

BUROS

CENTER FOR TESTING

TEST REVIEWS

Myers-Briggs Type Indicator® Step II (Form Q)

Myers-Briggs Type Indicator(r) Step II (Form Q)

Purpose

"Provides an in-depth personalized account of personality preferences."

Population

Ages 18 and over.

Publication Date

2001.

Acronym

MBTI Step II-Form Q.

Scores, 24

Four dichotomies (Extraversion vs. Introversion, Sensing vs. Intuition, Thinking vs. Feeling, Judging vs. Perceiving), 20 facets: 5 Extraversion-Introversion facets (Initiating-Receiving, Expressive-Contained, Gregarious-Intimate, Active-Reflective, Enthusiastic-Quiet), 5 Sensing-Intuition facets (Concrete-Abstract, Realistic-Imaginative, Practical-Conceptual, Experiential-Theoretical, Traditional-Original), 5 Thinking-Feeling facets (Logical-Empathetic, Reasonable-Compassionate, Questioning-Accommodating, Critical-Accepting, Tough-Tender), 5 Judging-Perceiving facets (Systematic-Casual, Planful-Open-Ended, Early Starting-Pressure-Prompted, Scheduled-Spontaneous, Methodical-Emergent).

Administration

Individual or group.

Price Data, 2001

\$85 per preview kit including item booklet, prepaid profile answer sheet, and manual (2001, 202 pages); \$97 per interpretive (Form Q) preview kit including item booklet, prepaid interpretive answer sheet, and manual; \$50 per 25 item booklets; \$150 per 10 profile answer sheets; \$110 per 5 interpretive answer sheets; \$15 per profile web administration; \$22 per interpretive web administration; \$65 per manual.

Time

(25-35) minutes.

Comments

Includes all items comprising Step I, Form M of MBTI(r); can be used to generate all reports produced by Step I, Form M of MBTI(r); provides advice for enhancing communication, conflict and change management, and decision making skills; computer or web-administered scoring available.

Authors

Katharine C. Briggs, Isabel Briggs Myers, Naomi L. Quenk (Profile Form, Interpretive Form, MBTI Step II Manual), Jean Kummerow (Profile Form, Interpretive Form), Allen L. Hammer (MBTI Step II Manual), and Mark S. Major (MBTI Step II Manual).

Publisher

Consulting Psychologists Press, Inc.

Cross References

For information about other components of this program, see T5:1755 (78 references) and T4:1702 (45 references); for reviews by John W. Fleenor and Paul M. Mastrangelo of Form M, see 14:251 (1 reference); for a review by Jerry S. Wiggins, see 10:206 (42 references); for a review by Anthony J. DeVito, see 9:739 (19 references); see also T3:1555 (42 references); for a review by Richard Coan, see 8:630 (115 references); see also T2:1294 (120 references) and P:177 (56 references); for reviews by Gerald A. Mendelsohn and Norman D. Sundberg and an excerpted review by Laurence Siegel, see 6:147 (10 references).

Review of the Myers-Briggs Type Indicator(r) Step II (Form Q) by ALLEN K. HESS, Professor, Auburn University at Montgomery, Montgomery, AL:

DESCRIPTION

Form Q Step II is the most recent iteration of the Myers-Briggs type Indicator (MBTI). Step II develops the facet approach; uses IRT (item response theory) in selecting items; presents a national normative sample; uses a graphical array to show profile results; and reports results via a four-letter personality typing system.

Recently, personality inventories that take a "basic personality" type or trait approach (e.g., the "Big Five" of the NEO-PI-R) have shredded the basic personality traits or types into components. The Form Q Step II uses the term "facets" to depict these components in the MBTI. Two thinking types might have different item and facet endorsements so we might see one as more logical, reasonable, and questioning, whereas the other is more logical, questioning, and critical. The facet scores allow a sharper definition of personality.

DEVELOPMENT

The Form Q Step II uses 93 items from Form M plus 51 more items to compose the total 144-item pool. This allows for the test user to import much of the interpretive wisdom gathered from prior research in using the current Step II edition. Also, it appears that the Step II items are scored on one and only one scale, as opposed to some inventories that use the same item on several scales. The latter compromises any statistical analyses and aborts differential assessment (i.e., makes it hard to distinguish between types of people). Using the item on one and only one scale is excellent and does not inflate scale-to-scale correlations, as is the case with some other inventories.

TECHNICAL

Sample

A national sample of 1,380 people is a marked advance for the MBTI. Still, the sample is weighted toward the older, the female, and towards Caucasians. When interpretations are made on the basis of

the 16 combinations of the types (e.g., ESFP or Extraverted, Sensing, Feeling, and Perceiving person), the normative sample is decomposed. Thus, the interpretation has a smaller sample base if one considers how many of the 1,380 are ESFP types. Future work establishing a sounder normative sample to use in typical interpretation is needed.

Reliability

Internal consistency, or the degree to which items within a scale adhere to each other, is good, particularly given the small number of items (from 5 to 9) on each facet. Alpha coefficients range from .57 to .85 with a median of .77. Test-retest reliability, or the degree to which scores are stable from time to time, is good. The manual presents a more telling stability measure. Because the MBTI measures types rather than traits, the degree to which people move from one category to another or remain in the same category was computed. The MBTI demonstrates a comforting degree of stability of classification. Finally, the manual presents a Polarity index or a measure of the degree to which people endorse extremes or central positions on the items, a useful feature in interpreting individual profiles.

Validity

Validity concerns the degree to which interpretations drawn from the test are empirically supported. The manual describes the facet scales as correlating with each other (convergent validity) and not correlating with facet scales (discriminant validity) from other dimensions (e.g., Expressive facet with other Extraversion facet scales and not with Thinking scales). Research on factorial purity of the preference scales (e.g., Extraversion-Introversion) is not presented.

The bulk of the validation studies concerns how the MBTI facet and preference scales correlate with several other personality measures: the California Psychological Inventory (CPI), Fundamental Interpersonal Relations Orientation-Behavior (FIRO-B), Eysenck Personality Questionnaire (EPQ), and the Rotter Locus of Control Scale. Also, research based on observer's descriptions of test takers was correlated with MBTI data. Most of the correlations are in the expected direction, providing some evidence of concurrent validity. One problem with this kind of data is that results in the other direction often can be explained away too readily. The research reported above was either unpublished or published in only one journal, a journal that is not in the mainstream of personality and assessment venues. Thus, predictive and experimental research replicated by various researchers and published in several of the more recognizable journals should provide the type of validation data comforting to test consumers. Finally, the manual asserts that test-taking set is part of personality. No one approaches a test with a test-neutral attitude. Whether for employment, parole, child custody, or, more prosaically, for self-awareness, the test-taker has various attitudes he or she brings to the test items. No validity scales

(e.g., social desirability scales) are included, which does not let the test interpreter know the test-taking set of the person being assessed.

COMMENTARY

The MBTI was originally developed by the mother-daughter team of Katherine Cook Briggs and Isabel B. Myers to measure Jungian types so people could know themselves. As such, the Form Q Step II manual stresses that the MBTI measures types (that takes a categorical or more bipolar approach), as opposed to traits (that take dimensional form). This means that a profile should be interpreted holistically rather than from a discrete trait approach. Simply put, a higher score on trait measures of Introversion means a person will show more introversive traits and behaviors. In the type approach, a higher score on Introversion (assuming it as the dominant function or most energized and accessible part of a personality) can only be interpreted in light of the auxiliary function (which balances the dominant function), the tertiary function (opposite to the auxiliary function), and the inferior function (opposite to dominant function and least accessible and least known to the person). For example, a person who is extraverted would manifest this dominant function quite differently if he or she has a thinking versus an intuiting auxiliary function. To be fair, any adept trait-oriented test interpreter would take various intensities of traits into account. But the type approach, and the MBTI in particular, builds this holistic interpretation into the fibers of the process. Quite clearly, the person intending to use the MBTI must be immersed in the Jungian perspective to best use the MBTI. The manual provides a good beginning toward orienting the test user in enough Jungian theory to use the MBTI.

SUMMARY

The MBTI has been useful in counseling and employment settings for a half-century. Much "heuristic" validity accumulates when professionals use a test. As test validity depends on the interpretations drawn from a test, the well-written manual provides ample interpretive wisdom. As the test has been used as a stimulus for dialog in counseling and employment, the MBTI might find a place in such collaborative clinical work as Finn and Tonsager (2002) and Fischer's (2000) approach wherein they openly discuss test findings with clients, blending the assessment and therapeutic functions of the clinician. On the other hand, use of the test in making clinical, employment, or forensic decisions without further validation research, particularly of the predictive kind, might be perilous.

REVIEWER'S REFERENCES

Finn, S. E., & Tonsager, M. E. (2002). How therapeutic assessment became humanistic. *Humanistic Psychologist*, 30, 10-22.

Fischer, C. T. (2000). Collaborative individualized assessment. *Journal of Personality Assessment*, 74, 2-14.

Cite this review

Hess, A. K. (2003). [Test review of Myers-Briggs Type Indicator(r) Step II (Form Q)]. In B. S. Plake, J. C. Impara, & R. A. Spies (Eds.), *The fifteenth mental measurements yearbook*. Retrieved from <http://marketplace.unl.edu/buros/>

REVIEW 2 OF 2

Review of the Myers-Briggs Type Indicator(r) Step II (Form Q) by KEVIN LANNING, Associate Professor of Psychology, Wilkes Honors College of Florida Atlantic University, Jupiter, FL:

The Myers Briggs Type Indicator(r) Step II (Form Q) (MBTI Step II-Form Q) represents the latest iteration in the continuing development of the MBTI. The MBTI includes scales that measure the dimensions Extraversion-Introversion (E-I), Sensing-Intuition (S-N), Thinking-Feeling (T-F), and Judging-Perceiving (J-P). These measures are typically treated dichotomously, giving rise to 16 types denoted by acronyms (e.g., ESTJ). The four dimensions and resultant type are scored on all three versions of the MBTI currently available from the test publisher, the 126-item Form G, the 93-item Form M (which supplanted Form G as the standard version of the test in 1998), and the 144-item Form Q presently under review. This review includes a brief discussion of the MBTI typology as well as comments specific to Step II-Form Q.

As a Step II measure of the MBTI, Form Q is the product of a lineage that runs largely parallel to that of the more widely used MBTI Forms G and M. Though Form Q is identical to Form M in the scoring of type, Form Q includes additional items derived from the 1989 Form K and its associated MBTI Expanded Analysis Report which, in 1996, was renamed the Expanded Interpretive Report. The additional items allow the scoring of five subscales ("facets") for each of the four dichotomies, which are reported to users in a brief Step II Profile or a more lengthy Step II Interpretive Report. In principle, the inclusion of the Step II facet scales permits more nuanced interpretation than would be afforded by describing persons in terms of type alone.

The facet scales vary in breadth: For each of the four dichotomies, there is a core facet as well as additional, narrower facets. For Extraversion-Introversion, the core facet is Initiating-Receiving. For the functions of S-N and T-F, the core facets are Concrete-Abstract and Logical-Empathetic, respectively,

whereas for J-P, the core is Systematic-Casual. For purposes of interpretation, scores are reduced to three categories: "in-preference" (i.e., in the same direction as the broader dichotomy of which the facet is a part), "midzone," and "out-of-preference." Because in-preference scores are expected given the structure of the instrument, it is the out-of-preference scores that are the main source of interest. These latter scores are used as a supplement to the four-letter type to provide an individualized type description, as, for example, in "Enthusiastic, Questioning, Casual INFJ." In addition to the 20 facets, a largely atheoretical Polarity Index measures the omnibus extremity of the facet scales, or the tendency to respond to each scale in a relatively homogeneous way.

DEVELOPMENT

The facet scales were derived from the administration of items from several prior forms of the test to a national sample of 1,380 adults recruited through random-digit dialing. Although the sample was intended to reflect U.S. Census demographics, the average age of respondents is approximately 50. An initial version of the facets was developed based on confirmatory factor analyses of Form K and other sets of items. From these analyses, items were retained for Form Q based on several criteria, including conceptual fit, the item response theory (IRT) discrimination parameter, and contribution to the internal consistency of the scale. Correlations between the Form Q measures and corresponding earlier Form K measures suggest some substantial changes from the prior measure; these range from .57 to .95 with a median of .82.

TECHNICAL

All 20 of the facets are brief, ranging from five to nine items in length, but are presented to users on an 11-point scale. Because these scores are computed using IRT, no hand-scoring is available. The manual includes descriptive statistics for the facets that are broken down by gender and age. The authors also report the kurtosis of the distribution of facet scores, but are incorrect in their assertion that these values, with a range of .09 to -1.41, support the supposed bimodality of the scales.

The correlational structure of Form Q appears largely appropriate to its intended form. In a confirmatory factor analysis, an adjusted goodness-of-fit index (AGFI) of .95 was found for the four factor by five facet model. Correlations among facet scales associated with different MBTI dichotomies are typically modest (most $< |.20|$), whereas those between facets within a dichotomy are typically higher (most $> .5$). For all but one of the facet scales, correlations with the corresponding higher order dichotomy range between .5 and .9; the remaining facet (Questioning-Accommodating) is also the only facet scale that does not share items with its broader dichotomy (Thinking-Feeling). The internal consistencies for the facets, with a median of .77, are generally high given the short length of these measures.

In addition to reliability data for the continuous scales, the manual also provides some information regarding the stability of the tripartite categories of in-preference, midzone, or out-of-preference. Across the 20 facets, the median percentage of adults remaining in the same category over a 1-month span was 73%, with only 3% moving across the two extremes.

Because the Form Q facet scales are new, evidence for their validity remains limited. The manual includes correlations with two other self-report measures, the California Psychological Inventory (CPI) and the Adjective Check List (ACL). These correlations are typically in the expected direction, particularly for the facets of Extraversion-Introversion, but do not yet support a differentiated understanding of the meaning of the facets. Similarly, hierarchical regression analyses were undertaken to assess the incremental contribution of the facets beyond the basic MBTI dimensions in the prediction of consequential outcomes such as reported hypertension and punctuality. These results indicate that the facets, when considered together, contribute to the validity of the test, but they do not speak to the meaning of the individual scales. Finally, the manual also presents correlations between the prior Form K facet scales and observer ratings and additional self-report measures; however, given the apparent differences between the Form K and Q facet scales, these are of uncertain relevance to the present instrument.

COMMENTARY

Reviews of previous versions of the MBTI have taken issue with the position that it is a measure of types rather than a measure of traits (see in particular the reviews by Mendelsohn, 1965; Mastrangelo, 2001; & Wiggins, 1989). Because the manual and interpretive materials continue to stress this point, it is appropriate to briefly consider its conceptual and empirical support.

Evidence supporting a typological interpretation of the MBTI would include, but is not limited to, both multidimensional clustering and interactions between type variables in the predictions of outcomes. The four-dimensional space defined by the E-I, S-N, T-F, and J-P axes would ideally be populated by 16 clusters, each discrete and distant from all four axes, with interactions reflected in different patterns of correlations in different clusters (such that, for example, the meaning of Thinking would effectively be different for Extraverts and Introverts). This is a very strong conception of type, but one that is suggested by the manual in the statement that "scores are expected to be bimodal-few scores at the midpoint," "the numerical portion of MBTI results has no ... diagnostic meaning," and "the four type preferences interact dynamically to form a whole that is different from the sum of the parts" (manual, p. 17).

To those raised on the Central Limit Theorem and the belief that characteristics such as extraversion are the product of polygenes and immeasurable experiences, the a priori likelihood that the distribution of MBTI scores would cleave naturally into 16 roughly equal clusters is certainly small. Unfortunately, the

manual does not provide empirical evidence which would reduce skepticism on this point. Empirical support for bimodality in the individual scales is lacking (Bess & Harvey, 2002), as is solid evidence favoring a configural interpretation of the instrument (the MMY review by Mastrangelo, 2001; cf. Reynierese & Harker, 2001). But although there is little support for treating the MBTI types as real entities, there is value in using the types as descriptive labels. The MBTI is now considered the most popular test of personality in history. This popularity stems not only from the attractiveness of Jungian concepts, but also from the appeal and cognitive economy of categories over dimensions. "ENFP" is easier to understand than, for example, t-scores of 65 on Extraversion, 35 on Sensing, 32 on Thinking, and 28 on Judging, and this probably accounts for much of the continuing popularity of the instrument. The value of the MBTI typology to consumers cannot be dismissed, despite the fact that it is the product of arbitrary cutting points rather than Aristotelian essences.

There are costs associated with treating the MBTI measures as dichotomies rather than as continua. For example, the reliabilities of continuous scales overestimate those of corresponding dichotomies by a factor of approximately 1.25 (see MacCallum, Zhang, Preacher, & Rucker, 2002). This problem is ameliorated somewhat by the fact that the interpretive materials accompanying the MBTI encourage recognition of the fallibility of scores on the instrument. The various profile sheets for the MBTI include bar charts and tables that plainly show the extremity (or "preference clarity") of scores. On the Step II Profile Sheet (though not on the Step II Interpretive Report), users are explicitly told that "The longer the bar ... the more likely it is that the instrument has accurately reflected your preference." On the Report (though not on the Profile Sheet), users are given the option of simply entering the client's "verified type" rather than deriving it from responses to the test.

In the MBTI, the categorization of persons into types is inevitably imperfect. Because most individuals will score near the mean on at least one of the major dimensions of the test, type classification may be unstable, with some one-third of all individuals changing type over a span of 1 month (see the MMY reviews by Fleenor, 2001, and Mastrangelo, 2001).

Though inclusion of the facet scales addresses this problem, it does not solve it. For example, consider a woman whose classification as an ESTJ masks only a slight preference for Sensing over Intuition. An examination of her facet scores might clarify the meaning of the moderate S-N score, reflecting, perhaps, two facets in the Sensing direction (Realistic and Practical), two in the Intuitive direction (Theoretical and Original), and one in the midzone. A problem arises when this pattern is reduced to a summary description. Because her preference for Sensing over Intuition is small, different responses to a mere one or two items could lead her to be classified as an ENTJ rather than an ESTJ. The facet scales that were previously out-of-preference would then become in-preference, and so she would change from being classified as a "Theoretical, Original ESTJ" to an "Realistic, Practical ENTJ."

Recognizing the problem that arises when facets are interpreted in light of type, the manual stresses the importance of verifying a client's type prior to providing a substantive interpretation of facet scores. But the notion of "verification" presupposes that the MBTI types really exist in nature, that there are underlying bifurcations on the major dimensions of the test. The evidence does not, again, support this claim; the world includes ambiverts as well as extraverts and introverts. Unfortunately, the midzone scores that are recognized for the facets are not applied to the higher order measures of the MBTI.

SUMMARY

There is a trade-off between simplicity and accuracy in the description of any person, between comprehensibility and completeness. The standard version of the MBTI has achieved popular success by positioning itself near the first of these poles. MBTI Step II represents a step towards the second pole, towards a more thorough if less simple description of personality. Although there is much that is attractive about this step, time will tell whether this new extension of the test retains the popular success of the prior version, and whether its narrow bandwidth facet scales warrant discrete interpretation.

REVIEWER'S REFERENCES

- Bess, T. L., & Harvey, R. J. (2002). Bimodal score distributions and the Myers-Briggs Type Indicator: Fact or artifact? *Journal of Personality Assessment*, 78, 176-186.
- Fleenor, J. W. (2001). [Review of the Myers-Briggs Type Indicator, Form M]. In B. S. Plake & J. C. Impara (Eds.), *The fourteenth mental measurements yearbook* (pp. 816-818). Lincoln, NE: Buros Institute of Mental Measurements.
- MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, 7, 19-40.
- Mastrangelo, P. M. (2001). [Review of the Myers-Briggs Type Indicator, Form M]. In B. S. Plake & J. C. Impara (Eds.), *The fourteenth mental measurements yearbook* (pp. 818-819). Lincoln, NE: Buros Institute of Mental Measurements.
- Mendelsohn, G. A. (1965). [Review of the Myers-Briggs Type indicator]. In O. K. Buros (Ed.), *The eighth mental measurements yearbook* (pp. 321-322). Highland Park, NJ: Gryphon Press.
- Reynierese, J. H., & Harker, J. B. (2001). The interactive and additive nature of psychological type. *Journal of Psychological Type*, 58, 6-32.
- Wiggins, J. S. (1989). [Review of the Myers-Briggs Type Indicator]. In J. C. Conoley & J. J. Kramer

(Eds.), *The tenth mental measurements yearbook* (pp. 537-538). Lincoln, NE: Buros Institute of Mental Measurements.

Cite this review

Lanning, K. (2003). [Test review of Myers-Briggs Type Indicator(r) Step II (Form Q)]. In B. S. Plake, J. C. Impara, & R. A. Spies (Eds.), *The fifteenth mental measurements yearbook*. Retrieved from <http://marketplace.unl.edu/buros/>