

Western University

Scholarship@Western

Management and Organizational Studies
Publications

DAN Department of Management and
Organizational Studies

6-1-2012

Item format and social desirability: Implications for interpretation of the Myers-Briggs Type Indicator

Edward Helmes
James Cook University

Julie Aitken Schermer
Western University, jharris@uwo.ca

Maryann Fraboni
Nipissing University

Follow this and additional works at: <https://ir.lib.uwo.ca/mospub>

Citation of this paper:

Helmes, Edward; Schermer, Julie Aitken; and Fraboni, Maryann, "Item format and social desirability: Implications for interpretation of the Myers-Briggs Type Indicator" (2012). *Management and Organizational Studies Publications*. 89.

<https://ir.lib.uwo.ca/mospub/89>

Item Format and Social Desirability: Implications for Interpretation of the MBTI*

Edward Helmes
James Cook University,
Queensland, Australia

Julie Aitken Schermer
The University of Western Ontario,
Ontario, Canada

Maryann Fraboni
Nipissing University,
Ontario, Canada

The MBTI (Myers-Briggs Type Indicator) remains widely used in many counseling applications despite extensive criticism of its basic nature and psychometric properties. The present study was designed to examine specifically the accuracy of the claimed minimal influence of social desirability on Form G of the MBTI. Undergraduate students ($n = 26$) judged the desirability of each item option of Form G, which was compared across the 60 item pairs in which both options were scored. The rated values were approximately equal for two domains, while *J* and *E* item options were judged to be more desirable than their *P* and *I* paired response options. A second study ($n = 52$) evaluated the social desirability of the 16 MBTI type descriptors, finding most descriptions to be above the neutral range in desirability. These results suggest that stylistic responding contaminates MBTI profiles and interpretative material. Consequently, users should consider alternative measures and at the very least, take great care in interpreting the MBTI because of its flawed structure.

Keywords: social desirability, MBTI (Myers-Briggs Type Indicator), test construction, personality, personality assessment

The MBTI (Myers-Briggs Type Indicator) (Myers & McCaulley, 1985; Myers, McCaulley, Quenck, & Hammer, 1998) represents an interesting mix of theory and practice in its development, and its current popularity in some areas of psychological practice suggests that a large number of its users are pleased with it (Quenck, 2000). Such acceptance may in part be due to the wide range of possible applications. The MBTI manual (Myers, McCaulley, Quenck, & Hammer, 1998) suggests many applications in career counseling and organizational development, despite there being little actual empirical evidence of validity for decisions affecting individuals in such uses and repeated published cautions regarding its use. For example, the MBTI has been repeatedly criticized on fundamental psychometric grounds (Stricker & Ross, 1962, 1964; McCaulley, 1991; Merenda, 1991; Boyle, 1995; Pittenger, 1993). Some criticisms attack the basic rationale for the structure of the MBTI, which in turn calls into question of the validity of the most basic interpretations of the scale. An early example of findings critical of the structure of the MBTI was provided by Lorr (1991), who attempted to

* Portions of these results were presented at the 26th International Congress of Psychology, 1996, Montréal and the 2006 joint conference of the Australian Psychological Society and the New Zealand Psychological Society, Auckland.

Edward Helmes, Ph.D., professor, Department of Psychology, James Cook University.

Julie Aitken Schermer, Ph.D., DAN Program in Management & Organizational Studies, The University of Western Ontario.

Maryann Fraboni, Ph.D., Department of Psychology, Nipissing University.

replicate the 16 presumed types using cluster analysis with two samples of male clerks and two methods of clustering. While there was reasonable convergence between the two methods of clustering, neither did a set of results show notable agreements with the standard MBTI types. Pittenger (1993) expressed related concerns, noting that the test could not empirically support many of the interpretations of the MBTI by its users. Research and comments along these lines are consistent with the recommendation expressed by Hunsley, Lee, and Wood (2003) about the MBTI that, "Psychologists are advised to rely on personality and vocational interest tests that have a sounder empirical basis" (p. 64).

While Hunsley et al. (2003) regarded the MBTI as a personality inventory; Myers et al. (1998) preferred to describe it as a measure of preference for personal styles, even though this distinction may not be one that is accepted by all and many do feel that it can be accurately described as a personality measure. Indeed, the distinction between a personality trait and a personal style is not obvious and Myers et al. did not provide a clear definition or distinction. Such differences in perspective, with MBTI advocates apparently feeling that criticisms can be dismissed by redefinition of terms, are not rare in the literature on the MBTI. The study reported here does not take a stand as to whether the MBTI measures traits or preferences in order to explore aspects of the MBTI items that may influence the interpretation of its scores.

It is likely that the firm typological beliefs of Katherine Cook Briggs and her daughter, Isabel Briggs Myers, were a major influence on forming the original definition of the types that has dominated the development of the MBTI, despite the lack of empirical evidence for the existence of those types (Myers et al., 1998). Indeed, even the claim to be based upon Jung's (1921) theories of personality types may not be fully justified (Barbuto, 1997; Merenda, 1991, p. 197). Barbuto and Merenda made several points about the conceptual weaknesses of the MBTI, one of the most relevant being that Jungian types vary in their accessibility to conscious awareness and therefore the claim that a conscious preference for a behavioral style is consistent with Jungian theory is misleading, if not erroneous.

Regardless of the degree to which the MBTI types reflect Jung's (1921) views on personality, the MBTI's collection of four dichotomous preference scales is intended to reflect "basic dispositions". This claim is made with little reference to other major personality or developmental theories to support the assertion of their supposed fundamental nature. The scale is intended "not to measure people or the traits they are said to 'have' or possess, but rather to sort people into groups to which, in theory, they already belong" (Myers et al. 1998, p. 127). Despite the claimed primacy of the measure as reflecting fundamental aspects of personality, the manual and most literature on the MBTI does not refer to other theories of personality or to other measures of personality (Carlyn, 1977; Carlson, 1985), although the relevant research to link the MBTI to other personality models does exist (Francis, Craig, & Robbins, 2007).

The MBTI item format is based on pairs of items (or in some cases in Form G, three items), from which respondents are to select the one member of the pair that is most descriptive of them. From a psychometric perspective, the MBTI is unusual in that it uses continuous scales to assess what are presumed to be underlying categorically discrete types. The continuous scores on the eight constituent scales that form the four bipolar scales (*EI* (extroversion-introversion), *SN* (sensing-intuition), *TF* (thinking-feeling), and *JP* (judging-perceiving)) are converted to categorical group membership by using the most extreme score from pairs of scales to define the dimension and the difference between the sum of scale scores to define the strength of the preference. For example, if the sum of keyed item responses for *E* (extroversion) is 17 and the sum for *I*

(introversion) is 19, then the resulting type would be *I* and the preference for *EI* would be *I2*. This process is repeated for the other three pairs of scales. The resulting four categorical types are then combined to create the 16-member typology of the MBTI.

The forced-choice response format of the MBTI should theoretically provide some advantages for the measure if it can thereby avoid undue influences of such stylistic characteristics as acquiescence, or yea-saying, and social desirability (Paulhus, 2007). The forced-choice format can only control for the influence of such factors, however, if the two options are roughly equivalent in their susceptibility to them. The MBTI manual claims that social desirability is controlled by means of item selection on the basis of "prediction ratios" and item weights. Prediction ratios are basically the proportion of the responses to the particular item by the keyed type divided by the sum of item option endorsements by that keyed group and the bipolar opposite group. This ratio is used to determine scoring weights for the MBTI items throughout the development of the various forms of the MBTI. In actuality of course, this practice confuses endorsement rate with social desirability and does not actually control the potential confound. The introduction of weights in scoring is intended to balance the popularity between the members of the item pairs, but the latest manual (Myers et al., 1998) does not present any data to support the claim of such balancing, nor does it refer to any of the substantial literature on social desirability in personality assessment. Girelli and Stake (1993) also made the point that item weighting was unlikely to be successful in removing the influence of social desirability.

The unique MBTI scoring method is rationalized with the argument that the continuous scores reflect the degree of preference for a particular categorical type. Doing so in the manner used by the MBTI assumes that the two constructs in question in fact are related to each other in order to justify the mathematical operation of subtraction. It makes no sense to calculate the difference between two scores if those scores are not related. In order to test this assumption of the MBTI structure, Girelli and Stake (1993) used Likert-scale versions of the separate members of the MBTI item pairs and found that the Likert-format *SN* and *TF* scales did not correlate at all, while support for the bipolarity of the introversion-extroversion pair was weak. They argued that the forced-choice format imposes the bipolarity on the scales, but the assumed underlying structural bipolarity is not supported when the component scales are separated. Tzeng, Ware, and Chen (1989) also broke the item pairs into their constituent separate items and performed a factor analysis with varimax rotation on the resulting set of items. They found that 90% of items loaded on scales corresponding to the underlying constructs of *E*, *I*, *S*, *N*, etc.. While such an analysis does not address the basic issue of the bipolarity of the scales, these studies clearly reflect the different views of the MBTI in the literature, much of which challenges the basic conceptual model of the MBTI.

Even if there were support for the bipolar relationship of the MBTI scales, bias might be still introduced if the two elements of the actual item pairs were not matched in terms of their basic attraction, or social desirability. The intent of the first study was to obtain ratings of social desirability of the individual members of the item pairs of the MBTI Form G. This would allow the determination of the actual balance in desirability between the members of the item pairs. Social desirability can confound the interpretation of a scale, and thus complicate its use in practice. The authors also explore the extent of confounding of the material in MBTI interpretive materials with social desirability in a second study to determine how desirability may be operated at that level. The issue of equating the item options has not been addressed in previous research into the

structure of the MBTI and findings might help clarify some of the discrepancies in the literature on the scale.

Analyses of the Two Studies

Study 1

Participants. The raters were 26 university undergraduates (10 males), with a mean age of 31.3 years old ($SD = 12.98$, range from 17 to 56 years old, median 23.5 years old). Participants completed the ratings of the 262 items in exchange for partial course credit. Edwards (1970) reported high levels of agreement among those asked to make such judgements, such that fairly small samples provide highly consistent social desirability scale values that generalize well to other samples and settings.

Procedure. The items of the three sections of Form G were separated into their constituent single statement or single adjective as appropriate. Desirability was judged on a 9-point scale, using the following instructions, adapted from Helmes, Reed, and Jackson (1977), "Please read the following questions and rate each one as to how favorable or desirable it would be for a person to say 'Yes' to the question", where 1 = "Extremely undesirable", through 5 = "Neither desirable nor undesirable", to 9 = "Extremely desirable". Labels for the values of 4 and 6 were simply "Undesirable" and "Desirable", while other intermediate points were labeled as "Very" and "Moderately", such that all rating points had a corresponding verbal label.

The scoring templates for Form G were analyzed to determine the scoring pattern for the bipolar items. Of the 126 item pairs, only 94 are actually scored on Form G. Of these 94 item pairs, 60 items have both elements of the pair scored; for 32 items, only one element of the pair is actually scored. One item has two options that are keyed for the same preference with different weights, while the other item has all three options scored for three different preferences. These latter two items were not included in the analysis. Only when both options of a pair are scored for the bipolar opposites can the issue of differential desirability of the domains be evaluated.

Results. Ratings were obtained for a total of 262 statements and adjectives from Form G of the MBTI¹. The overall mean for social desirability was 6.1 on the nine-point scale ($SD = 0.97$). Table 1 reports the outcome of 60 paired *t*-tests for the four bipolar preferences of the MBTI for judged social desirability. There were 17 pairs for the *JP* domain, 14 item pairs for the *SI* domain, 17 item pairs for the *EI* domain, and 12 pairs for the *TF* domain. Part 2 of Form G (word pairs) provided most of the item pairs for the *SN* and *TF* domains, while the *JP* and *EI* preference domains were more equally spread across all three sections of the MBTI.

Table 1

Comparisons of Item Pair Members (A vs. B) for Four Domains of the MBTI

MBTI preference pair	Percent of items with A > B in rated desirability	Percent of items with A = B in rated desirability	Percent of items with A < B in rated desirability
<i>JP</i>	52% ($n = 9$)	47% ($n = 8$)	0
<i>SI</i>	14% ($n = 2$)	86% ($n = 12$)	0
<i>EI</i>	59% ($n = 10$)	35% ($n = 6$)	6% ($n = 1$)
<i>TF</i>	8% ($n = 1$)	84% ($n = 10$)	8% ($n = 1$)

The paired items were roughly equivalent for social desirability for the *TF* and *SN* domains, while the *J* (judging) and *E* (extroversion) areas were more desirable than their *P* (perceiving) and *I* (introversion) opposites. Interestingly, the latter preference choices are also those to be used when tied scores occur (Myers

¹ Copies of the item means and standard deviations are available from Edward Helmes.

et al., 1998).

Discussion. The study provided basic information on social desirability ratings for the constituent items for Form G of the MBTI. The results suggest that the item pairs of the *JP* and *EI* domains are not balanced in terms of social desirability and that interpretations of these scales may be confounded. Such confounding means that interpretations of half the MBTI may be questionable. Confounded interpretations in turn suggest that the MBTI should not be used for decision-making for individuals. Enough evidence exists, including that from this study, to support Hunsley et al.'s (2003) conclusions that measures with better evidence for validity than the MBTI be used for such purposes.

In many applications of the test, the information provided to clients is not based on actual responses to the items themselves, but on the descriptions of the 16 types that are assumed to exist by the developers of the MBTI and underlie the logic of the preference scales. Such descriptions would of course figure heavily in many organizational applications of the MBTI. If these descriptions are confounded in terms of social desirability, it is so plausible that acceptance of the validity and accuracy of the MBTI type descriptors by the respondent is based as much on their desirability as on their overt content. Interpretations of the accuracy of the MBTI for individuals based upon such confounded descriptive material would be of questionable validity at best.

Therefore, the brief descriptions used to describe each of the 16 types were rated in Study 2 for social desirability. This study thus extends the first into an examination of the broader association of social desirability with the MBTI.

Study 2

Participants and procedure. Fifty-two undergraduate university students (15 males, 36 females, one did not report gender; mean age = 20.5 years old, $SD = 4.95$, range from 18 to 51 years old, median 19 years old) judged the desirability of each description of the 16 types of the MBTI on a 1-7 scale, with 1 = "Extremely undesirable", through 4 = "Neither desirable nor undesirable", to 7 = "Extremely desirable". Intermediate values used the adjectives "Moderately" and "Very". The type descriptions were taken from the summaries of type provided with many of the interpretive materials for the MBTI (see Table 4.25 in Myers et al., 1998, p. 64). The versions used were those current for the second edition of the MBTI manual, which are appropriate for Form G (Myers & McCaulley, 1985). The latest versions differ in some parts of the descriptions from those used in this study, but have the majority of the descriptive statements in common.

Results. The average judged desirability ratings of the 16 types are presented in Table 2 and suggest that all of the types were rated as above the mid-point (3.5 numerically, neither "Desirable" nor "Undesirable") of the scale in desirability. The mean rating was 4.9 ($SD = 0.58$) on a seven-point scale, and only the *ISFJ* type approached the neutral point in judged social desirability. In particular, the four most desirable type descriptions with ratings of 5.5 out of seven or higher are *ENTJ* (extraverted, judging, and intuitive with thinking), *INFP* (introverted, perceptive, and intuitive with feeling), *INTJ* (introverted, judging, and intuitive with thinking), and *ESTP* (extraverted, perceiving, and intuitive with sensing).

Table 2 also reports the percentage of the MBTI respondents who fall into each of the 16 types based upon the national representative sample figures (see Table 14.1 in Myers et al., 1998, p. 379). These figures are based on a US national sample of over 3,000 adults. The correlation with the norm frequency of occurrence for judged social desirability was 0.62.

Table 2

Frequency of Occurrence and Judged Social Desirability for 16 MBTI Types

Type	Percentage in norm sample (%) (rank)	Judged mean desirability* (SD)
<i>ISTJ</i>	11.6 (3)	4.6 (1.1)
<i>ISFJ</i>	13.8 (1)	3.6 (1.2)
<i>ISTP</i>	5.4 (8)	4.7 (1.0)
<i>ISFP</i>	8.8 (4)	4.6 (1.0)
<i>ESTP</i>	4.3 (10)	5.5 (1.2)
<i>ESFP</i>	8.5 (6)	4.5 (1.2)
<i>ESTJ</i>	8.7 (5)	5.1 (1.0)
<i>ESFJ</i>	12.3 (2)	4.8 (0.9)
<i>INFJ</i>	1.5 (16)	4.6 (0.8)
<i>INTJ</i>	2.1 (14)	5.5 (1.1)
<i>INFP</i>	4.4 (9)	5.8 (1.1)
<i>INTP</i>	3.3 (11)	4.7 (1.0)
<i>ENFP</i>	8.1 (7)	5.0 (1.1)
<i>ENTP</i>	3.2 (12)	5.4 (1.0)
<i>ENFJ</i>	2.5 (13)	4.8 (1.1)
<i>ENTJ</i>	1.8 (15)	5.9 (0.9)

Notes. *E*: Extroversion; *F*: Feeling; *I*: Introversion; *J*: Judging; *N*: Intuition; *P*: Perceiving; *S*: Sensing; *T*: Thinking; * 1 = "Extremely undesirable" to 7 = "Extremely desirable".

Discussion. The more desirable types from the MBTI reported here differ from the "Fake good" profile found by Furnham (1990a; 1990b) who reported that type *ESTJ* (extraverted, judging, and sensing with thinking) was the "ideal fake good" type. We found the *ESTJ* type had moderate ratings for social desirability, ranking in the middle of the distribution, but *ESTP* was judged to be one of the most desirable types and differed from *ESTJ* in only one element. This contrast in results between "Fake good" and "Desirable" targets suggests that social desirability judgments may differ from asking individuals to consciously distort their responses to a given measure (Paulhus, 2007; Streiner & Norman, 2008).

General Discussion

The claim that social desirability was controlled on the MBTI by its developers is not supported by the results of Study 1, which found that two of the four bipolar scales were unbalanced for social desirability across the constituent item pairs and that most item options were above the neutral point in terms of judged social desirability. The majority of the MBTI items were also rated as being above the neutral point in social desirability and at least two of the preference pairs were badly matched on social desirability. The confounding of content with social desirability extended to the MBTI interpretive material. Study 2 found that all of the types as commonly described were rated as above the neutral point in desirability.

It appears that the authors of the MBTI have confounded test content with desirability through not linking their work to the broader literature in personality and personality assessment. Their claims of control by interpreting popularity as desirability are mistaken and deal with two characteristics that were identified as conceptually and empirically distinct long ago (Jackson & Messick, 1969). The high susceptibility of the MBTI to the association with social desirability also likely contributes to the ease with which responses to the test can be deliberately distorted (Furnham, 1990a; 1990b). In addition, that influence also likely leads to mistaken

attributions of accuracy and validity by respondents, with many likely to accept their descriptive type in much the same way that respondents accept “Barnum” personality descriptions.

Study 1 showed that the pairings of MBTI items were not matched for social desirability and that common outcomes were associated with differences in item desirability. This imbalance is present primarily for the *JP* and *EI* domains. This finding suggests social desirability is one source of uncontrolled error variance on the MBTI and that some scales are more susceptible to this influence than others. Basic methods of controlling for social desirability, such as Jackson’s (1970) DRI (differential reliability index), were not used during the development of the MBTI, despite the implications for confounded interpretations and the recommendations for the use of a method such as the DRI (Holden & Passey, 2009; Streiner & Norman, 2008). Study 2 demonstrated that the descriptors of the 16 types in the interpretive material was also confounded with social desirability ratings, once again suggesting that the developers of the MBTI paid insufficient attention to such a source of irrelevant variance during its development.

Limitations

One limitation of this study is that it did not use the most current version of the MBTI. This research is based upon Form G, while the current form is M (Myers et al., 1998), although materials for Form G remain for sale and remain in use. New item analyses were conducted for Form M and items with three response options were converted back to two options during the revision of items for Form M, while retaining most of the items from Form G unchanged. The analyses for the revision included prediction ratios as for previous forms, factor analyses, and analyses based upon item response theory. The latter analyses did, unfortunately, in some ways, continue to score the pairs of items as one response. It does not appear to have been any analysis on the individual statement or adjective that form the paired elements in the development of Form M. This process appears to be a continuation of the underlying assumption of the reality of the four bipolar preferences, despite the predominantly negative empirical evidence on this point. These data best represent Canadian undergraduates, but the evidence on the generalizability of judgements of social desirability suggests that the conclusions would hold for other English-speaking countries.

Conclusions

The net conclusion concerning the MBTI is a case of theory driving results, a process not consistent with scientific progress. In addition, current and prospective users need to become more aware of the serious limitations of the test. The MBTI is flawed at many levels, conceptually and empirically. It confounds response styles content with little recognition so that such confounding is problematic. This confound extends to interpretative material for the test, which can lead to erroneous interpretations by users of the test who receive such materials. Some may argue that the use of the MBTI for some applications, such as for “individual development” as part of a larger program of staff development, would still be appropriate even though its validity for many other purposes was unsupportive. The nature of the criticisms of the MBTI is such that we feel that no use of the MBTI is without serious risk of erroneous interpretations. The authors, therefore, agree with Hunsley et al. (2003) that the use of an alternative measure would be preferable. At best, the authors feel that users of the MBTI need to evaluate the cost of errors caused by the use of the test, costs to both the administrator and the client.

References

- Barbuto, J. E. (1997). A critique of the Myers-Briggs Type Indicator and its operationalization of Carl Jung's psychological types. *Psychological Reports, 80*, 611-625.
- Boyle, G. J. (1995). Myers-Briggs Type Indicator (MBTI): Some psychometric limitations. *Australian Psychologist, 30*, 71-74.
- Carlson, J. G. (1985). Recent assessments of the Myers-Briggs Type Indicator. *Journal of Personality Assessment, 49*, 356-365.
- Carlyn, M. (1977). An assessment of the Myers-Briggs Type Indicator. *Journal of Personality Assessment, 41*, 461-473.
- Edwards, A. E. (1970). *The measurement of personality traits by scales and inventories*. New York: Holt, Rinehart & Wilson.
- Francis, L. J., Craig, C. L., & Robbins, M. (2007). The relationship between psychological type and the three major dimensions of personality. *Current Psychology: Research and Reviews, 25*, 257-271.
- Furnham, A. (1990a). Faking personality questionnaires: Fabricating different profiles for different purposes. *Current Psychology: Research and Reviews, 9*, 46-55.
- Furnham, A. (1990b). The fake ability of the 16 PF, Myers-Briggs and Firo-B personality measures. *Personality and Individual Differences, 11*, 711-716.
- Girelli, S. A. & Stake, J. E. (1993). Bipolarity in Jungian type theory and the Myers-Briggs Type Indicator. *Journal of Personality Assessment, 60*, 290-301.
- Helmes, E., Reed, P. L., & Jackson, D. N. (1977). Desirability and frequency scale values and endorsement proportions for items of personality research Form-E. *Psychological Reports, 41*, 435-444.
- Holden R. R., & Passey, J. (2009). Social desirability. In M. R. Leary, & R. H. Hoyle (Eds.), *Handbook of individual differences in social behaviour* (pp. 441-454). New York: Guilford.
- Hunsley, J., Lee, C. M., & Wood, J. M. (2003). Controversial and questionable assessment techniques. In S. O. Lilienfeld, S. J. Lynn, & J. M. Lohr (Eds.), *Science and pseudoscience in clinical psychology* (pp. 39-76). New York: Guilford.
- Jackson, D. N. (1970). A sequential system for personality scale development. In C. D. Spielberger (Ed.), *Current topics in clinical and community psychology* (Vol. 2, pp. 61-96). New York: Academic Press.
- Jackson, D. N., & Messick, S. (1969). A distinction between judgments of frequency and of desirability as determinants of response. *Educational and Psychological Measurement, 29*, 273-293.
- Jung, C. G. (1921). *Psychological types*. New York: Harcourt, Brace.
- Lorr, M. (1991). An empirical evaluation of the MBTI typology. *Personality and Individual Differences, 12*, 1141-1145.
- McCaulley, M. H. (1991). Additional comments regarding the MBTI: A response to comments. *Measurement and Evaluation in Counseling and Development, 23*, 182-185.
- Merenda, P. F. (1991). Additional comments regarding the MBTI. *Measurement and Development in Counseling and Development, 23*, 179-181.
- Myers, I. B., & McCaulley, M. H. (1985). *A guide to the development and use of the Myers-Briggs Type Indicator*. Palo Alto, C. A.: Consulting Psychologists Press.
- Myers, I. B., McCaulley, M. H., Quenk, N. L., & Hammer, A. L. (1998). *MBTI manual: A guide to the development and use of the Myers-Briggs Type Indicator* (3rd ed.). Palo Alto, C. A.: Consulting Psychologists Press.
- Paulhus, D. L. (2007). The self-report method. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 224-239). New York: Guilford.
- Pittenger, D. J. (1993). The utility of the MBTI. *Review of Educational Research, 63*, 467-488.
- Pittenger, D. J. (2005). Cautionary comments regarding the Myers-Briggs Type Indicator. *Consulting Psychology Journal: Practice and Research, 57*, 210-221.
- Quenck, N. L. (2000). *Essentials of Myers-Briggs Type Indicator assessment*. New York: Wiley.
- Streiner, D. L., & Norman, G. R. (2008). *Health measurement scales: A practical guide to their development and use* (4th ed.). Oxford: Oxford University Press.
- Stricker, L. J. & Ross, J. (1962). *A description and evaluation of the MBTI* (Research Bulletin No. RB-62-6). Princeton, N. J.: Educational Testing Service.
- Stricker, L. J., & Ross, J. (1964). An assessment of some structural properties of the Jungian personality typology. *Journal of Abnormal and Social Psychology, 68*, 62-71.
- Tzeng, O. C. S., Ware, R., & Chen, J. M. (1989). Measurement and utility of continuous unipolar ratings for the Myers-Briggs Type Indicator. *Journal of Personality Assessment, 53*, 727-738.