

cognitive behavioral interventions for anxiety in diverse populations. *Behavior Therapist*, 36, 101–108.

## Multidimensional Personality Questionnaire

Nicole M. Cain,<sup>1</sup> Mark R. Lukowitsky,<sup>2</sup> and Aidan G.C. Wright<sup>3</sup>

<sup>1</sup>Long Island University, Brooklyn, U.S.A., <sup>2</sup>Albany Medical College, U.S.A., and <sup>3</sup>University of Pittsburgh, U.S.A.

The Multidimensional Personality Questionnaire (MPQ; Tellegen, 1982) is a self-report instrument that provides a comprehensive assessment of personality. The original MPQ contained 300 items, but the measure has been subsequently revised to consist of 276 mostly true-false items. A central issue in contemporary personality research centers on how personality traits as assessed by self-report inventories are linked to phenomena at other levels of analysis, such as psychological processes and biological systems. The MPQ is notable in that it provides a fine-grained analysis of personality by measuring a range of dispositions at the lower level via 11 primary trait scales. Additionally, these trait scales coalesce into four orthogonal higher order factors: Positive Emotionality (PEM), Negative Emotionality (NEM), Constraint (CON), and Absorption (ABS). PEM and NEM are explicitly linked to temperament, and CON encompasses traits related to behavioral restraint and reversed impulsivity. In this way, the MPQ is able to inform our understanding of the structure of personality; its genetic, neurobiological, and psychological underpinnings; and its relationship to psychopathology.

### Test Construction and the Development of the MPQ

The MPQ was developed through an *exploratory approach* to test construction. Tellegen and

Waller (2008) described the exploratory approach as allowing constructs to evolve and change as a planned part of the test construction process as compared with the more structured approach taken by the deductive and external methods of test construction. The exploratory approach does not lack focus or direction; rather, it is a bidirectional process whereby ideas are data and data are ideas. Initial constructs guide the development of an item pool, and data collection with this item pool then guides construct revision, followed by yet another round of data collection, and so on until a scale is developed to match the constructs.

The exploratory test construction process began with a belief that personality traits are real and important. This realist perspective conceives of personality traits as psychobiological structures underlying behavioral dispositions. Dispositions are not generalized, situation-free action tendencies but, rather, proclivities to behave in certain ways in certain situations (Tellegen, 1991). In this way, a dimensional trait occurs in essentially the same qualitative form in different people, but with quantitative variations in the amount or level. Developing a measure based on this premise requires an overinclusive item pool drawing from different conceptualizations or alternative theories of the target traits. This approach provides information on convergent and discriminant correlation patterns that will ultimately lead to better construct elaboration and demarcation. It is a laborious process, requiring several iterations of data collection and item revisions. By the end of the iterative process, the exploratory approach has ideally led to better constructs and to scales that match these constructs.

It was in this manner that the MPQ was developed over a period of 10 years. Although the MPQ was eventually developed into an omnibus personality measure, this was not the original goal of its developers. Rather, it was to clarify and highlight the nature of several trait dimensions that were repeatedly described in the personality literature but were

not well understood, were given conflicting interpretations, or were psychometrically neglected. Thus, items from various preexisting personality measures were assembled and administered to participants along with newly developed items using a series of seven gradually expanded questionnaires (Questionnaires I–VII) in order to clarify the structure and content of major personality dimensions. From each successive questionnaire, the items that were the least psychometrically sound were omitted and replaced by new items. The new items aimed to clarify meaning and to bolster measurement of the already included trait dimensions. In addition, new items were generated to identify additional significant trait dimensions that were not already being measured. The process of generating new items was conducted by Tellegen and colleagues as well as by graduate students participating in Tellegen's personality assessment seminar at the University of Minnesota. Item developers discussed the trait constructs at length but also submitted items that reflected their own interpretations of the previously identified trait markers; thus, alternative theories of the same construct were encouraged. Out of 1,082 items, only 300 items survived this iterative process to become original MPQ items (Tellegen & Waller, 2008).

Factor analytic work in the development of the MPQ employed an orthogonal (Varimax) rotation to clarify the number and location of the dimensions needed to represent the most robust item clustering. Tellegen and colleagues required that each factor dimension and each item identified as a factor marker be replicated in each round of data collection and analysis. This process was evaluated separately for both sexes in samples exceeding 250 participants each to reduce the probability of discarding psychometrically sound items. Factor replication was conducted in two steps. First, a subset of items consisting of established markers of each scale was factor analyzed. Second, the remaining items, usually recently added items, were correlated with the replicated marker scales. Items demonstrating good convergent

and discriminant correlational patterns in both sexes were ultimately retained. Tellegen and colleagues did not aim to create scales of equal length, but, rather, the number of items on each scale represented the number of items needed to achieve adequate internal reliability and how many distinct content areas had been identified in the course of construct elaboration. Also, mostly true-keyed items and mostly false-keyed items were written for each scale; however, balanced keys were not required. Items for the MPQ were solely determined psychometrically, with the key to item inclusion being its convergent and discriminant properties (Tellegen & Waller, 2008).

In addition to the primary substantive scales, the True-Response Inconsistency (TRIN) scale was developed to detect characteristic inconsistencies in response patterns to MPQ content that are the result of a fixed-response style. It comprises pairs of items that are similar in content if scored in opposite directions, so that an imbalance of true or false answers across the item pairs is indicative of inconsistency—either “yea-saying” or “nay-saying,” respectively. The MPQ also contains a Variable-Response Inconsistency (VRIN) scale consisting of item pairs in which members are similar and keyed in the same direction, so that nonmatching answers (i.e., “true” for one and “false” for the other) reflect inconsistency of response. The TRIN and VRIN scales are valuable because they address the basic issue of whether the respondent attended sufficiently to the content of the items to answer in a consistent fashion.

### **MPQ Primary Scales and Higher Order Factors**

In its original format, the MPQ contained 300 items rated on a mostly true-false scale (Tellegen, 1982). However, subsequent iterative revisions resulted in the final version of 276 binary, mostly true-false items that factor into 11 primary scales that range from 19 to 34 items each. Nine items from the primary scales comprise the TRIN and VRIN validity pairs. The MPQ

also contains a 14-item index of social desirability, the Unlikely Virtues (UNVIR) scale.

Tellegen and Waller (2008) describe the development of the MPQ scales and note that it began with an exploration of individual differences in hypnotic susceptibility, which was a construct that had not received much attention in personality assessment. Questionnaire I was administered to subjects and included items developed to represent a variety of hypnotic tendencies, including trust, ability to relax, dissociative tendencies, and the like. Research participants concurrently completed measures that reflected Neuroticism (N) and Extraversion (E). Factor analyses suggested three independent factors, the first two being N and E and a third that was labeled *Absorption* as it reflected a tendency to become immersed in self-involving and self-altering experiences that are triggered by either internal or external stimuli.

Questionnaire II was administered to further augment the N, E, and Absorption factors, and it also included items that broadly reflected Eysenck's Psychoticism, which, despite its name, is characterized by antagonism and low constraint as opposed to aberrant perception and cognition. Factor analyses led to a more clearly defined N factor, which subsequently became the MPQ Stress Reaction scale, which describes individual differences in the frequency and intensity of negative emotional states (e.g. anxiety, anger, and distress). Psychoticism items did not add variance to Absorption, nor did they suggest a separate Psychoticism scale. Thus, additional item sets were generated that eventually led to the development of the Alienation scale, which describes people who endorse the belief that they are victims of malicious intent by others. Psychoticism items were also captured by the development of a Control-versus-Impulsivity scale, which contrasts individuals who are reflective, careful, or plodding with those who are impulsive, reckless, or careless. Factor analyses of Questionnaire II also suggested that the items reflecting E could be defined by three distinct dimensions, which became

the Social Potency, Social Closeness, and Control-versus-Impulsivity scales of the current MPQ.

The purpose for Questionnaire III was to further improve upon the previously developed scales and also to administer items to further demarcate styles of thinking that were related to conforming, conventional, and authoritarian beliefs and attitudes. Through an iterative process, a new scale called Traditionalism emerged with high scores suggesting high moral standards, the endorsement of religious standards, strict child-rearing practices, and the opposition of rebelliousness. By contrast, low scorers do not give strong importance to morals, do not believe in punitive discipline, and are not very concerned with what is considered "proper."

Questionnaire IV continued to improve the previously developed scales and further demarcated constructs of interest. For example, upon further exploration of the Social Potency scale, it became clear that its content included items reflecting interpersonal effectiveness as well as a desire to make an impact on others. Questionnaire IV thus included items to clarify the relationship between dominance and achievement needs. Factor analyses suggested the emergence of a scale that was labeled Achievement, which reflected individual differences in the propensity to work hard, persist, and set high and ambitious standards. The Achievement scale was thus related to but distinct from Social Potency, which captured a tendency to be forceful, be persuasive, and enjoy taking on leadership roles where one has the opportunity to influence others. Questionnaire IV also set out to further explore and clarify interpersonal themes found in personality measures. Interpersonal theory had traditionally placed friendliness as opposite to hostility. However, the items making up the Social Closeness scale of the MPQ contrasted warmth or a preference for interpersonal closeness with distance and a preference for solitude. The question thus arose as to whether hostility should be considered a distinct dimension. Questionnaire IV thus

included a number of items reflecting aggressive qualities. Factor analyses suggested that these items were distinct from Social Closeness and suggestive of an Aggression factor, which captured a tendency toward physical aggression and vindictive qualities. Thus, the MPQ has four distinguishable interpersonal scales: Social Potency, Social Closeness, Aggression, and Alienation. Data collected from Questionnaire IV also served to further demarcate Stress Reaction from other forms of general "anxiety." The addition of motivational items pulled from various preexisting scales that reflected avoidance reactions allowed for the development of the current Harmavoidance scale, which captures an approach-avoidance motivational trait reflecting a tendency to seek or avoid physically dangerous activities (e.g., skydiving) and individual differences in fearfulness versus enjoyment of those dangerous situations. Harmavoidance is thus both empirically and conceptually distinct from the Stress Reaction scale, which is not considered a motivational-behavioral trait and rather reflects an expectation of catastrophe but not a pattern of avoidance.

Following the development of these 10 scales, it became apparent to Tellegen and colleagues that the content of the items largely reflected emotional states and dispositions. To evaluate whether the MPQ was largely an emotional-temperament inventory, the already developed scales were administered along with a mood inventory in the form of Questionnaire V. Several substantive correlations were found between the MPQ scales and mood items. For example, Stress Reaction correlated strongly with several negative mood state items, as did Alienation and Aggression. Social Potency, Achievement, and Social Closeness also correlated positively with positive mood states, albeit to a lesser degree. To make the MPQ a more complete inventory, the authors decided to include a scale reflecting the positive counterpart to the Stress Reaction scale. Thus, items were generated to reflect dispositional Wellbeing and were administered to participants as part of Questionnaire VI. This addition proved

fruitful, and thus an additional round of item development and data collection as part of Questionnaire VII led to the development of the Wellbeing scale, which represented the 11th developed scale of the MPQ.

Once the 11 MPQ scales were generated, attempts were made to further elaborate the structure of the scales and to identify higher order factors. The higher order factor analysis of the MPQ scales suggested a four-factor solution: PEM, NEM, CON, and ABS. PEM includes the scales Wellbeing, Social Potency, Achievement, and Social Closeness. Individuals who score high on PEM endorse feeling efficacious in life, involved in social and work-related activities, and open to positive experiences and emotions. NEM is associated with Stress Reaction, Alienation, and Aggression. Individuals with high scores on NEM endorse feeling stressed and tend to react with strong negative emotions to everyday difficulties. The CON factor is composed of the Control, Harmavoidance, and Traditionalism scales, and to a lesser extent with Social Potency and (reversed) Aggression. Individuals high on CON describe themselves as being cautious and conventional, avoiding danger, and adhering to traditional values. Individuals who score low on CON are characterized by impulsivity, danger seeking, and a rejection of traditional rules and behaviors. ABS is both a trait scale and a higher order factor. It includes items that reflect two subscales: sentient and prone to imaginative and altered states. Individuals who score high on ABS describe themselves as being responsive to evocative stimuli and deeply immersed in their thoughts and imagination, and as experiencing states of altered awareness.

### Validity of the MPQ

Substantial evidence exists for the relationship between MPQ personality variables and indices of overt behavior. In a sample of 228 college students, Harkness, Tellegen, and Waller (1995) examined agreement between self-report MPQ trait scores and informant ratings on the MPQ

trait scales collected from significant others (e.g., parent, friend, or romantic partner). Correlations between self-report scores and informant ratings were significant for all MPQ trait scales, with correlations averaging .43. Similarly, Kamp (1986) examined correlations between MPQ trait and higher order factor scores and specific behaviors using a biographical inventory in 224 college students. He found substantial relationships between MPQ scores and the frequency of behaviors in various domains, such as leadership (.45 correlation with Social Potency), destructive behaviors (.50 correlation with Aggression), neurotic behaviors such as frequent absences from school (.51 correlation with Stress Reaction), social activities (.61 correlation with PEM), use of alcohol and drugs (–.51 correlation with CON), and aesthetic activities (.46 correlation with Absorption).

There is also a growing literature on the relationship of MPQ traits to psychopathology and maladjustment. Tellegen (1985) noted that anxiety and depression could be differentiated based on the higher order factors of NEM and PEM. He argued that depression could be identified by low PEM, whereas anxiety could be identified by high NEM. Based on a factor analysis of mental disorders that identified two broad dimensions of psychopathology, an internalizing dimension (e.g., anxiety and mood disorder) and an externalizing dimension (e.g., alcohol and drug use, and antisocial personality disorder), Krueger, McGue, and Iacono (2001) found that the internalizing dimension was positively correlated with MPQ NEM—and, among women, negatively correlated with PEM—whereas the externalizing dimension was negatively correlated with MPQ CON.

Finally, Tellegen et al. (1988) reported significant genetic variance components for all MPQ primary trait scales (range = .39 to .55) and the higher order factors (range = .40 to .58) in samples of monozygotic and dizygotic twins reared together and apart, thus demonstrating substantial heritability for the primary traits and higher order factors. Krueger (2000)

also reported high levels of correspondence between the higher order factors of the MPQ and the structure that emerged when genetic variance components for each primary trait scale were extracted and factor analyzed. Correlations between the PEM, NEM, and CON factor scores and scores on the corresponding genetic factors were .97, .96, and .98, respectively, leading Krueger (2000) to conclude that the structure of personality as assessed by the MPQ closely corresponds with the underlying etiological structure of personality.

**SEE ALSO:** Construct Validity; Factor Analysis; Internalizing and Externalizing; Multitrait–Multimethod Analysis; Reliability; Scale Development; Self-Report Questionnaires

## References

- Harkness, A. R., Tellegen, A., & Waller, N. G. (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, 185–204.
- Kamp, J. D. (1986). *A demonstration of principles for improving the prediction of behavior from personality inventories*. Unpublished manuscript, University of Minnesota, Minneapolis.
- Krueger, R. (2000). Phenotypic, genetic, and nonshared environmental parallels in the structure of personality: A view from the Multidimensional Personality Questionnaire. *Journal of Personality and Social Psychology*, 79, 1057–1067.
- Krueger, R., McGue, M., & Iacono, W. G. (2001). The higher-order structure of common DSM mental disorders: Internalization, externalization, and their connections to personality. *Personality and Individual Differences*, 30, 1245–1259.
- Tellegen, A. (1982). *Brief manual for the Multidimensional Personality Questionnaire*. Unpublished manuscript, University of Minnesota, Minneapolis.
- Tellegen, A. (1985). Structures of mood and personality and their relevance to assessing anxiety, with an emphasis on self-report. In A. H. Tuma and J.D. Master (Eds.), *Anxiety and the anxiety disorders* (pp. 681–706). Mahwah, NJ: Erlbaum.

- Tellegen, A. (1991). Personality traits: Issues of definition, evidence, and assessment. In W. M. Grove & D. Cicchetti (Eds.), *Thinking clearly about psychology: Essays in honor of Paul Everett Meehl: Vol. 2. Personality and psychopathology* (pp. 10–35). Minneapolis: University of Minnesota Press.
- Tellegen, A., Lykken, D. T., Bouchard, T. J., Wilcox, K., Segal, N., & Rich, S. (1988). Personality similarity in twins reared apart and together. *Journal of Personality and Social Psychology*, 54, 1031–1039.
- Tellegen, A., & Waller, N. G. (2008). *Exploring personality through test construction: Development of the Multidimensional Personality Questionnaire*. Minneapolis, MN: University of Minnesota Press.

### Further Reading

- Tellegen, A. (1988). The analysis of consistency in personality assessment. *Journal of Personality*, 56, 621–663.

## Multiple Determinants and Effect Size

Michael J Strube

Washington University, U.S.A.

An effect size estimates the magnitude of a treatment's influence on an outcome or estimates the magnitude of the relationship between two variables (e.g., a measure of a trait and a measure of a behavior that the trait influences). Effect sizes are typically reported in a standardized metric, such as a correlation coefficient or a standardized mean difference, allowing easy interpretation of effect magnitude regardless of the original measurement scales or methods. The standardized metric makes it easy to compare effects across studies and to combine the effect sizes from multiple studies, as is necessary in meta-analyses. Although effect sizes are intended to be easy to interpret, their magnitude is nonetheless influenced by several factors that represent the

context within which an effect size needs to be considered.

One of the easiest ways to understand what is meant by the context of an effect size is to consider the simple correlation coefficient and the limitation that reliability of measurement places on the values that the correlation can take. A well-known equation from classical measurement theory (e.g., Nunnally & Bernstein, 1994) shows how reliability attenuates the magnitude of an obtained correlation between two measures relative to the theoretical or error-free (latent) value of that correlation:

$$r_{xy} = \rho_{xy} \sqrt{r_{xx} r_{yy}}$$

In this equation,  $r_{xy}$  is the obtained correlation between two measures (e.g., a trait measure and a behavior measure);  $\rho_{xy}$  is the theoretical, latent, or error-free correlation between the variables; and  $r_{xx}$  and  $r_{yy}$  are the reliabilities of the two measures. As is evident from the equation, increasing measurement error (i.e., lower reliability) attenuates the magnitude of the obtained correlation relative to its true underlying value. For example, assume that the underlying true correlation between the trait of neuroticism and the behavior of compulsive checking of door and window locks is .40. Assume, however, that the measure of neuroticism that is used in research has a reliability of .80 and the measure of compulsive lock checking has a reliability of .70. The magnitude of the *obtained* correlation ( $r_{xy}$ ) between the measures will be .30. The less-than-perfect reliability of the measures has limited the ability to recover the underlying true correlation. Another way to view attenuation due to measurement error is that it places more realistic boundaries on the values that a correlation can take. Unless measurement is error free, obtained correlations do not have bounds of  $\pm 1.00$  but instead have bounds of  $\pm (r_{xx} r_{yy})^{1/2}$ . Whether a correlation is considered *small* or *large*, then, depends in part on the measurement reliability context. An obtained correlation of .3 might seem quite small in the context of highly reliable measures,