 1984; Heilman & Okimoto, 2007). Future studies are encouraged to investigate factors that may influence the measurement equivalence of the MPQ-35 across gender groups.

Regarding age invariance, the results indicated that the MPQ-35 demonstrated measurement invariance across the three age groups up to the level of latent variance-covariance but not at the level of latent mean. This suggests a consistent factor structure for the MPQ-35 across different age groups, with a likelihood of mean differences. The findings are consistent with previous studies that have reported high level of measurement invariance across different age groups for other personality inventories (Dong & Dumas, 2020; Gustavsson et al., 2008; Picconi et., 2018). This indicates that the MPQ-35 is a valid measure of personality traits across different age groups. It is important to note, however, that the present study examined only the age range of 39–93 years and, therefore, it remains unclear whether the same level of invariance observed for the MPQ-35 applies to individuals younger than this age range. Further research is needed to assess the measurement invariance of the MPQ-35 in younger age groups, in order to determine the generalizability of the findings.

External construct validity

This research contributes to the understanding of the MPQ-35's construct validity by investigating the associations between its 10 primary traits and 19 unique personality and psychological constructs. This offers important insights into the nature of these connections, furthering our understanding of such relationships in a broader context. The results revealed that PEM domain scales positively correlated with positive affect and well-being constructs, while negatively correlating with neuroticism, pessimism, and negative affect. In contrast, NEM domain scales displayed a reverse association, negatively correlating with positive psychological functioning and positively with neuroticism, pessimism, and negative affect. The CON domain showed a mixed pattern, with mostly weak or insignificant relationships to most constructs, indicating weaker associations compared to the PEM and NEM domains.

The findings align with earlier research on various MPQ versions (Church & Burke, 1994; Eigenhuis et al., 2013; Javdani et al., 2014; Sellbom et al., 2022; Tellegen, 1982; Tellegen & Waller, 2008), highlighting the MPQ-35's external validity and the significance of the PEM and NEM domains in capturing positive and negative psychological functioning. Future research should scrutinize the CON scales within the MPQ-35, analyze their connections to other personality and psychological dimensions, and investigate more precise alternative measures to enhance

external construct validity. Differing from prior MPQ psychometric studies, this research explored a wider array of constructs related to normal personality, in line with the MPQ-35's intended function as a research instrument. In contrast, studies of longer MPQ forms have typically focused on a smaller number of constructs, primarily emphasizing pathological personality traits, where extended versions of the MPQ are more suitable for providing comprehensive evaluations.

Finally, despite the existence of other brief forms of the MPQ, the psychometric examination of the MPQ-35 remains significant. The MPQ-35 was introduced in the MIDUS study in 2004/2006 (Ryff et al., 2021) and has been in use for almost two decades, preceding many of the other abbreviated versions of the MPQ. The availability of MIDUS data as open-source material has resulted in a substantial body of research and numerous publications from various MIDUS waves, which necessitates a thorough validation effort for the MPQ-35. Additionally, designed as a concise measure for Tellegen's model of personality, the MPQ-35 can be a practical choice for assessing personality traits within this framework, especially suitable for comprehensive surveys and screening contexts. Shorter questionnaires like the MPQ-35 are often found to improve response and completion rates, thereby enhancing the overall survey design's value (Galesic & Bosnjak, 2009; Kost & da Rosa, 2018). Participants generally prefer shorter surveys as they are more engaging and easier to complete (Dillman et al., 2014), a preference that becomes particularly pronounced in online environments, where shorter attention spans may lead to the abandonment of lengthy surveys (Kees et al., 2017). Consequently, this validation of the MPQ-35, the shortest form of the MPQ, highlights its strengths and limitations, providing the MIDUS research team with evidence informing its use in future surveys.

Limitations

Despite several strengths in the present study, it is not without limitations. First, validation studies typically analyze multiple samples for cross-validation purposes. However, this wasn't feasible in the current analysis due to the reliance on preexisting data. Nonetheless, this limitation was partially addressed by the use of a large sample size and dividing the data into two random halves for factor analysis, which may help to minimize the drawback. Second, the absence of a full-length MPQ form as a validator might have constrained our understanding of the new scales' performance compared to the original scales or the brief version's equivalence to the full-length version. While incorporating another validated MPQ form would have offered

a more robust assessment, this wasn't possible due to using readily available data. However, including 19 alternative constructs, such as the five big personality factors, might have partly compensated this limitation. Future research could explore MPQ-35's differential validity by comparing it with longer MPQ versions (e.g., MPQ-55, MPQ-155, MPQ-276) and assessing the impact of questionnaire length on MPQ's measurement properties.

Third, the exclusion of the absorption domain in the MPQ-35 may concern researchers interested in studying the complete spectrum of Tellegen's personality model. It may also raise questions about its factor structure and validity compared to longer MPQ versions. However, since absorption is not associated with the three higher-order factors proposed by Tellegen (1982), its absence likely has minimal impact on the MPQ-35's factor structure. Fourth, although the MPQ-35 replicates the three higher-order factors of the MPQ model, this study did not create composite scores to evaluate external construct validity. One reason for this is that constructing such scores typically requires regression-based equations that are derived from a comprehensive MPQ inventory. Additionally, scores for the higher-order factors were not readily computed in the MIDUS data set, unlike those for primary traits. This suggests that creating composite scores for the higher-order factors may not be recommended for a brief inventory like the MPQ-35.

Fifth, the predominantly White composition of the sample and the age range of 39–93 years may restrict the generalization of the findings to younger and more diverse populations. Although the large sample size and diverse socioeconomic backgrounds of the participants help to address concerns about generalizability, caution should still be exercised. Future research could benefit from deliberately oversampling underrepresented groups to further enhance the generalizability of the outcomes. Last, although it is known that the MPQ-35 items were adapted from the original MPQ for use in the MIDUS, the lack of detailed information on the item selection process might raise concerns. However, the MPQ developers are renowned for their rigor and caution in constructing, improving, or abbreviating MPQ forms, as demonstrated in various studies (e.g., Patrick et al., 2002; Tellegen & Waller, 2008). The validation studies of MPQs clearly show that proposing abbreviated MPQ forms typically follows rigorous item selection methods, such as classical test theory and item response theory, along with comprehensive psychometric assessment. These detailed reports, combined with the robust psychometric properties demonstrated in the current study, provide confidence in the methods used to select the items included in the MPQ-35, which may address concerns regarding item selection.

As a final point, it is important to acknowledge that the MPQ-35, a short version adapted from the full-length MPQ, is a proprietary instrument owned by the University of Minnesota Press. Although I obtained approval to use the MPQ-35 in this research, the form was specifically developed for the MIDUS surveys and is not an official alternative form of the MPQ. None of the MPQ's copyrighted developers were involved in this study, and using the MPQ-35 outside the MIDUS context requires approval from the copyright holder.

Conclusion

The psychometric examination of the MPQ-35 conducted in this study holds significance. It provides support for the construct validity of the MPQ-35 and its utility in assessing 10 personality traits. The use of rigorous statistical analyses on a large, diverse sample enhances the validity evidence obtained in this research. The 10-factor structure validated in this study, along with its high level of invariance across gender and age groups, suggests that the MPQ-35 corresponds to the theorized Tellegen's model underlying its structure. The correlations between the MPQ-35 scale scores and those of other scales measuring the same or different constructs indicate that many possess adequate convergent and discriminant validity. This outcome implies that the MPQ-35 scales, particularly within the PEM and NEM domains, effectively measure intended traits and differentiate them from unrelated constructs. Nonetheless, scales within the CON domain warrant further analysis through comparisons with more closely related or divergent constructs to fully ascertain their validity. Despite certain limitations, the findings of the current study suggest that the MPQ-35, with its advantages of brevity and ease of use, represents a promising tool for researchers aiming to assess Tellegen's model of personality. Future research could aim at enhancing the psychometric properties of the MPQ-35 by addressing the limitations identified in this study and further validating its applicability across more diverse populations and contexts.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Author's note

Data were obtained from the third wave of the National Survey of Midlife Development in the United States (MIDUS III), 2013–2014, which are made available by the Inter-University Consortium for Political and Social Research (ICPSR) at the University of Michigan (https://www.icpsr.umich.edu).

ORCID

Khalid ALMamari http://orcid.org/0000-0001-9699-5783

References

Abele, A. E., Hauke, N., Peters, K., Louvet, E., Szymkow, A., & Duan, Y. (2016). Facets of the fundamental content dimensions: Agency with competence and assertiveness – Communion with warmth and morality. *Frontiers in Psychology*, *7*, 1810. https://doi.org/10.3389/fpsyg.2016.01810

Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology*, *93*(5), 751–763. https://doi.org/10.1037/0022-3514.93.5.751

Abele, A. E., & Wojciszke, B. (2014). Communal and agentic content in social cognition: A dual perspective model. In *Advances in experimental social psychology* (Vol. 50, pp. 195–255). Academic Press.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling*, *16*(3), 397–438. https://doi.org/10.1080/10705510903008204

Ben-Porath, Y. S., Almagor, M., Hoffman-Chemi, A., & Tellegen, A. (1995). A cross-cultural study of personality with the Multidimensional Personality Questionnaire. *Journal of Cross-Cultural Psychology*, *26*(4), 360–373. https://doi.org/10.1177/0022022195264002

Booth, T., & Hughes, D. J. (2014). Exploratory structural equation modeling of personality data. *Assessment*, *21*(3), 260–271. https://doi.org/10.1177/1073191114528029

Bouchard, T. J., Jr,., & Waller, N. (2017). The recaptured scale technique applied to the multidimensional personality questionnaire constructs: A replication across item format and gender. *Personality and Individual Differences*, *109*, 77–82. https://doi.org/10.1016/j.paid.2016.11.057

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford publications.

Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research*, *36*(1), 111–150. https://doi.org/10.1207/S15327906MBR3601_05

Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment*. Sage publications.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Chen, X., Li, M., & Wei, Q. (2019). Agency and communion from the perspective of self versus others: The moderating role of social class. *Frontiers in Psychology*, *10*, 2867. https://doi.org/10.3389/fpsyg.2019.02867

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233–255.

Church, A. T. (1994). Relating the Tellegen and five-factor models of personality structure. *Journal of Personality and Social Psychology*, *67*(5), 898–909. https://doi.org/10.1037/0022-3514.67.5.898

Church, A. T., & Burke, P. J. (1994). Exploratory and confirmatory tests of the big five and Tellegen's three-and four-dimensional models. *Journal of Personality and Social Psychology*, *66*(1), 93–114. https://doi.org/10.1037/0022-3514.66.1.93

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334. https://doi.org/10.1007/BF02310555

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302. https://doi.org/10.1037/h0040957

Cross, S. E., & Madson, L. (1997). Models of the self: Self-construals and gender. *Psychological Bulletin*, *122*(1), 5–37. https://doi.org/10.1037/0033-2909.122.1.5

Del Rosario, P. M., & White, R. M. (2005). The narcissistic personality inventory: Test–retest stability and internal consistency. *Personality and Individual Differences*, *39*(6), 1075–1081. https://doi.org/10.1016/j.paid.2005.08.001

De Vries, R. E. (2013). The 24-item brief HEXACO inventory (BHI). *Journal of Research in Personality*, *47*(6), 871–880. https://doi.org/10.1016/j.jrp.2013.09.003

DiLalla, D. L., Gottesman, I. I., Carey, G., & Vogler, G. P. (1993). Joint factor structure of the Multidimensional Personality Questionnaire and the MMPI in a psychiatric and high-risk sample. *Psychological Assessment*, *5*(2), 207–215. https://doi.org/10.1037/1040-3590.5.2.207

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method*. John Wiley & Sons.

Dong, Y., & Dumas, D. (2020). Are personality measures valid for different populations? A systematic review of measurement invariance across cultures, gender, and age. *Personality and Individual Differences*, *160*, 109956. https://doi.org/10.1016/j.paid.2020.109956

Donnellan, M. B., Hopwood, C. J., & Wright, A. G. (2012). Revaluating the evidence for the general factor of personality in the multidimensional personality questionnaire: Concerns about Rushton and Irwing (2009). *Personality and Individual Differences*, *52*(3), 285–289. https://doi.org/10.1016/j.paid.2011.10.018

Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology*, *46*(4), 735–754. https://doi.org/10.1037/0022-3514.46.4.735

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. https://doi.org/10.1177/1073191112444920

Eigenhuis, A., Kamphuis, J. H., & Noordhof, A. (2015). Personality differences between the United States and the Netherlands: The influence of violations of measurement invariance. *Journal of Cross-Cultural Psychology*, *46*(4), 549–564. https://doi.org/10.1177/0022022115570671

Eigenhuis, A., Kamphuis, J. H., & Noordhof, A. (2017). Personality in general and clinical samples: Measurement invariance of the multidimensional personality questionnaire. *Psychological Assessment*, *29*(9), 1111–1119. https://doi.org/10.1037/pas0000408

Eisenbarth, H., Lilienfeld, S. O., & Yarkoni, T. (2015). Using a genetic algorithm to abbreviate the Psychopathic Personality Inventory–Revised (PPI-R). *Psychological Assessment*, *27*(1), 194–202. https://doi.org/10.1037/pas0000032

Fonseca-Pedrero, E., Paino, M., Lemos-Giráldez, S., Sierra-Baigrie, S., & Muñiz, J. (2011). Measurement invariance of the Schizotypal Personality Questionnaire-

Brief across gender and age. *Psychiatry Research*, *190*(2-3), 309–315. https://doi.org/10.1016/j.psychres.2011.05.021

Fulkerson, J. A., Keel, P. K., Leon, G. R., & Dorr, T. (1999). Eating-disordered behaviors and personality characteristics of high school athletes and nonathletes. *International Journal of Eating Disorders*, *26*(1), 73–79. https://doi.org/10.1002/(SICI)1098-108X(199907)26:1<73::AID-EAT9>3.0.CO;2-F

Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly*, *73*(2), 349–360. https://doi.org/10.1093/poq/nfp031

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, *4*(1), 26–42. https://doi.org/10.1037/1040-3590.4.1.26

Gomez, R., Griffiths, M. D., & Stavropoulos, V. (2020). Internal and external validity of the brief version of the Multidimensional Personality Questionnaire: Exploratory structural equation modelling. *PLoS One*, *15*(4), e0229470. https://doi.org/10.1371/journal.pone.0229470

Gomez, R., Watson, S., Brown, T., & Stavropoulos, V. (2022). Personality inventory for DSM–5-Brief Form (PID-5-BF): Measurement invariance across men and women. *Personality Disorders*, *14*(3), 334–338. https://doi.org/10.1037/per0000569

Gosling, S. D., Rentfrow, P. J., & Swann, W. B.Jr, (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*(6), 504–528. https://doi.org/10.1016/S0092-6566(03)00046-1

Gustavsson, J. P., Eriksson, A.-K., Hilding, A., Gunnarsson, M., & Ostensson, C.-G. (2008). Measurement invariance of personality traits from a five-factor model perspective: Multi-group confirmatory factor analyses of the HP5 inventory. *Scandinavian Journal of Psychology*, *49*(5), 459–467. https://doi.org/10.1111/j.1467-9450.2008.00654.x

Harkness, A. R., Tellegen, A., & Waller, N. (1995). Differential convergence of self-report and informant data for Multidimensional Personality Questionnaire traits: Implications for the construct of negative emotionality. *Journal of Personality Assessment*, *64*(1), 185–204. https://doi.org/10.1207/s15327752jpa6401_13

Heilman, M. E., & Okimoto, T. G. (2007). Why are women penalized for success at male tasks?: The implied communality deficit. *The Journal of Applied Psychology*, *92*(1), 81–92. https://doi.org/10.1037/0021-9010.92.1.81

Hopwood, C. J., & Donnellan, M. B. (2010). How should the internal structure of personality inventories be evaluated? *Personality and Social Psychology Review*, *14*(3), 332–346. https://doi.org/10.1177/1088868310361240

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185. https://doi.org/10.1007/BF02289447

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Javdani, S., Finy, M. S., & Verona, E. (2014). Evaluation of the validity of the Multidimensional Personality Questionnaire–Simplified-Wording Form (MPQ-SF) in adolescents with treatment histories. *Assessment*, *21*(3), 352–362. https://doi.org/10.1177/1073191113504617

John, O. P. (1990). The "Big Five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. A. Pervin (Ed.), *Handbook of Personality: Theory and Research* (pp. 66–100). Guilford Press.

Johnson, W., Spinath, F., Krueger, R. F., Angleitner, A., & Riemann, R. (2008). Personality in Germany and Minnesota: An IRT-based comparison of MPQ self-reports. *Journal of Personality*, *76*(3), 665–706. https://doi.org/10.1111/j.1467-6494.2008.00500.x

Jonason, P. K., & Webster, G. D. (2010). The dirty dozen: A concise measure of the dark triad. *Psychological Assessment*, *22*(2), 420–432. https://doi.org/10.1037/a0019265

Jones, D. N., & Paulhus, D. L. (2014). Introducing the short dark triad (SD3) a brief measure of dark personality traits. *Assessment*, *21*(1), 28–41. https://doi.org/10.1177/1073191113514105

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*(1), 141–151. https://doi.org/10.1177/001316446002000116

Kees, J., Berry, C., Burton, S., & Sheehan, K. (2017). An analysis of data quality: Professional panels, student subject pools, and Amazon's Mechanical Turk. *Journal of Advertising*, *46*(1), 141–155. https://doi.org/10.1080/00913367.2016.1269304

Kleka, P., & Soroko, E. (2018a). How to avoid the sins of questionnaires abridgement. *In Survey Research Methods*, *12*(2), 147–160.

Kleka, P., & Soroko, E. (2018b). How to abbreviate questionnaires and avoid the sins? *In Survey Research Methods*, *12*(2), 147–160.

Kline, P. (2013). *Handbook of psychological testing* (3rd ed.). Routledge.

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford publications.

Kost, R. G., & da Rosa, J. C. (2018). Impact of survey length and compensation on validity, reliability, and sample characteristics for ultrashort-, short-, and long-research participant perception surveys. *Journal of Clinical and Translational Science*, *2*(1), 31–37. https://doi.org/10.1017/cts.2018.18

Lachman, M. E., & Weaver, S. L. (1998). The sense of control as a moderator of social class differences in health and well-being. *Journal of Personality and Social Psychology*, *74*(3), 763–773. https://doi.org/10.1037/0022-3514.74.3.763

Marquardt, C. A., Kramer, M. D., Tellegen, A., & Arbisi, P. A. (2021). Development of the Multidimensional Personality Questionnaire-55 abbreviated research scales. *Psychological Assessment*, *33*(10), 952–961. https://doi.org/10.1037/pas0001039

Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J., Trautwein, U., & Nagengast, B. (2010). A new look at the big five factor structure through exploratory structural equation modeling. *Psychological Assessment*, *22*(3), 471–491. https://doi.org/10.1037/a0019227

Marsh, H. W., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A. J., & Trautwein, U. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling*, *16*(3), 439–476. https://doi.org/10.1080/10705510903008220

McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum.

McGue, M., Slutske, W., Taylor, J., & Iacono, W. G. (1997). Personality and substance use disorders: I. Effects of gender and alcoholism subtype. *Clinical and Experimental Research*, *21*(3), 513–520. https://doi.org/10.1097/00000374-199705000-00019

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *The Journal of Applied Psychology*, *93*(3), 568–592. https://doi.org/10.1037/0021-9010.93.3.568

Miller, M. W., Greif, J. L., & Smith, A. A. (2003). Multidimensional Personality Questionnaire profiles of veterans with traumatic combat exposure: Externalizing and internalizing subtypes. *Psychological Assessment*, *15*(2), 205–215. https://doi.org/10.1037/1040-3590.15.2.205

Milojev, P., Osborne, D., Greaves, L. M., Barlow, F. K., & Sibley, C. G. (2013). The Mini-IPIP6: Tiny yet highly stable markers of Big Six personality. *Journal of Research in Personality*, *47*(6), 936–944. https://doi.org/10.1016/j.jrp.2013.09.004

Mroczek, D. K., & Kolarz, C. M. (1998). The effect of age on positive and negative affect: A developmental perspective on happiness. *Journal of Personality and Social Psychology*, *75*(5), 1333–1349. https://doi.org/10.1037/0022-3514.75.5.1333

Muthén, B., & Muthén, L. K. (2023). *Mplus user's guide.* version 810. Muthén & Muthén. (Original work published 1998)

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: Measures of effect size for some common research designs. *Psychological Methods*, *8*(4), 434–447. https://doi.org/10.1037/1082-989X.8.4.434

Patrick, C. J., Curtin, J. J., & Tellegen, A. (2002). Development and validation of a brief form of the Multidimensional Personality Questionnaire. *Psychological Assessment*, *14*(2), 150–163. https://doi.org/10.1037/1040-3590.14.2.150

Patrick, C. J., Kramer, M. D., Tellegen, A., Verona, E., & Kaemmer, B. A. (2013). Development and preliminary validation of a simplified-wording form of the Multidimensional Personality Questionnaire. *Assessment*, *20*(4), 405–418. https://doi.org/10.1177/1073191112467051

Perera, H. N., McIlveen, P., Burton, L. J., & Corser, D. M. (2015). Beyond congruence measures for the evaluation of personality factor structure replicability: An exploratory structural equation modeling approach. *Personality and Individual Differences*, *84*, 23–29. https://doi.org/10.1016/j.paid.2015.01.004

Picconi, L., Jackson, C. J., Balsamo, M., Tommasi, M., & Saggino, A. (2018). Factor structure and measurement invariance across groups of the Italian Eysenck Personality Questionnaire-Short form (EPP-S). *Personality and Individual Differences*, *123*, 76–80. https://doi.org/10.1016/j.paid.2017.11.013

Prenda, K. M., & Lachman, M. E. (2001). Planning for the future: A life management strategy for increasing control and life satisfaction in adulthood. *Psychology and Aging*, *16*(2), 206–216. https://doi.org/10.1037/0882-7974.16.2.206

Rammstedt, B., Danner, D., Soto, C. J., & John, O. P. (2018). Validation of the short and extra-short forms of the Big Five Inventory-2 (BFI-2) and their German adaptations. *European Journal of Psychological Assessment*, *36*(1), 149–161. https://doi.org/10.1027/1015-5759/a000481

Romero, E., Villar, P., Gómez-Fraguela, J. A., & López-Romero, L. (2012). Measuring personality traits with ultra-short scales: A study of the Ten Item Personality Inventory (TIPI) in a Spanish sample. *Personality and Individual Differences*, *53*(3), 289–293. https://doi.org/10.1016/j.paid.2012.03.035

Rosenberg, M. (1965). *Society and the adolescent self-image* (Vol. 11, p. 326). Princeton University Press.

Rucker, D. D., Galinsky, A. D., & Magee, J. C. (2018). The agentic–communal model of advantage and disadvantage: How inequality produces similarities in the psychology of power, social class, gender, and race. *Advances in Experimental Social Psychology*, *58*, 71–125.

Ryff, C. D., Almeida, D. M., Ayanian, J. Z., Carr, D. S., Cleary, P. D., Coe, C., Davidson, R. J., Krueger, R. F., Lachman, M. E., Marks, N. F., Mroczek, D. K., Seeman, T. E., Seltzer, M. M., Singer, B. H., Sloan, R. P., Tun, P. A., Weinstein, M., & Williams, D. R. (2021). *Midlife in the United States (MIDUS 2), 2004-2006.* Inter-university Consortium for Political and Social Research.

Ryff, C., Almeida, D., Ayanian, J., Binkley, N., Carr, D. S., Coe, C., Davidson, R., Grzywacz, J., Karlamangla, A., Krueger, R., Lachman, M., Love, G., Mailick, M., Mroczek, D., Radler, B., Seeman, T., Sloan, R., Thomas, D., Weinstein, M.,

& Williams, D. (2019). *Midlife in the United States (MIDUS 3), 2013–2014*. Inter-university Consortium for Political and Social Research.

Ryff, C. D., & Keyes, C. L. M. (1995). The structure of psychological well-being revisited. *Journal of Personality and Social Psychology*, *69*(4), 719–727. https://doi.org/10.1037/0022-3514.69.4.719

Sellbom, M., Corey, D. M., & Ben-Porath, Y. S. (2022). Incremental validity of the multidimensional personality questionnaire in the preemployment assessment of police officer candidates. *Criminal Justice and Behavior*, *49*(7), 1050–1069. https://doi.org/10.1177/00938548211033630

Sellbom, M., & Tellegen, A. (2019). Factor analysis in psychological assessment research: Common pitfalls and recommendations. *Psychological Assessment*, *31*(12), 1428–1441. https://doi.org/10.1037/pas0000623

Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A reevaluation of the Life Orientation Test. *Journal of Personality and Social Psychology*, *67*(6), 1063–1078. https://doi.org/10.1037/0022-3514.67.6.1063

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

Somaraju, A. V., Nye, C. D., & Olenick, J. (2022). A review of measurement equivalence in organizational research: What's old, what's new, what's next? *Organizational Research Methods*, *25*(4), 741–785. https://doi.org/10.1177/10944281211056524

Spörrle, M., & Bekk, M. (2014). Meta-analytic guidelines for evaluating single-item reliabilities of personality instruments. *Assessment*, *21*(3), 272–285. https://doi.org/10.1177/1073191113498267

Stanton, J. M., Sinar, E. F., Balzer, W. K., & Smith, P. C. (2002). Issues and strategies for reducing the length of self-report scales. *Personnel Psychology*, *55*(1), 167–194. https://doi.org/10.1111/j.1744-6570.2002.tb00108.x

Tellegen, A. (1982). *Brief manual for the differential personality questionnaire* [Unpublished manuscript]. University of Minnesota.

Tellegen, A. (2011). *Multidimensional Personality Questionnaire 276 (MPQ-276), research kit*. University of Minnesota Press.

Tellegen, A., & Waller, N. G. (2008). Exploring personality through test construction: Development of the Multidimensional Personality Questionnaire. In G. J. Boyle, G. Matthews, & D. H. Saklofske (Eds.), *The SAGE handbook of personality theory and assessment: Vol. 2. Personality measurement and testing* (pp. 261–292). SAGE.

Thalmayer, A. G., Saucier, G., & Eigenhuis, A. (2011). Comparative validity of brief to medium-length Big Five and Big Six Personality Questionnaires. *Psychological Assessment*, *23*(4), 995–1009. https://doi.org/10.1037/a0024165

The Jamovi project. (2021). *Jamovi* (Version 2.3.28) [Computer software].

Thørrisen, M. M., Sadeghi, T., & Wiers-Jenssen, J. (2021). Internal consistency and structural validity of the Norwegian translation of the Ten-Item Personality Inventory. *Frontiers in Psychology*, *12*, 723852. https://doi.org/10.3389/fpsyg.2021.723852

Trapnell, P. D., & Paulhus, D. L. (2012). Agentic and communal values: Their scope and measurement. *Journal of Personality Assessment*, *94*(1), 39–52. https://doi.org/10.1080/00223891.2011.627968

Trapnell, P. D., & Wiggins, J. S. (1990). Extension of the Interpersonal Adjective Scales to include the Big Five dimensions of personality. *Journal of Personality and Social Psychology*, *59*(4), 781–790. https://doi.org/10.1037/0022-3514.59.4.781

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, *41*(3), 321–327. https://doi.org/10.1007/BF02293557

Waller, N. G., DeYoung, C. G., & Bouchard, T. J.Jr, (2016). The recaptured scale technique: A method for testing the structural robustness of personality scales. *Multivariate Behavioral Research*, *51*(4), 433–445. https://doi.org/10.1080/00273171.2016.1157753

Woods, S. A., & Hampson, S. E. (2005). Measuring the Big Five with single items using a bipolar response scale. *European Journal of Personality*, *19*(5), 373–390. https://doi.org/10.1002/per.542

Yarkoni, T. (2010). The abbreviation of personality, or how to measure 200 personality scales with 200 items. *Journal of Research in Personality*, *44*(2), 180–198. https://doi.org/10.1016/j.jrp.2010.01.002