

ITEM ANALYSIS OF THE MYERS-BRIGGS TYPE INDICATOR

GARY J. SIPPS, RALPH A. ALEXANDER, AND LARRY FRIEDT
The University of Akron

The current study explores the item structure of the Myers-Briggs Type Indicator (MBTI). A factor analysis of responses of a large sample ($n = 1291$) yielded six salient factors, four resembling the four scales of the MBTI. Kuder-Richardson 20 coefficients are reported for the scales and factors. Pearson r s are also reported for the factor-scale relationships. All analyses yield only limited support for the item validity of the MBTI. Relevant issues involving the construction of the test are discussed and suggestions are made for future research.

IN 1977 Carlyn reviewed the then current psychometric status of the Myers-Briggs Type Indicator (MBTI; Myers, 1962). From among the more than 300 studies (Buros, 1978) conducted using the MBTI, Carlyn (1977) presents those studies investigating the reliability or validity of the instrument.

Based on the available studies of construct validity, Carlyn (1977) concludes "that the individual scales of the Myers-Briggs Type Indicator measure important dimensions of personality which seem to be quite similar to those postulated by Jung" (p. 471). The relevant studies involved correlating subjects' scores on one or more of the four MBTI scales with their scores on measures of various other constructs. Representative characteristics include intellectualism and impulsiveness (Webb, 1964), vocational interests

The authors thank Justine Paulchel for her assistance in organizing the data and the results.

Requests for reprints should be sent to: Dr. Gary J. Sippes, Department of Psychology, The University of Akron, Akron, OH 44325.

Copyright © 1985 Educational and Psychological Measurement

(Myers, 1962; Stricker and Ross, 1964), and "behavioral styles" (Cohen, Cohen, and Cross, 1981). A few studies have used factor analysis to explore the nature of the measured constructs. However, with one exception, these analyses have employed scale scores of the MBTI and other instruments, and not individual items. Tzeng, Outcalt, Boyer, Ware, and Landis (1984) conducted an item analysis of the MBTI which produced "clear simple structures" yielding factors which "matched almost perfectly with the theoretical scales of the MBTI." The authors conducted three other item analyses.

Studies involving content validity have all focused on the four scales' scores in relation to a number of variables, including expert self ratings (Bradway, 1964), interview ratings (Harrison, 1976), and concurrent measures of the same typology (Stricker and Ross, 1964). Carlyn (1977) notes that "the question has not been settled" in regard to the appropriateness of the scored items in assessing Jung's typology.

Omizo (1978) suggests that the psychometric properties of the MBTI should indeed be evaluated. After a critical analysis of the instrument and the pertinent literature, he makes the recommendation, among many, that factor analytic procedures be used to establish the construct validity of the instrument.

The purpose of this study is to perform an item analysis of the MBTI using factor analytic procedures. Through the use of individual items in the analysis, stronger statements can be made regarding the psychometric properties of the MBTI than could be made using solely scale scores. Results of this examination were integrated to provide an empirical basis for determining the degree of convergent validity present in the construction of the MBTI. This was accomplished by examining both the number of factors produced and the degree to which the items from each scale loaded on one and only one common factor. In addition, KR-20 reliability coefficients were calculated to further explore the psychometric properties of the instrument.

Method

Subjects

The sample for this study consisted of 1291 subjects who were students at The University of Akron. There were 466 male and 825 female participants. A large age range was obtained, with a minimum of 17 and a maximum of 65 years. The resulting mean age was 24.3 years with a corresponding standard deviation of 10.9 years. It

is worth noting that the age distribution was rather positively skewed, approximating the general population.

Measures

The Myers-Briggs Type Indicator is a 166 item forced-choice inventory. The score for each scale is calculated by summing the indicated responses for relevant items. Only 95 of the 166 items on the MBTI are scored—the purpose of the remaining 71 items is unexplained in the manual (Myers, 1962). The items were rationally produced such that endorsement of a particular response is counted toward the score for the corresponding trait. For example, the item, Which word in each pair appeals to you more? “116. A. sociable B. detached” corresponds to E and I, respectively.

Procedure

Subjects were instructed to complete the MBTI, recording responses on provided University of Akron computer sheets. Subjects were tested in small and large groups, and debriefing was made available at a later date. To verify the accuracy of the encoded data, a sampling of the stored data was compared item-by-item with raw responses as recorded.

Analysis

Subject responses to the 95 scored items were submitted to a principle components factor analysis. Both oblique and orthogonal rotations were performed in order to determine the best-fit solution. Inspection of the resulting factor matrix revealed a limited number of relevant factors. Using Cattell's (1966) scree test, six factors were retained for further consideration. Examination of the inter-factor correlation matrix from oblique rotation revealed little inter-factor correlation (r 's ranging from $-.11$ to $.19$). Accordingly, the orthogonal solution was used. This was of concern because the authors report inter-scale correlation coefficients as high as $r = .44$ (Myers, 1962).

Results

The six factors obtained account for 27.4% of the variance. Individually, Factor 1 accounted for 6.26% (eigenvalue = 8.76), Factor 2 accounted for 4.91% (eigenvalue = 5.66), Factor 3 account-

TABLE 1
Loadings of Scored MBTI Items for Six Resultant Factors

MBTI item	Factors					
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
1	-0.04237	0.54315*	-0.05521	0.01012	0.12933	0.01609
2	-0.17902	-0.23540	0.24413	-0.01042	-0.40285*	-0.00048
4	-0.38569*	0.06702	0.16014	-0.00169	0.14727	-0.04661
6	0.00335	0.03537	-0.06032	0.49725*	-0.03172	-0.01681
9	0.02892	0.42465*	-0.01972	0.10571	-0.00174	0.00590
11	0.03775	0.10696	0.10756	-0.03115	0.42903*	0.00205
13	0.04117	-0.45440*	-0.04933	0.12308	-0.08031	-0.00600
15	-0.16930	0.07274	-0.00252	0.46799*	-0.02886	0.01211
17	-0.12062	-0.04264	0.09996	0.01760	-0.38626*	0.06296
19	-0.06629	0.04960	0.01155	-0.47819*	-0.02416	0.44885*
20	0.07656	-0.27105	-0.11182	0.07694	-0.25460	0.01004
25	-0.08487	0.10028	-0.06236	0.54223*	-0.04261	0.07566
26	-0.30396	-0.14553	0.32006	-0.02300	0.06606	-0.05774
27	0.07216	0.46360*	0.02348	0.02093	0.06537	0.02110
29	-0.31386	-0.15630	0.26861	0.03499	-0.09025	-0.00732
33	-0.21656	0.06727	0.06963	0.37866*	0.16933	0.46474*
35	0.00534	0.49346*	-0.15691	0.00805	-0.03145	-0.08433
37	0.12238	0.09336	-0.03542	-0.05764	0.32721	-0.00413
41	0.01162	0.11623	0.02261	-0.33988	0.07567	0.01550
42	0.12021	-0.44686*	-0.04567	0.16264	-0.29933	0.02354
47	0.00524	-0.01225	-0.08466	0.17236	-0.11980	0.29626
49	-0.07984	0.49374*	0.11893	0.02131	0.12197	0.04587
50	-0.08035	0.03041	-0.02770	0.65701*	0.03616	-0.02985
53	-0.05816	0.00003	0.10161	-0.03812	0.36763*	0.02362
55	-0.01999	0.45986*	-0.04096	0.02329	0.13621	0.02868
58	0.10476	0.07523	0.16955	-0.23183	-0.09617	-0.00281
60	0.17201	0.54161*	0.10874	0.03889	-0.01133	-0.08687
64	0.10291	0.02437	0.27565	0.01062	0.36562*	0.43334*
66	0.07643	0.05286	0.23331	0.37456*	-0.13391	0.07224
68	0.16810	-0.35381*	0.15745	0.00080	-0.04005	-0.00478
70	0.22443	-0.00233	0.18992	-0.04021	-0.37610*	-0.00861
72	0.41839*	0.01855	-0.07316	-0.01863	-0.07727	0.54521*
73	-0.02128	0.00196	0.35376*	0.10795	-0.31022	0.04852
74	0.43490*	0.44786*	0.04129	-0.05336	0.10967	0.05332
76	0.24071	-0.15044	0.14750	0.08117	-0.49832*	0.06708
77	0.05491	-0.17154	0.18440	0.28778	0.14478	0.45094*
78	0.31796	0.09613	0.16598	-0.05213	0.36666*	0.03673
79	0.31354	0.07754	-0.16938	-0.03966	-0.17762	0.50693*
81	0.36119*	-0.05446	-0.00326	0.08817	-0.04179	0.12146
84	0.08169	0.02772	0.38888*	-0.02085	0.01041	0.00632
85	0.34376	0.62167*	0.05085	-0.00294	0.12352	0.02863
86	0.55290*	0.11307	-0.17761	0.03160	-0.10232	0.17005
87	0.37613*	0.08743	0.07056	-0.40940*	-0.05691	0.05251
88	0.20054	0.11994	0.16775	0.00115	0.33114	0.43102*
89	-0.05683	0.12719	0.46677*	0.02104	-0.01370	0.07847
90	0.48700*	-0.01342	0.07788	0.05483	0.27997	0.02045
91	-0.05504	0.10940	0.50887*	0.07666	-0.11122	-0.01533
92	0.07514	0.00194	0.19221	0.40428*	-0.14210	0.08409
93	0.03501	-0.04984	0.46282*	0.06847	0.05226	-0.01611
94	-0.02623	-0.36040*	0.34478	0.05311	-0.12850	0.46672*
95	0.21641	-0.02095	0.11350	0.38707*	0.04521	0.05798
97	0.28541	0.36109*	-0.03872	-0.00745	0.05625	0.44854*

TABLE 1 (Continued)

MBTI item	Factors					
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
98	0.36897*	0.16782	-0.01745	-0.05273	0.25953	0.45545*
99	0.05402	-0.19697	0.19049	0.06077	-0.29952	0.12026
100	0.25435	-0.01461	-0.08512	0.05392	-0.10465	0.53647*
102	0.52254*	0.12373	-0.03229	0.03267	0.29905	0.10700
103	-0.05595	0.00020	0.60950*	0.01266	0.16215	-0.04206
104	0.39875*	0.21167	0.03222	0.02293	0.43266*	0.505895
105	0.38900*	0.04804	-0.06232	0.08362	0.05469	0.08588
106	0.39719*	0.12739	0.14347	-0.39309*	0.00841	0.02663
107	0.47927*	-0.00930	0.14623	-0.01701	0.33655	0.09427
108	0.50967*	-0.08450	0.10007	-0.03030	-0.11335	0.13160
109	0.59178*	0.27291	-0.07156	0.05726	-0.07192	0.02541
111	-0.07583	0.05150	0.47965*	0.03211	-0.02847	0.01847
112	0.27002	0.19203	0.15419	-0.01680	0.19112	0.01450
113	0.16787	-0.24021	0.11976	0.09397	-0.11120	0.50522*
114	0.42579*	0.11035	-0.19786	-0.01255	-0.11225	0.48553*
115	0.46476*	-0.08206	0.12081	0.05612	-0.27231	0.09101
116	0.01794	0.10125	0.41115*	0.33738	0.01418	0.04200
117	0.38675*	0.03918	0.13542	0.02656	0.21966	0.07963
118	0.49900*	0.34572	-0.11684	0.01905	-0.06963	0.07326
119	0.39411*	0.18727	0.08325	0.02115	0.16248	0.04133
120	-0.00679	0.11364	0.48419*	0.01180	-0.09377	0.04146
121	0.14263	0.12686	0.40202*	-0.01068	0.28549	-0.01486
122	0.26541	-0.00849	0.38325*	-0.01430	0.27374	-0.01467
124	0.23359	0.31610	0.16054	0.13110	0.07194	0.06217
126	0.09307	0.01191	0.09243	0.46504*	-0.03406	0.40834*
128	0.17133	0.19330	0.14144	-0.05564	0.35615*	0.39581*
129	0.24214	-0.01384	0.25949	0.28711	0.02454	0.01585
132	0.20864	0.52298*	0.16448	0.07250	-0.01388	0.03794
133	0.53752*	-0.09038	0.08227	0.00771	0.03210	0.08274
134	0.15431	0.01599	0.18599	0.47131*	-0.11534	0.02813
138	0.19631	0.07788	0.21114	0.40294*	-0.11037	0.06850
140	0.08448	0.20017	0.24806	0.07267	-0.33605	0.04444
142	0.20177	0.39693*	0.17594	0.02844	0.02663	0.00186
145	0.26112	0.15722	0.19532	0.01021	0.42831*	0.07954
147	0.26059	-0.02544	0.28913	-0.00411	-0.08464	0.05830
148	0.19527	0.04018	0.06655	0.51960	-0.14092	0.03037
149	0.09132	-0.15343	0.19967	0.09124	-0.30287	0.04448
151	0.20460	0.39661*	0.17787	0.01402	0.06352	-0.02462
153	0.27962	-0.26488	0.19168	-0.09203	0.02799	-0.02770
154	0.03327	-0.09815	0.44341*	0.07825	-0.08621	0.01477
158	0.01915	-0.05668	0.36412*	-0.02791	-0.04317	0.43263*
160	0.04941	0.10162	0.28602	0.35885*	-0.04207	0.08472
165	0.06435	-0.01838	0.20172	0.12709	-0.33309	0.07754

ed for 4.50% (eigenvalue = 4.31), Factor 4 accounted for 4.05% (eigenvalue = 3.51), Factor 5 accounted for 3.86% (eigenvalue = 2.84), and Factor 6 accounted for 3.81% (eigenvalue = 2.31).

The resultant factor matrix for the six orthogonal factors is displayed in Table 1. Factor 1 yielded 21 factor loadings exceeding

TABLE 2
Number and Percent of MBTI Items With Loadings Greater Than .35 For Six Factors

Factor	MBTI Scale											
	EI			SN			TF			JP		
	<i>n</i>	%	<i>r</i>	<i>n</i>	%	<i>r</i>	<i>n</i>	%	<i>r</i>	<i>n</i>	%	<i>r</i>
1	2	9.5	.01	7	33.3	.25	9	42.9	.08	3	14.3	.20
2	—	—	.09	—	—	.12	—	—	-.02	17	100.0	.33
3	1	7.6	-.02	2	15.4	-.05	10	77.0	-.13	—	—	-.05
4	16	100.0	.40	—	—	.15	—	—	.02	—	—	.12
5	—	—	-.04	11	100.0	.13	—	—	.04	—	—	-.03
6	4	25.0	.3	4	25.0	.20	5	31.2	.05	3	18.8	.12

the criterion, absolute values greater than or equal to .35. Seventeen MBTI items produced factor loadings exceeding the criterion for Factor 2. For Factor 3, 13 items are noteworthy; for Factor 4, 16 items; for Factor 5, 11 items; and for Factor 6, 16 items stand out.

How these items are distributed by MBTI scale is shown in Table 2. Factors 1 and 6 are comprised of a heterogeneous grouping of items. Factors 2 through 5, however, each resemble one MBTI scale. Pearson correlations were computed to examine these factor-scale score relationships. The values of interest ranged from $-.13$ to $.40$ (see Table 2).

To further explore the inter-item relationships for each scale, Kuder-Richardson 20 (KR-20) coefficients were calculated for the four MBTI scales. The KR-20 coefficients were .595 (EI), .664 (SN), .628 (TFMale), .628 (TFFemale), and .531 (JP). For the six Factors, KR20 coefficients for the clusters formed by the indicated items ranged from .325 (Factor 5) to .825 (Factor 1). Coefficients for the other factors were .472 (Factor 2), .712 (Factor 3), .580 (Factor 4), and .804 (Factor 6).

Discussion

Factors 2 through 5 correspond to the JP, TF, EI and SN scales, respectively. The result supports the inclusion of the relevant items on the indicated scale. Although the items do tend to cluster in a way consistent with the structure of the MBTI, there is by no means a one-to-one correspondence, contrary to the findings of Tzeng et al. (1984). Seven items (29.2%) from the JP scale did not meet the criterion for Factor 2, 13 items (56.5%) for the TF scale failed to appear on Factor 3, six EI items (27.3%) did not load significantly on Factor 4, and 15 SN items (57.7%) did not yield significant loadings for Factor 5. This raises the question of the appropriateness of the

use of these items on the questionnaire. This question of item inclusion is reinforced by examination of the factor-scale score correlation matrix. The coefficients for the relationships between a factor resembling a particular scale and that scale's score are neither large nor exclusive. For example, the Factor 5 – SN correlation yields an $r = .13$. However, three of the remaining five correlations with SN produce larger coefficients although Factor 5 has the largest and purest overlap with SN scale items. This clearly suggests that the SN scale is not factorially pure. This problem also exists for the Factor 3-TF relationship, though this relationship is, admittedly, weaker because only 77% of the Factor 3 items are from the TF scale. The EI and JP relationships are less troublesome.

This problem is not simply a result of the factor structure. Examination of the KR-20 coefficients calculated on the salient items for each of the six factors (.325~.825) compare favorably with those coefficients calculated on the four MBTI scales (.531~.664)—neither range is satisfactory. The latter coefficients also differ considerably from those cited by Tzeng et al. (1984), i.e., .74~.85.

A second point of diversion from the results of Tzeng et al. (1984), and one inconsistent with the intended structure of the MBTI, is the appearance of factors 1 and 6. Factor 1 (eigenvalue = 8.76) appears to be a measure of unconditional positive regard with MBTI items such as “109. A. orderly B. easy-going,” “133. Do you feel it is a worse fault A. to show too much warmth B. not to have warmth enough,” and “86. A. convincing B. touching” loading very highly. Factor 6 (eigenvalue = 2.31), though not readily interpretable, is important in that it also does not correspond to any of the MBTI scales. It is apparent that the rationally contrived structure of the MBTI is inconsistent with the factor structure produced in this study.

The third noteworthy result is that there was no distinction within the resultant factor structure between items scored for males and those scored for females on the TF scale. This lack of discrimination suggests that the validity of separate TF scales for males and females is questionable and requires further investigation.

A fourth issue requiring inspection is the weighting of particular responses. Eighty-six responses (various items, various scales) are scored as 2-points whereas the remainder are scored at unity. Study of the factor matrix reveals that the items with double-weighting show no greater discriminability than the remaining items.

A fifth issue requiring examination is the inclusion of the 71 items on the questionnaire which are not scored. There is no rationale given in the manual for their appearance. If they were included for

test development, the 22 years of use since the inception of the current test form would certainly have provided ample opportunity for further item selection. As of this writing, there have been no studies of the affect that the item arrangement might have on respondent scores.

The findings of this study support the multi-scale structure of the MBTI. Recognizing the existence of two general factors (1 and 6) overlapping the four scales, four factors were obtained which correspond to the four MBTI scales, to a degree. However, because there was no where near a one-to-one correspondence between items from each of the MBTI scales and the salients appearing for the resulting factors, and because of the apparent lack of satisfactory internal consistency, it is advisable that the MBTI be retested for cross validation and be reanalyzed for revision. Future research should continue to address the nature of the MBTI at the item level (cf Tzeng et al., 1984) and not solely the summative score level, as has been predominant in the past.

REFERENCES

- Bradway, K. (1964). Jung's psychological types: Classification by test versus classification by self. *Journal of Analytical Psychology*, 9, 129-135.
- Buros, O. K. (Ed.) (1978). *The eighth mental measurements yearbook*. Lincoln: University of Nebraska.
- Carlyn, M. (1977). An assessment of the Myers-Briggs Type Indicator. *Journal of Personality Assessment*, 41, 461-473.
- Cohen, D., Cohen, M., and Cross, H. (1981). A construct validity study of the Myers-Briggs Type Indicator. *EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT*, 41, 883-891.
- Harrison, N. W. (1976). Validation of Jung's typological framework. *Dissertation Abstracts International*, 37(3-A), 1467.
- Myers, I. B. (1962). *The Myers-Briggs Type Indicator Manual*. Princeton, NJ: Educational Testing Service.
- Omizo, M. M. (1978). An evaluation of the Myers-Briggs Type Indicator. *Dissertation Abstracts International*, 39(1-A), 200-201.
- Stricker, L. J. and Ross, J. A. (1964). Some correlates of a Jungian personality inventory. *Psychological Reports*, 14, (Monograph Supplement 6).
- Tzeng, O. C. S., Outcalt, D., Boyer, S. L., Ware, R., and Landis, D. (1984). Item validity of the Myers-Briggs Type Indicator. *Journal of Personality Assessment*, 48, 255-256.
- Webb, S. C. (1964). An analysis of the scoring system of the Myers-Briggs Type Indicator. *EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT*, 24, 765-781.