EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT
1986, 46

CONSTRUCT VALIDITY OF THE MYERS-BRIGGS TYPE INDICATOR

BRUCE THOMPSON AND GLORIA M. BORRELLO
University of New Orleans

The Myers-Briggs Type Indicator (MBTI) is widely recognized as an important measure of normal or non-pathological variations in personality. However, the construct validity of the measure has not been clearly established using factor analytic techniques. The present study investigated the structure and item performance of the instrument using data from 359 college students. Factor analysis was applied to the 95 scored MBTI items. Factor adequacy and invariance coefficients were computed, and the appropriateness of the recommended item weights was examined. The results strongly supported the instrument's construct validity.

PSYCHOLOGICAL measures that emphasize variations in normal behavior rather than psychopathology have been increasingly emphasized in research and in educational practice. Form F of the Myers-Briggs Type Indicator (MBTI) (Briggs and Myers, 1976) is representative of these types of measures. The MBTI has particular appeal since the test has both instructional (McCaulley, 1981, p. 295) and counseling applications (Myers, 1962, pp. 4–5, 76–82).

The MBTI is also noteworthy since the measure is grounded in a recognized personality theory, Jung's (1971/1921) theory of psychological types. The Extraversion-Introversion (EI) scale of the MBTI measures characteristics that Jung considered fundamental to personality—preferences for the outer world of people as opposed to the inner world of ideas. The MBTI also includes two scales measuring what Jung regarded as orienting functions. The Thinking-Feeling (TF) scale measures a preference to be guided by logic rather than by needs for affiliation and warmth. The Sensing-

Intuition (SN) scale measures a preference for orienting toward the observable as opposed to orienting via insight. The final scale, Judging-Perceiving (JP), measures a preference for order and rules as opposed to a preference for flexibility.

Interest in the MBTI has stimulated a number of reliability and validity studies that have been summarized by others (Carlyn, 1977; Carskadon, 1979). In various studies internal stability reliability coefficients for the four scales have tended to vary between .80 and .90, while stability coefficients have tended to be slightly lower. However, as McCaulley (1981, p. 319) notes, "construct validity studies are most relevant in establishing the validity of the MBTI, since the Indicator was constructed specifically to implement a theory." A variety of these constuct validity studies have been conducted and have involved diverse phenomena including short-term memory (Carlson and Levy, 1973) and clinical orientations of psychotherapists (Levin, 1979). Myers and Davis (1964) and later McCaulley (1977) conducted an impressive longitudinal study of the specialties selected by several thousand medical students over several decades.

Nunnally (1967, p. 100) notes that factor analysis can shed considerable light on validity issues and that some researchers even refer to construct validity as "factorial validity." Thus it is surprising that almost no validity studies involving the MBTI have used these methods. Thompson and Borrello (1986) report a study of MBTI performance using second-order factor-analytic methods. As Kerlinger (1984, p. xivv) noted, "while ordinarily factor analysis is probably well understood, second-order factor analysis, a vitally important part of the analysis, seems not to be widely known and understood." Basically, second-order methods involve the extraction and oblique rotation of "first-order" factors, followed by the extraction and rotation of "second-order" factors from the matrix of intercorrelations among the first-order factors.

Thompson and Borrello (1986) extracted 32 first-order factors based on Guttman's eigenvalues-greater-than-one criterion. Four second-order factors were then extracted from the interfactor correlation matrix. Subjective interpretation and factor adequacy coefficients (Thompson and Pitts, 1981/82) indicated that the MBTI measures four dimensions and that items generally measure the scales the items are expected to measure.

Although the results of the analysis were highly supportive of the MBTI's construct validity, as Gorsuch (1983) has noted, "there is nothing sacred about either primary or higher-order factors" (p. 254)

and in some cases "interest would be in both primary and secondary factors" (p. 255). As Gorsuch (1983, p. 240) emphasizes:

> The essential difference between the primary factors and the higher-order factors is that the primary factors are concerned with narrow areas of generalization where the accuracy is great. The higher-order factors reduce accuracy for an increase in the breadth of generalization. In some analyses, the reduction in accuracy when going from primary to second-order factors will be small; in other studies it may be quite great. It depends upon the data being analyzed.

The present paper reports a reanalysis of the Thompson and Borrello (1986) data to address two research questions. First, how robust is the factor structure of the MBTI over first-order as against second-order extraction methods? Reanalysis of the same data set might yield evidence that the "breadth" and the "accuracy" of structures at both levels were similar; such evidence would represent potent additional support regarding the validity of the MBTI.

In typical usage the MBTI is employed to classify subjects into types based on each of the four MBTI scales by consulting weighted scores on responses to selected items. The validity of this classification procedure to some extent presumes the validity of the recommended scoring weights. Thus, the study's second research question asked, how closely do scores derived using conventional scoring correspond with scores computed using the empirically-derived factor structures to define item weights?

Method

Subjects

All the subjects (n = 359) were students enrolled in an urban university in the southern United States. The students were enrolled in a health course, "Personal and Community Health," required for several majors and frequently chosen as an elective by other students. The sample included 103 (28.7%) non-white students and 65 males (18.1%).

Results

Form F of the MBTI consists of 166 items. However, 71 items were added as part of a research program (McCaulley, 1981, p. 312) and are not scored. Continuous scale scores were computed in the

recommended manner using the remaining 95 items (McCaulley, 1981, p. 314; Myers, 1962, p. 9).

Determining the number of factors to extract from the correlation matrix is a fundamental issue in any analysis. An increasingly common practice is to terminate extraction when the eigenvalues for successive factors begin to "level off" or begin explaining similar proportions of variance in the correlation matrix. The eigenvalues for the first seven factors were: 7.78, 5.06, 4.37, 3.65, 2.12, 2.08, and 1.98. Therefore, based on application of Cattell's "scree test," four principal components were extracted and rotated to the varimax criterion. Table 1 presents factor pattern coefficients for those items with factor pattern coefficients greater than .40 in absolute value. The table also indicates the scales that the items are intended to measure.

How closely obtained factors correspond with theoretically expected factors can be empirically evaluated by computing factor validity or "adequacy" coefficients. As Thompson and Pitts (1981/82, p. 101) explain:

> The calculated factors can be rotated to a position of "best fit" with a theoretically derived target matrix. The target matrix delineates how many factors are expected and the expected correlation between each item and each factor. The cosines of the angles between the hypothetical and the actual measures can be interpreted as validity [or correlation] coefficients.

The four factor adequacy coefficients were each greater than .95. The mean cosine (.88, SD=.11) between actual and expected results for the 95 items was also quite high.

However, as Nunnally (1967, p. 280) notes, "one tends to take advantage of chance in any situation where something is optimized from the data at hand." This occurs in correlational methods, including multiple regression, canonical correlation analysis, and factor analysis. Thus, methods to estimate the sampling specificity or invariance of analyses have been recommended by various methodologists.

In order to investigate the invariance of the factor analytic results, the sample was randomly split into roughly equal-sized subgroups ($n1$=174; $n2$=185) and then separate factor analyses were conducted with both data sets. When these two factor solutions are rotated to positions of "best fit" with each other, the cosines of the angles between the factors are correlation or invariance coefficient. All four factor invariance coefficients were greater than .95. The mean item cosine (.87; SD=.16) was also quite satisfactory.

Finally, an analysis was conducted to evaluate the scoring proce-

TABLE 1
Rotated Pattern Coefficients for Salient Items

Item	Scale	I	II	III	IV
85	JP	.67	−.05	.08	.17
27	JP	.57	.08	.05	.05
151	JP	.56	.04	.05	−.02
60	JP	.56	.08	−.04	−.03
55	JP	.53	.03	.06	.11
49	JP	.52	−.01	.02	.00
1	JP	.52	.02	.11	.02
132	JP	.52	.04	−.05	−.04
35	JP	.50	−.04	−.18	.15
42	JP	.46	−.09	.14	−.12
68	JP	.43	−.04	.06	−.07
94	JP	.43	−.12	.24	.21
20	JP	.43	−.06	.20	−.12
13	JP	.42	−.05	−.08	.06
113	JP	.40	−.16	.11	−.11
50	EI	−.06	.76	−.09	.07
87	EI	−.09	.59	−.04	−.14
148	EI	−.03	.58	−.04	.04
33	EI	.11	.58	.14	−.13
6	EI	.01	.53	.00	.14
19	EI	−.07	.50	.06	.02
15	EI	.14	.50	−.15	−.03
92	EI	−.12	.49	−.04	−.04
134	EI	−.02	.46	.08	.11
26	EI	.06	.46	−.10	.02
95	EI	−.12	.46	.12	.07
138	EI	.09	.46	−.20	−.09
25	EI	.03	.45	−.16	−.01
77	EI	.00	.43	.24	−.11
128	SN	.14	−.06	.56	.00
64	SN	−.11	.02	.50	−.08
17	SN	.04	.01	.49	.04
145	SN	.07	.03	.45	−.01
2	SN	.22	−.03	.44	.23
88	SN	.09	.10	.44	.02
104	SN	.27	−.06	.44	.15
11	SN	.01	−.07	.43	−.19
107	SN	−.08	−.09	.43	.11
76	SN	.13	.01	.42	.01
78	SN	.01	−.08	.42	−.12
70	SN	−.04	.09	.41	.12
98	SN	.24	−.06	.40	.24
114	TF	.13	−.13	.08	.66
72	TF	.01	−.01	.09	.54
86	TF	−.02	−.02	.21	.49
26	TF	.08	−.11	.01	.48
103	TF	−.06	.02	−.11	.47
89	TF	−.16	−.05	−.02	.45
79	TF	.06	.09	−.08	.45
154	TF	.10	−.11	.09	.43
91	TF	−.06	.02	.01	.42
118	JP	.30	−.11	.06	.40

dures used in some MBTI applications. Computing continuous scores in the recommended manner involves differentially weighting selected responses and then summing item scores (McCaulley, 1981, pp. 310–311). In order to evaluate this scoring procedure factor scores were computed for the 359 subjects on each of the four factors. Correlation coefficients were then computed between the factor scores and the weighted, summated scale scores. Table 2 reports the convergent and divergent validity coefficients, which in effect evaluate the MBTI's conventional weighting and scoring procedure against the scores computed using the empirically-derived factor pattern coefficients as item weights.

Discussion

It is particularly important to evaluate the construct validity of measures for which specific applications of findings have been recommended as regards both instruction and counseling (Myers, 1962, pp. 76–82). The present study provides consistent supportive evidence regarding the construct validity of the Myers-Briggs Type Indicator (Briggs & Myers, 1976).

The factor analytic results summarized in Table 1 suggest that items are related to factors in the expected fashion. Twenty-two of the 24 Judging-Perceptive (JP) items had a correlation with factor I greater than .30 in absolute value. Twenty of the 22 Extraversion-Introversion items had a correlation with factor II greater than .30 in absolute value. Twenty-two of the 26 Sensing-Intuition (SI) items had a correlation with factor III greater than .30 in absolute value. Sixteen of the 23 Thinking-Feeling (TF) items had a correlation with factor IV greater than .30 in absolute value. Furthermore, only

TABLE 2
Convergent and Divergent Validity Coefficients

Scale Score	Factor I	II	III	IV
Judging	.92*	−.06	.13	.20
Perceptive	−.94*	.06	−.14	−.18
Extraversion	−.03	.96*	−.06	−.02
Introversion	.03	−.06*	.06	−.01
Sensing	.17	−.08	.93*	.18
Intuition	−.24	.05	−.85*	−.14
Thinking	−.03	−.06	.10	.94*
Feeling	−.06	−.01	−.04	−.84*

Note. Convergent validity coefficients are indicated with asterisks. For one "end" of each of the four scale continua, one set of scale scores is expected to be negatively related with its corresponding factor.

factor IV had coefficients greater than .30 in absolute value for any items which were not expected to "load" on the scale; three non-TF items were associated with the factor.

The clarity of these results made the empirical evaluation using the factor adequacy coefficients less surprising. The correspondence between actual and expected results is clear. The invariance coefficients suggest that the results are not sample specific. The analysis presented in Table 2 suggests that the conventional scoring procedures are appropriate.

These results are particularly impressive since the results were so robust over both first-order and second-order factor analytic methods. The use of both methods with a single data set allowed a determination that the structure of the MBTI is *both* "generalizable" and "accurate" in the senses the terms have been used by Gorsuch (1983). Thus, the results indicate that the various recommended applications of MBTI findings at least will be based on data that have measurement validity.

REFERENCES

Briggs, K. C. and Myers, I. B. (1976). *Myers-Briggs Type Indicator: Form F.* Palo Alto, CA: Consulting Psychologists Press.

Carlson, R. and Levy, N. (1973). Studies of Jungian typology: Memory, social perception, and social action. *Journal of Personality*, 41, 559–576.

Carlyn, M. (1977). An assessment of the Myers-Briggs Type Indicator. *Journal of Personality Assessment*, 41, 461–473.

Carskadon, T. G. (1979). Clinical and counseling aspects of the Myers-Briggs Type Indicator: A research review. *Research in Psychological Type*, 1, 2–31.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Jung, C. G. (1971). Psychological types. In R. F. C. Hull (Ed.), *The collected works of C. G. Jung* (Vol. 6). Princeton, NJ: Princeton University Press. (Originally published 1921).

Kerlinger, F. N. (1984). *Liberalism and conservatism: The nature and structure of social attitudes.* Hillsdale, NJ: Erlbaum.

Levin, L. S. (1979). Jungian personality variables in psychotherapists of five different theoretical orientations. *Dissertation Abstracts International*, 39, 4042B–4043B. (University Microfilms 79-01,823).

McCaulley, M. H. (1977). *The Myers longitudinal medical study.* Gainesville, FL: Center for Applications of Psychological Type.

McCaulley, M. H. (1981). Jung's theory of psychological types and the Myers-Briggs Type Indicator. In P. McReynolds (Ed.), *Advances in personality assessment* (Vol. 5). San Francisco: Jossey-Bass.

Myers, I. B. (1962). *The Myers-Briggs Type Indicator: Manual*. Palo Alto, CA: Consulting Psychologists Press.

Myers, I. B. and Davis, J. A. (1964, September). *Relation of medical students' psychological type to their specialties twelve years later*. Paper presented at the annual meeting of the American Psychological Association, Los Angeles.

Nunnally, J. C. (1967). *Psychometric theory*. New York: McGraw-Hill.

Thompson, B. and Borrello, G. M. (1986). Second-order factor structure of the MBTI: A construct validity assessment. *Measurement and Evaluation in Counseling and Development*, 18, 148–153.

Thompson, B. and Pitts, M. C. (1981/82). The use of factor adequacy coefficients. *Journal of Experimental Education*, 50, 101–104.